

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ALGORITHMES EFFICACES D'ALLOCATION DE RESSOURCES POUR LES FUTURS RÉSEAUX MOBILES SANS FIL
À BANDES MULTIPLES

THÈSE
PRÉSENTÉE
COMME EXIGENCE PARTIELLE
DU DOCTORAT EN INFORMATIQUE

PAR
CIRINE CHAIEB

OCTOBRE 2024

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.12-2023). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je tiens tout d'abord à exprimer ma profonde gratitude envers moi-même pour la patience, l'engagement et la persévérance que j'ai investis dans mon parcours universitaire.

À mon directeur de recherche Professeur Wessam Ajib et à ma codirectrice professeure Fatma Abdelkefi, je suis reconnaissante pour leur excellent encadrement, leurs conseils et leur précieux soutien continu tout au long de mes travaux de recherche.

Un grand merci à mes chers parents, à mon mari et à toute ma famille pour leurs encouragements, leur présence et leur amour.

Enfin, un grand merci à toute l'équipe du laboratoire TRIME et à tous mes amis, en particulier à mon amie Yasmina, pour leurs conseils, leur soutien et les moments de joie partagés !

TABLE DES MATIÈRES

TABLE DES FIGURES	vi
LISTE DES TABLEAUX	viii
LISTE DES ABRÉVIATIONS, SIGNES ET ACRONYMES	ix
RÉSUMÉ	xiii
CHAPITRE 1 INTRODUCTION GÉNÉRALE	1
1.1 Mise en contexte et motivations	1
1.2 Définitions pertinentes	8
1.3 Problématiques	9
1.3.1 Problématique générale	9
1.3.2 Problématiques spécifiques	10
1.4 Contributions et méthodologie	11
1.4.1 Maximisation du nombre d'utilisateurs associés dans un réseau à bandes multiples	11
1.4.2 Techniques d'accès multiple dans un réseau à bandes multiples	12
1.4.3 Efficacité énergétique dans un réseau à bandes multiples	13
1.5 Organisation	14
1.6 Conclusion	15
CHAPITRE 2 MAXIMISATION DU NOMBRE D'UTILISATEURS ASSOCIÉS DANS UN RÉSEAU À BANDES MULTIPLES	16
2.1 Résumé	16
2.2 Introduction	17
2.2.1 Revue de littérature	17
2.2.2 Contributions	20
2.2.3 Organisation	21

2.3	Modèle du système	21
2.4	Formulation du problème	25
2.5	NP-difficulté	26
2.6	Solutions algorithmiques	27
2.6.1	Un algorithme centralisé (C-GUA)	27
2.6.2	Des algorithmes distribués (D-NBS et D-RLA)	29
2.7	Résultats de simulation	33
2.8	Conclusions	39
CHAPITRE 3 TECHNIQUES D'ACCÈS MULTIPLE DANS UN RÉSEAU À BANDES MULTIPLES		40
3.1	Résumé	40
3.2	Introduction	41
3.2.1	Revue de littérature	42
3.2.2	Contributions	44
3.2.3	Organisation	45
3.3	Modèle du système	45
3.4	Formulation du problème	49
3.5	NP-difficulté	50
3.6	Solutions algorithmiques	51
3.6.1	Un algorithme glouton (GRA)	51
3.6.2	Une métaheuristique (GA)	53
3.6.3	Un algorithme d'allocation de puissance (WFA)	54
3.6.4	Solutions d'apprentissage par renforcement profond (DRLA et MDRLA)	58
3.7	Résultats de simulation	61

3.7.1	Ajustement des hyperparamètres du réseau de neurones et convergence.....	62
3.7.2	Comparaison des performances des algorithmes proposés.....	66
3.8	Conclusions	69
CHAPITRE 4 EFFICACITÉ ÉNERGÉTIQUE DANS UN RÉSEAU À BANDES MULTIPLES		70
4.1	Résumé.....	70
4.2	Introduction.....	71
4.2.1	Revue de littérature.....	71
4.2.2	Contributions.....	73
4.2.3	Organisation	74
4.3	Modèle du système	75
4.4	Formulation du problème	77
4.5	NP-difficulté	78
4.6	Solutions algorithmiques	79
4.6.1	Solution itérative (IRAA)	79
4.6.2	Algorithmes gloutons (CGRA et DGRA)	83
4.6.3	Solution d'apprentissage par renforcement profond (MADRA).....	84
4.7	Résultats de simulation	85
4.8	Conclusions	95
CHAPITRE 5 CONCLUSIONS ET TRAVAUX FUTURS		96
5.1	Conclusions	96
5.2	Travaux futurs	98
BIBLIOGRAPHIE		100

TABLE DES FIGURES

Figure 1.1	Trafic mondial de données mobiles en exaotets (Ericsson, 2022).	2
Figure 1.2	Nombre d'appareils mobiles par technologie en milliards (Ericsson, 2022).	3
Figure 1.3	Comparaison des performances entre la 5G (IMT-2020) et la 6G (Vaezi <i>et al.</i> , 2022).	3
Figure 2.1	Une illustration d'un réseau hétérogène à bandes multiples.	23
Figure 2.2	Pourcentage d'utilisateurs associés en fonction du nombre de SBSs.	35
Figure 2.3	Pourcentage d'utilisateurs associés et de créneaux temporels utilisés en fonction du nombre de créneaux temporels disponibles pour chaque station de base.	35
Figure 2.4	Pourcentage d'utilisateurs associés en fonction du nombre d'utilisateurs.	36
Figure 2.5	Convergence de D-RLA.	37
Figure 2.6	Pourcentage d'utilisateurs associés en fonction du nombre d'utilisateurs.	38
Figure 2.7	Pourcentage de créneaux temporels utilisés en fonction du nombre d'utilisateurs.	38
Figure 3.1	Une illustration d'un réseau à bandes multiples utilisant les deux techniques OMA et MSC-NOMA.	46
Figure 3.2	Architecture du modèle d'apprentissage.	63
Figure 3.3	L'impact du taux d'apprentissage sur les résultats de l'algorithme DRLA ($L = 3$, $N_1 = 512$, $N_2 = 256$ et $N_3 = 64$).	63
Figure 3.4	L'impact du nombre de couches cachées sur les résultats de l'algorithme DRLA ($\alpha = 0.009$).	64
Figure 3.5	L'impact du nombre de neurones sur les résultats de l'algorithme DRLA ($\alpha = 0.009$, $U = 10$, $B_s = B_m = 1$ et $C_s = C_m = 2$).	65
Figure 3.6	Convergence de GA ($U = 10$, $B_s = B_m = 1$, $C_s = C_m = 2$ et $D_u^{max}/D_u^{min} = 20$).	65

Figure 3.7	Comparaison des performances des algorithmes proposés en termes de débit de données ($U = 10, B_s = B_m = 1$ et $C_s = C_m = 2$).	66
Figure 3.8	Comparaison des performances des algorithmes proposés en termes de nombre d'utilisateurs associés ($U = 10, B_m = B_s = 1$ et $C_m = C_s = 2$).	67
Figure 3.9	Indice de Jain en fonction de D_u^{max}/D_u^{min} ($U = 10, B_s = B_m = 1$ et $C_s = C_m = 2$).	68
Figure 4.1	Un exemple du modèle du système ($B_s = 1, B_m = 2$ et $B_z = 4$).	75
Figure 4.2	Positionnement des stations de base.	87
Figure 4.3	Efficacité énergétique ($\log_{10}(\eta)$) de CGRA pour différents scénarios.	88
Figure 4.4	Efficacité énergétique ($\log_{10}(\eta)$) de CGRA pour différentes exigences de qualité de service.	89
Figure 4.5	Nombre d'utilisateurs associés de CGRA pour différents scénarios.	89
Figure 4.6	Nombre d'utilisateurs associés de CGRA pour différentes exigences de qualité de service.	90
Figure 4.7	Puissance consommée de CGRA pour différents scénarios.	90
Figure 4.8	Puissance utilisée de CGRA pour différentes exigences de qualité de service.	91
Figure 4.9	Convergence de MADRA – Fonction de récompense individuelle.	92
Figure 4.10	Comparaison de l'efficacité énergétique pour MADRA.	93
Figure 4.11	Comparaison de l'efficacité énergétique des différents algorithmes.	93
Figure 4.12	Comparaison de nombre d'associations des différents algorithmes.	94
Figure 4.13	Comparaison de la consommation de puissance des différents algorithmes.	94
Figure 4.14	Comparaison de l'efficacité énergétique de IRAA et S-IRAA	95

LISTE DES TABLEAUX

Table 2.1	Notations et paramètres de simulation.....	22
Table 2.2	Complexité asymptotique des algorithmes.....	33
Table 3.1	Principaux travaux de recherche.....	44
Table 3.2	Notations et paramètres de simulation.....	62
Table 3.3	Impact de la mobilité sur les performances des algorithmes proposés.	69
Table 4.1	Notations et paramètres de simulation.....	87

LISTE DES ABRÉVIATIONS, SIGNES ET ACRONYMES

$\sqrt{\cdot}$ Racine carrée.

\sum Symbole somme.

$\binom{n}{k}$ Coefficient binomial.

\cup Union.

\cap Intersection.

\forall Pour tout.

\in Appartient à.

\mathcal{O} La notation grand O.

max Maximum de.

min Minimum de.

arg max Argument du maximum.

arg min Argument du minimum.

\times Produit cartésien.

\log_{10} Logarithme décimal.

\log_2 Logarithme binaire.

$\lceil x \rceil$ Partie entière supérieure de x .

\mathbf{A}_i i -ème ligne de la matrice \mathbf{A} .

\mathbf{A}^i i -ème colonne de la matrice \mathbf{A} .

$|\mathbf{X}^m|$ Somme des éléments du vecteur \mathbf{X}^m .

A Lettre en gras majuscule pour désigner une matrice.

a Lettre minuscule soulignée pour désigner un vecteur.

$\overline{\mathcal{V}}$ Cardinalité de l'ensemble \mathcal{V} .

\neq Différent de.

\subseteq Inclus dans ou égal à.

\approx Approximativement.

∞ Infini.

\emptyset Ensemble vide.

$[x]^+$ Retourne x si positif ou zéro.

p.ex. Par exemple.

etc. Et cetera.

resp. Respectivement.

FWA *Fixed wireless access.*

5G Cinquième génération.

6G Sixième génération.

IMT *International Mobile Telecommunications.*

3GPP *Third Generation Partnership Project.*

LTE *Long-term evolution.*

OMA *Orthogonal multiple access.*

NOMA *Non-orthogonal multiple access.*

MHz Mégahertz.

GHz Gigahertz.

THz Térahertz.

dB Décibel.

Mbps Mégabit par seconde.

mmWave *Millimeter-wave.*

Sub-6 GHz Fréquences inférieures à 6 GHz.

MIMO *Multiple-input, multiple-output.*

FDMA *Frequency division multiple access.*

TDMA *Time division multiple access.*

MSC-NOMA *Multiple sub-channel non-orthogonal multiple access.*

QL *Q-learning.*

DQL *Deep Q-learning.*

P *Temps polynomial (polynomial time).*

NP *Non déterministe en temps polynomial (non-deterministic polynomial time).*

AMPL *A mathematical programming modeling language.*

IEEE *Institute of Electrical and Electronics Engineers.*

SINR *Signal to interference plus noise ratio.*

SNR *Signal to noise ratio.*

LoS *Line of sight.*

NLoS *Non-line of sight.*

SIC *Successive interference cancellation.*

GAP *Generalized assignment problem.*

3-DMP *3-dimensional matching problem.*

MKP *Multiple knapsack problem.*

AGWN *Additif white Gaussian noise.*

CSI *Channel state information.*

sMBS *Sub-6 GHz macro-base station.*

SBS *Small base station.*

mSBS *Millimeter-wave small base station.*

mBS *Millimeter-wave base station.*

zBS *Terahertz base station.*

sBS *Sub-6 GHz base station.*

KKT *Karush-Kuhn-Tucker.*

UARAP *User association and resource allocation problem.*

OPT *Optimum solution.*

C-HUA *Centralized greedy user association algorithm.*

D-NBS *Distributed nearest base station algorithm.*

D-RLA *Distributed reinforcement learning algorithm.*

D-BSA *Distributed base station algorithm.*

RAP *Resource allocation problem.*

GRA *Greedy resource allocation algorithm.*

GA *Genetic algorithm.*

WFA *Water filling algorithm.*

DRLA *Deep reinforcement learning algorithm.*

MDRLA *Multi-agent deep reinforcement learning algorithm.*

IRAA *Iterative resource allocation algorithm.*

S-IRAA *Separated iterative resource allocation algorithm.*

MADRA *Multi-agent deep reinforcement learning algorithm.*

CGRA *Centralized greedy algorithm.*

DGRA *Distributed greedy algorithm.*

RÉSUMÉ

Au cours des dernières décennies, plusieurs générations de réseaux cellulaires se sont succédé pour répondre au mieux aux exigences des applications mobiles émergentes et aux défis de la rareté du spectre radio. Différentes technologies telles que la communication à ondes millimétriques (mm-Wave), le déploiement des réseaux hétérogènes et des réseaux à bandes multiples sont des solutions prometteuses pour les futurs réseaux cellulaires.

Dans cette thèse, nous proposons des algorithmes efficaces d'allocation de ressources et d'association des utilisateurs aux stations de base pour les réseaux cellulaires à bandes multiples de prochaines générations. Dans le premier chapitre, nous introduisons de manière générale le projet de thèse en abordant les motivations, les problématiques, les contributions ainsi que la méthodologie adoptée. Dans le deuxième chapitre, nous étudions le problème conjoint d'association des utilisateurs et d'allocation des créneaux temporels. Chaque utilisateur ne peut être associé qu'à une seule station de base utilisant des fréquences inférieures à 6 GHz (sub-6 GHz) ou des fréquences millimétriques (mmWave), et chaque station de base ne peut associer qu'un nombre limité d'utilisateurs. L'objectif est de maximiser le nombre d'utilisateurs associés tout en garantissant un débit de données requis par utilisateur. Nous commençons par la formulation mathématique du problème comme un programme linéaire en nombres entiers, puis nous prouvons son appartenance à la classe des problèmes NP-difficile. Nous proposons un algorithme glouton centralisé et un autre distribué évolutif basé sur l'apprentissage automatique par renforcement. Dans le troisième chapitre, nous étudions le problème conjoint d'association des utilisateurs aux stations de base, d'allocation de puissance, d'affectation des sous-canaux et de la technique d'accès multiple. Nous considérons un réseau cellulaire utilisant la bande de fréquences sub-6 GHz et la bande de fréquences mmWave. Ainsi, les deux techniques d'accès multiple orthogonal et non orthogonal sont utilisées. La fonction objectif est d'optimiser le débit de données tout en respectant les exigences d'un débit minimum par utilisateur et les contraintes de puissance. Nous prouvons que le problème est NP-difficile par restriction et nous proposons des solutions algorithmiques. Tout d'abord, un algorithme glouton ainsi qu'une métaheuristique sont détaillés. Puis, nous proposons deux autres algorithmes basés sur la technique d'apprentissage automatique par renforcement profond. Ces algorithmes sont ensuite évalués et comparés. Dans le quatrième chapitre, nous considérons un réseau cellulaire utilisant trois bandes de fréquences sub-6 GHz, mmWave et térahertz et nous étudions le problème d'optimisation conjointe d'allocation de ressources et d'association des utilisateurs aux stations de base. Une nouvelle mesure d'efficacité énergétique qui quantifie le nombre d'utilisateurs satisfaits par unité de puissance est considérée. L'objectif est de maximiser l'efficacité énergétique du système tout en respectant les contraintes du système et de qualité de service. Le problème conjoint d'association des utilisateurs aux stations de base, d'allocation des canaux et de puissance est formulé comme un programme non linéaire partiellement en nombres entiers. Nous montrons que le problème appartient à la classe des problèmes NP-difficile et nous proposons des solutions algorithmiques peu complexes et évolutives. La première solution consiste à diviser le problème en deux sous-problèmes pour le résoudre de manière itérative. Tout d'abord, nous résolvons le sous-problème d'association des utilisateurs aux stations de base et d'attribution des canaux. Ensuite, nous résolvons le sous-problème d'allocation de puissance. Nous proposons aussi une solution multiagent basée sur l'apprentissage par renforcement profond. Chaque utilisateur est un agent qui apprend la meilleure association indépendamment des autres utilisateurs en maximisant son score. Nous présentons des résultats de simulation Monte-Carlo de chaque problème étudié afin d'illustrer les performances des solutions proposées.

Mots-clés : Réseaux cellulaires à bandes multiples, ondes millimétriques, fréquences térahertz, allocation de ressources, association des utilisateurs aux stations de base, techniques d'accès multiples, NOMA, efficacité énergétique, solutions algorithmiques, apprentissage par renforcement profond, NP-difficulté, algorithmes gloutons.

CHAPITRE 1

INTRODUCTION GÉNÉRALE

Ce chapitre introduit le projet de ma thèse de doctorat en informatique d'une manière générale. Il est divisé en cinq parties portant sur (i) la mise en contexte et les motivations, (ii) quelques définitions pertinentes, (iii) les problématiques, (iv) les contributions et la méthodologie utilisée pour mener à bien la recherche et finalement (v) l'organisation de la thèse.

1.1 Mise en contexte et motivations

Différentes générations de systèmes cellulaires utilisant diverses technologies se sont succédé et ont évolué ces dernières décennies pour satisfaire au mieux les exigences de couverture et de capacité. Ces exigences sont liées à la croissance explosive du nombre d'appareils mobiles connectés simultanément à l'internet ainsi qu'à la quantité de données mobiles échangées. Comme illustré sur la figure 1.1, le trafic mondial total de données mobiles, à l'exclusion du trafic généré par l'accès sans fil fixe (en anglais : *fixed wireless access*¹, FWA), est estimé à 325 exaoctets par mois en 2028, dont une augmentation d'un facteur de près de quatre par rapport à 2022. De plus, à titre de comparaison, le nombre d'appareils mobiles connectés passeront de 7,6 milliards en 2017 à 8,4 milliards en 2022, soit une augmentation de 9,5%. Une prévision sur cinq ans réalisée par *Ericsson* en novembre 2022 montre qu'une explosion du nombre d'appareils connectés passera des anciennes générations de systèmes cellulaires à la cinquième génération (5G) (Ericsson, 2022). Comme illustré sur la figure 1.2, ce nombre est estimé à environ 9,2 milliards d'ici la fin de 2028, dont la moitié sont des abonnements à la 5G. Par conséquent, il devient de plus en plus difficile de satisfaire les diverses demandes en termes de qualité de service (en anglais : *quality of service*). Ceci rend le problème d'allocation des ressources radio (p. ex., la puissance de transmission, les canaux, etc.) plus complexe. Cette complexité est due à la limitation des ressources dans les réseaux cellulaires. Plusieurs questions se posent dans ce contexte. Comment optimiser l'association des utilisateurs aux stations de base ? Comment transmettre plus de données en utilisant moins de ressources ? Comment choisir de manière optimale ou quasi-optimale la bande de fréquences de communication, la puissance et les canaux de communication afin de satisfaire au mieux une demande soumise à des contraintes de ressources ?

1. FWA est un type d'accès à internet à large bande utilisant les communications sans fil pour connecter deux emplacements fixes directement.

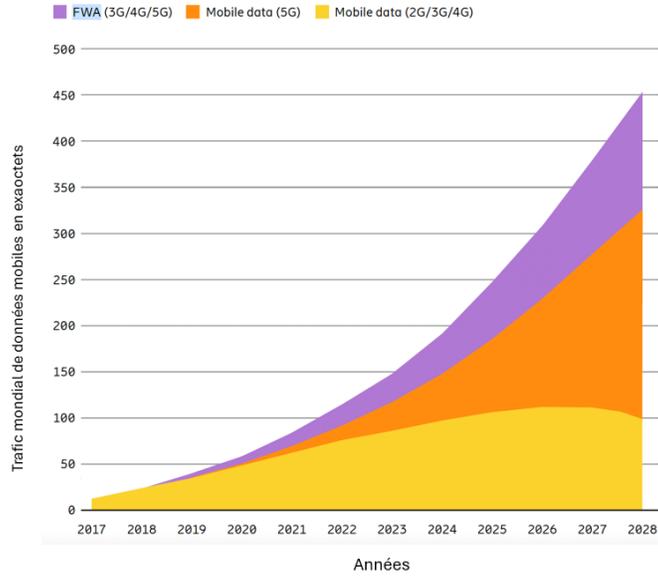


FIGURE 1.1 – Trafic mondial de données mobiles en exaotets (Ericsson, 2022).

Afin d’optimiser les performances des prochaines générations de systèmes cellulaires, telles que la 6G et ses évolutions futures, de nouvelles solutions algorithmiques d’allocation de ressources ainsi que l’intégration d’autres technologies et la combinaison de ces dernières sont nécessaires. Les performances de la 6G sont attendues pour être significativement améliorées par rapport aux générations précédentes.

Comme illustré sur la figure 1.3, ces améliorations incluent notamment : (i) un débit plus élevé allant jusqu’à 1 Tbit/s, (ii) une faible latence, (iii) une fiabilité de 99,9999999% et (iv) une capacité de connexion allant jusqu’à 10 millions d’appareils par kilomètre carré, soit 10 fois plus que la 5G (Vaezi *et al.*, 2022). En outre, selon le projet de partenariat de 3e génération (en anglais : *3rd Generation Partnership Project*, 3GPP), la 5G améliore la fiabilité de la connexion en situation de forte mobilité allant jusqu’à 500 Km/h et la 6G jusqu’à 1000 Km/h.

Plusieurs technologies sont proposées dans la littérature comme des solutions prometteuses pour atteindre les objectifs des futurs réseaux cellulaires. Nous citons à titre d’exemple : (i) le déploiement des réseaux denses hétérogènes de multi-niveaux (en anglais : *multi-tier and dense heterogeneous cellular networks*) où différents types de stations de base sont déployées dans la même zone géographique (p. ex., les stations de base macro-cellule, micro-cellule, pico-cellule, etc.) (Xu *et al.*, 2021; Adedoyin et Falowo, 2020), (ii) l’utilisation de plusieurs bandes de fréquences, notamment

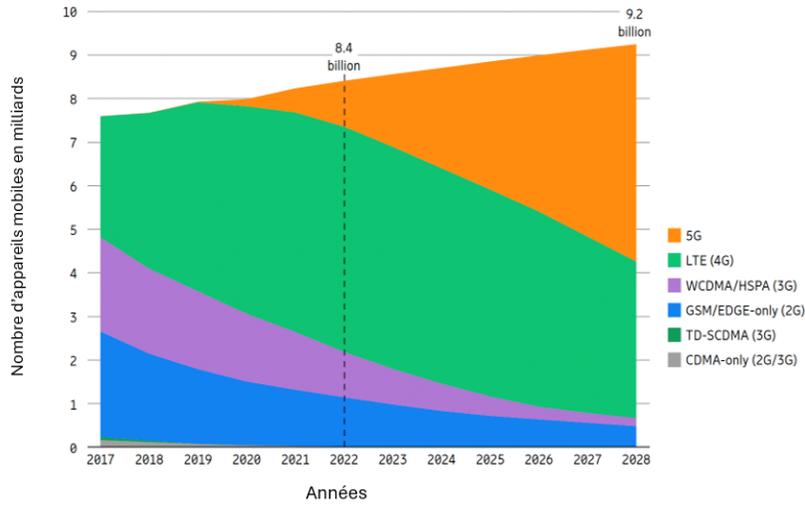


FIGURE 1.2 – Nombre d'appareils mobiles par technologie en milliards (Ericsson, 2022).

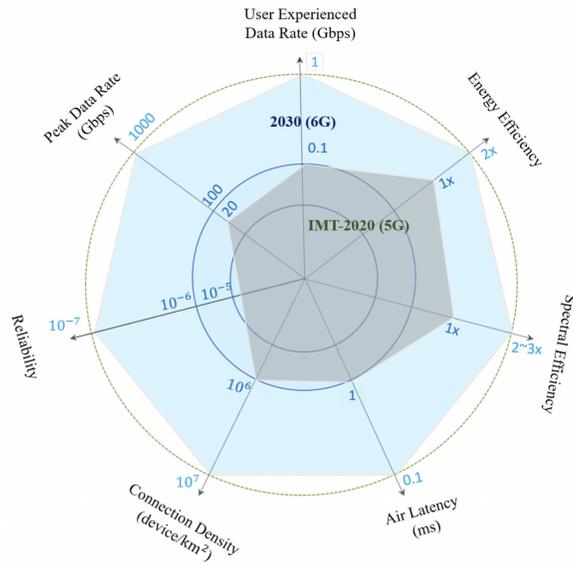


FIGURE 1.3 – Comparaison des performances entre la 5G (IMT-2020) et la 6G (Vaezi *et al.*, 2022).

la bande de fréquences millimétriques (en anglais : *millimeter-wave*, mmWave) et la bande de fréquences térahertz (en anglais : *terahertz*, THz) (Wang *et al.*, 2023; Vaezi *et al.*, 2022) et (iii) de la technique d'accès multiple non orthogonal (en anglais : *non-orthogonal multiple access*, NOMA) (Pei *et al.*, 2022; Dai *et al.*, 2018; Ding *et al.*, 2017).

Les réseaux cellulaires traditionnels (p. ex., 3G, 4G) opèrent sur les bandes de fréquences sub-6 GHz. À titre d'exemple, les bandes de fréquences hertziennes définies par le 3GPP pour le réseau d'évolution à long terme avancée (en anglais : *advanced long term evolution*, LTE) s'étalent de 450 MHz à 3,8 GHz, et pour la 5G la bande de fréquences sub-6 GHz (p. ex., 3,3 – 5 GHz) et la bande de très hautes fréquences entre 24,25 GHz et 71,0 GHz (p. ex., 24,25 – 29,5 GHz et 37 – 43,5 GHz) (3GPP, 2022). Étant donné que la bande de fréquences sub-6 GHz est largement utilisée par les réseaux cellulaires actuels ainsi que d'autres systèmes de communication, les chercheurs, les opérateurs téléphoniques et les fabricants d'équipements radio cherchent de plus en plus à exploiter d'autres bandes de fréquences plus élevées. Parmi ces bandes, la bande de fréquences mmWave dont la longueur d'onde est comprise entre 1 et 10 millimètres et la bande de fréquences THz. Bien que ces fréquences permettent des débits de données plus élevés en raison des larges bandes passantes qu'elles offrent, elles ont une portée plus courte et elles sont plus sensibles aux obstacles tels que les bâtiments et les arbres. En revanche, les fréquences sub-6 GHz sont parfaitement adaptées pour une couverture cellulaire étendue ainsi qu'une communication de longue portée (Xing et Rappaport, 2021; Akdeniz *et al.*, 2014).

La bande de fréquences mmWave s'étalent de 30 GHz à 300 GHz offrant une bande passante dix fois plus large que la bande de fréquences sub-6 GHz. De plus, les longueurs d'ondes millimétriques favorisent la miniaturisation des composants radio (p. ex., antennes, lignes d'adaptation des circuits radio-fréquence, amplificateurs, etc.) ce qui facilite l'intégration de la technique de multiplexage à entrées multiples et à sorties multiples (en anglais : *multiple-input multiple-output*, MIMO). La technique MIMO consiste à utiliser plusieurs antennes pour la transmission et la réception des signaux afin d'augmenter le débit de données, élargir la couverture et améliorer la fiabilité des communications (Tezergil et Onur, 2022; Zhou *et al.*, 2019; Xiao *et al.*, 2017). En contraste avec les caractéristiques de propagation des signaux sub-6 GHz, les signaux mmWave sont fortement atténués par la pluie et l'absorption atmosphérique en raison de l'interaction des molécules d'eau et d'oxygène avec ces ondes. À titre d'exemple, l'absorption d'oxygène introduit une atténuation

de 15 dB/km dans la bande de fréquences autour de 60 GHz (Rappaport *et al.*, 2011). De plus, les signaux mmWave sont susceptibles d'être affectés par des obstacles tels que la diffraction et l'obstruction, ce qui peut grandement dégrader la qualité de la liaison de communication mmWave. Le déploiement dense de petites cellules, couvrant chacune une zone limitée d'environ une dizaine de mètres, permet également d'étendre la couverture et de réduire les interférences (Theis *et al.*, 2021; Uwaechia et Mahyuddin, 2020; Akdeniz *et al.*, 2014).

Les systèmes de communications THz ont le potentiel d'offrir une large bande passante disponible entre 0,1 et 10 THz, pouvant supporter des débits de données ultra-élevés (plusieurs téraoctets par seconde). Récemment, l'exploitation de la bande de fréquences THz a reçu une attention considérable en tant que solution prometteuse pour les futurs systèmes cellulaires (p. ex., le projet européen Heat-X 6G flagship et le projet Next G Alliance des États-Unis) pour offrir des communications ultra-fiables à faible latence et une haute efficacité énergétique de 1 Tb/J (Maier, 2023; Uusitalo *et al.*, 2021). Similairement aux signaux mmWave, les signaux THz sont susceptibles d'être atténués par les conditions atmosphériques, en particulier par l'humidité de l'air, ce qui peut réduire la portée de transmission.

Pour améliorer les performances des liens mmWave et THz dans les réseaux sans fil, des études ont mené à l'utilisation des techniques de formation de faisceau (en anglais : *beamforming*), des antennes directionnelles et des surfaces intelligentes reconfigurables (en anglais, *reconfigurable intelligent surfaces*, RISs) afin de concentrer l'énergie radio dans une direction spécifique et de limiter les effets des trajets multiples et des pertes de trajet (Khan *et al.*, 2023). Les RISs sont constituées de milliers d'éléments de métamatériaux ou de petites antennes afin de contrôler dynamiquement les signaux radio en les réfléchissant, en les diffusant et en les amplifiant (pour les RISs actives) (Ma *et al.*, 2023; Rasilainen *et al.*, 2023).

De plus, l'utilisation des communications à bandes multiples présente plusieurs avantages. Premièrement, la bande de fréquences sub-6 GHz offre une meilleure capacité de pénétration à travers les bâtiments et les obstacles, tandis que les bandes de fréquences mmWave et THz permettent d'atteindre des débits de données plus élevés sur des courtes distances. Deuxièmement, l'utilisation de différentes bandes de fréquences permet d'augmenter la capacité du réseau et de satisfaire un plus grand nombre d'utilisateurs. Finalement, cette combinaison de fréquences permet d'améliorer

la fiabilité des réseaux cellulaires en proposant plusieurs solutions d'association. Ceci signifie que chaque utilisateur peut être associé à une ou plusieurs stations de base utilisant différentes bandes de fréquences. De plus, si un signal est perdu en raison d'un blocage ou d'un autre problème de propagation, l'utilisateur peut changer l'association (en anglais : *handover*) et la bande de communication pour continuer à bénéficier d'un service ininterrompu (Wang *et al.*, 2023; Moltchanov *et al.*, 2022; Agarwal *et al.*, 2022; Rasti *et al.*, 2022; Semiari *et al.*, 2019). Il est possible de combiner les bandes de fréquences de deux manières différentes : le déploiement hybride et le déploiement coexistant. Dans le déploiement hybride, chaque station de base peut opérer sur plusieurs bandes de fréquences simultanément, notamment les bandes sub-6 GHz et mmWave. En revanche, dans le déploiement coexistant, chaque station de base utilise une seule bande de fréquence et les stations de base sont déployées séparément dans la même zone géographique. Le déploiement des stations de base peut être fait d'une manière hétérogène. Les macro-stations de base sont généralement utilisées pour couvrir une large zone géographique, tandis que les petites stations de base sont utilisées pour couvrir une zone plus petite ou pour fournir une couverture intérieure. Les petites stations de base peuvent être aussi déployées dans des zones où les macro-stations de base ont des difficultés à fournir une capacité ou une couverture suffisante, par exemple dans les stades, les zones urbaines denses, les campus universitaires.

Les techniques d'accès multiple permettent à plusieurs utilisateurs d'utiliser le même canal de communication. Elles peuvent être classées en deux catégories : les techniques d'accès orthogonal et les techniques d'accès non orthogonal. Ces techniques peuvent être utilisées en fonction des besoins en termes de qualité de service et des contraintes du système de communication. Les techniques d'accès orthogonal (en anglais : *orthogonal multiple access*, OMA), telles que l'accès multiple par répartition temporelle (en anglais : *time division multiple access*, TDMA) et l'accès multiple par répartition en fréquence (en anglais : *frequency division multiple access*, FDMA), séparent les signaux des différents utilisateurs dans le temps ou la fréquence pour éviter les interférences. La technique TDMA consiste à allouer à chaque utilisateur une période de temps pendant laquelle il peut transmettre ses données. La technique FDMA consiste à allouer à chaque utilisateur une bande de fréquences. La technique d'accès non orthogonal (en anglais : *non-orthogonal multiple access*, NOMA) permet à plusieurs utilisateurs d'accéder simultanément à la même ressource de communication. Les signaux des différents utilisateurs sont superposés dans le domaine de puissance ou dans le domaine de code (en anglais : *power domain or code domain*) et sont séparés à l'aide des techniques adéquates

de traitement du signal. La technique NOMA dans le domaine de puissance est basée sur (i) un codage par superposition (en anglais : *superposition coding*) du côté de l'émetteur qui permet de superposer les données à des niveaux de puissance différents afin de les transmettre simultanément et (ii) sur une méthode d'annulation d'interférence successive (en anglais : *successive interference cancellation*, SIC) du côté du récepteur afin de séparer les différents signaux. La technique NOMA permet d'augmenter la capacité du système en termes de nombre d'associations et de débit, ainsi de réduire le temps d'attente et d'améliorer l'efficacité énergétique et spectrale (Islam *et al.*, 2018; Ding *et al.*, 2017; Morgado *et al.*, 2017). Cependant, les performances de NOMA sont directement liées au nombre d'utilisateurs partageant le même canal. Si le nombre d'utilisateurs est grand, la qualité de service peut se dégrader à cause de l'interférence (appelée aussi bruit NOMA). Les deux techniques OMA et NOMA peuvent être combinées pour contourner ce problème d'interférence. Cette technique hybride est appelée NOMA à sous-canaux multiples (en anglais : *multiple sub-channel-NOMA*, MSC-NOMA). La technique MSC-NOMA consiste à diviser les ressources fréquentielles d'une façon orthogonale et à former des groupes d'utilisateurs. Les utilisateurs de même groupe partagent les mêmes ressources alors que les différents groupes utilisent des ressources orthogonales. La technique MSC-NOMA est ainsi considérée une technique prometteuse pour les futurs réseaux cellulaires (Ding *et al.*, 2022; Zhu *et al.*, 2020; Liu *et al.*, 2020a).

Dans le contexte d'allocation des ressources dans les réseaux cellulaires, l'apprentissage automatique (en anglais : *machine learning*) suscite de plus en plus d'intérêt et fait l'objet de nombreuses études de recherche (Shi *et al.*, 2023; Gkonis, 2022; Bartsiakos *et al.*, 2022). En fait, l'apprentissage automatique est une branche de l'intelligence artificielle qui repose sur des approches mathématiques et statistiques. Le concept de l'apprentissage automatique est de concevoir des systèmes intelligents ayant la capacité d'apprendre automatiquement à partir des données et des expériences passées ou bien par interaction avec l'environnement (Russell et Norvig, 2010). Il existe différents types d'apprentissage automatique, qui se distinguent selon les informations disponibles lors de la phase d'apprentissage : apprentissage supervisé, non supervisé, par renforcement, etc. Compte tenu de la dynamique inhérente à l'environnement des réseaux mobiles cellulaires, en raison de la mobilité des utilisateurs et des obstacles, de la variation du gain du canal et de la charge du réseau, l'utilisation de l'apprentissage automatique par renforcement se révèle particulièrement bénéfique pour résoudre les problèmes d'optimisation. L'idée de l'apprentissage par renforcement est de déterminer la meilleure stratégie d'allocation de ressources qui maximise la récompense au cours du temps. Contrairement

à l'apprentissage supervisé qui demande des données correctement étiquetées, l'apprentissage par renforcement utilise une fonction de récompense pour trouver les meilleures décisions. De même, contrairement à l'apprentissage non supervisé qui vise à trouver des structures sous-jacentes à partir des données non étiquetées, l'apprentissage par renforcement cherche une stratégie qui optimise une fonction objectif par interaction avec un environnement initialement totalement inconnu (Russell et Norvig, 2010). L'apprentissage par renforcement consiste à faire apprendre à chaque agent (p. ex., un algorithme exécuté par une station de base, un appareil mobile ou un dispositif externe) à interagir avec l'environnement et à choisir une action parmi un ensemble d'actions possibles pour adapter son comportement aux changements. Les décisions prises par l'agent ont un impact sur les récompenses qu'il recevra à l'avenir (Sutton et Barto, 2018). Différents algorithmes d'apprentissage par renforcement ont été utilisés pour résoudre les problèmes d'allocation des ressources dans les réseaux sans fil, notamment l'algorithme de Q-apprentissage (en anglais : *Q-learning*, QL) dans (Alam *et al.*, 2023; Iqbal *et al.*, 2022; Majumdar *et al.*, 2021), l'algorithme de QL profond (en anglais : *deep Q-learning*, DQL) dans (Adeogun et Berardinelli, 2023; Tang *et al.*, 2020; He *et al.*, 2019) et l'algorithme du bandit manchot dans (He *et al.*, 2023; Lima *et al.*, 2022; Alizadeh et Vu, 2022).

1.2 Définitions pertinentes

Dans cette section, nous introduisons quelques définitions utiles qui sont utilisés dans cette thèse (Cormen *et al.*, 2010).

- Un algorithme glouton (ou vorace) est un algorithme itératif. À chaque itération, l'algorithme choisit un optimum local. Il ne revient jamais en arrière et il ne considère pas une solution si elle est rejetée. Bien que les algorithmes gloutons ne garantissent pas toujours une solution optimale à un problème d'optimisation, ils sont souvent efficaces pour trouver des solutions quasi-optimales en temps raisonnable.
- La complexité d'un problème est souvent mesurée en termes d'utilisation de la mémoire et du temps d'exécution qui dépend généralement des facteurs externes tels que le langage de programmation, le compilateur et le processeur. Dans cette thèse, nous nous intéressons particulièrement à la complexité temporelle, qui correspond au nombre d'opérations nécessaires à l'exécution des algorithmes. Cette complexité peut être étudiée dans le pire des cas, le cas moyen et le meilleur des cas. Les problèmes d'optimisation peuvent être classés en fonction de leurs difficultés. Les classes de complexité les plus connues sont P, NP, NP-complet et

NP-difficile.

- Un problème est considéré comme étant dans la classe P (en anglais : *polynomial time*) si une solution peut être trouvée en temps polynomial, c'est-à-dire en un temps proportionnel à la taille des entrées.
- Un problème est considéré comme étant dans la classe NP (en anglais : *non-deterministic polynomial time*) est un problème de décision qui peut être résolu par un algorithme polynomial non déterministe.
- Un problème est considéré comme étant NP-complet s'il est à la fois dans la classe NP et il est réductible au problème de satisfiabilité d'une expression logique par une transformation polynomiale.
- Un problème est considéré comme étant NP-difficile s'il est au moins aussi difficile que n'importe quel problème de la classe NP, mais il n'est pas nécessairement dans la classe NP. Ainsi, un problème est dit NP-difficile si tout problème de la classe NP peut être réduit en temps polynomial à ce problème.

1.3 Problématiques

Cette section détaille la problématique générale ainsi que les problématiques spécifiques de cette thèse.

1.3.1 Problématique générale

Dans un contexte où les ressources de communication sont limitées, l'enjeu majeur consiste à gérer de manière optimale ces ressources pour répondre aux demandes exigeantes des utilisateurs en termes de qualité de service. Dans nos travaux de recherche, nous nous concentrons sur l'optimisation de plusieurs ressources et qui sont :

- La bande passante : C'est la plage de fréquences utilisée par un signal électronique sur un moyen de transmission donné. Nous considérons les bandes de fréquences sub-6 GHz, mmWave et THz.
- Le canal de transmission : Il est créé en divisant la bande de fréquences disponible en des unités plus petites. Si la technique OMA est considérée, chaque canal est alloué à un seul utilisateur. En revanche, si la technique MSC-NOMA est considérée, chaque canal peut être partagé par plusieurs utilisateurs du même groupe.

— Le temps de transmission : Il désigne le créneau temporel alloué pour transmettre les données. Dans cette thèse, nous nous intéressons à déterminer la répartition optimale ou quasi-optimale des ressources radio afin d’optimiser les performances du lien descendant du système. L’objectif général est d’étudier minutieusement les problèmes d’allocation des ressources dans les réseaux cellulaires à bandes multiples et de les résoudre efficacement. Différentes fonctions objectif sont étudiées, notamment l’optimisation du nombre d’utilisateurs associés aux stations de base, le débit de données et l’efficacité énergétique. Les modèles du système considérés sont réalistes et relativement peu étudiés dans la littérature. Cependant, ils représentent des modèles prometteurs pour atteindre les objectifs des futurs réseaux cellulaires. Il est à noter que, dans un réseau cellulaire à bandes multiples, non seulement l’association des utilisateurs aux stations de base qui doit être prise en compte dans la procédure d’allocation des ressources, mais également la bande de fréquences de communication, le choix du canal et le niveau de puissance doivent être conjointement considérés. Les solutions algorithmiques proposées sont efficaces, innovantes et uniques dans la mesure où elles résolvent des problèmes complexes.

1.3.2 Problématiques spécifiques

Les problèmes d’allocation des ressources radio sont souvent de nature combinatoire et donc difficiles à résoudre, étant donné la multitude de variables d’optimisation à considérer. Ces variables peuvent prendre différentes formes, telles que des variables binaires, entières et continues. Par conséquent, des algorithmes efficaces et peu complexes sont nécessaires. Dans cette thèse, nous avons étudié trois problématiques distinctes, chacune portant sur un problème spécifique d’allocation de ressources. Dans ce qui suit, nous présentons ces trois problématiques.

- La première problématique consiste à étudier le problème conjoint d’association des utilisateurs et d’allocation des ressources temporelles dans les réseaux cellulaires à bandes multiples utilisant les deux bandes de fréquences sub-6 GHz et mmWave. Les ressources sont définies par le nombre de créneaux temporels alloués par les stations de base aux utilisateurs. La solution doit maximiser le nombre d’associations tout en respectant les contraintes de ressources disponibles.
- La deuxième problématique consiste à étudier le problème d’association des utilisateurs, d’allocation de puissance, d’affectation des sous-canaux et de la technique d’accès multiple à chaque sous-canal. Un réseau à bandes multiples est considéré où des stations de base utilisant

la bande de fréquences sub-6 GHz et d'autres utilisant la bande de fréquences mmWave sont déployées dans la même zone géographique. Pour chaque canal alloué, deux techniques d'accès multiple peuvent être utilisées : OMA ou MSC-NOMA.

- La troisième problématique consiste à optimiser l'efficacité énergétique dans les réseaux cellulaires à bandes multiples : sub-6 GHz, mmWave et THz. Nous introduisons une nouvelle métrique d'efficacité énergétique qui mesure le nombre d'utilisateurs qu'un watt peut satisfaire. Le problème d'optimisation regroupe l'association des utilisateurs aux stations de base, l'allocation de puissance et des canaux. Chaque utilisateur peut s'associer à une seule station de base utilisant un seul canal de transmission.

1.4 Contributions et méthodologie

Dans ce qui suit, nous discutons les différentes contributions de cette thèse. Chaque contribution est présentée dans une sous-section ainsi que la méthodologie et les publications scientifiques de nos travaux.

1.4.1 Maximisation du nombre d'utilisateurs associés dans un réseau à bandes multiples

L'objectif de la première contribution est de maximiser le nombre d'utilisateurs associés sous des contraintes de qualité de service. Ces contraintes sont définies par un débit de données minimum pour chaque utilisateur. Nous considérons un modèle du système composé d'une macro-station de base utilisant la bande de fréquences sub-6 GHz, des petites stations de base opérant sur des fréquences mmWave et d'autres partageant les mêmes fréquences que la macro-station de base. Les hypothèses adoptées sont choisies de la manière la plus réaliste possible. Nous formulons le problème d'association des utilisateurs et d'allocation des ressources temporelles sous forme d'un programme linéaire en nombres entiers et en variables binaires. Les variables binaires sont les variables d'association et les variables en nombre entiers représentent le nombre de créneaux temporels alloués par les stations de base aux utilisateurs. Ensuite, nous étudions la complexité du problème et nous prouvons son appartenance à la classe des problèmes NP-difficile. La résolution du problème est faite d'une manière optimale en utilisant le langage de modélisation algébrique AMPL (en anglais : *a mathematical programming modeling language*) (AMPL, 2020). Ce langage offre une interface commune pour de nombreux solveurs tels que CPLEX pour résoudre les programmes linéaires et MINOS pour les programmes non linéaires (IBM, 2017; MINOS, 2003). Pour résoudre le problème

étudié en temps raisonnable, nous proposons deux algorithmes. Le premier est un algorithme glouton centralisé, implémenté au niveau d'un point central. Le deuxième est un algorithme distribué basé sur l'apprentissage automatique par renforcement QL. L'apprentissage consiste à trouver la meilleure association pour chaque utilisateur en fonction de son expérience passée, de manière autonome et indépendante des autres. Les résultats de simulation Monte-Carlo confirment que les algorithmes proposés permettent d'approcher de la solution optimale avec une complexité de calcul considérablement réduite.

Cette contribution a été publiée dans le journal *IEEE Systems Journal*, 2020 (Chaieb *et al.*, 2020b) et une partie des résultats ont été préalablement publiée dans les actes de la conférence *IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications, Montreal, QC, Canada, 2017* (Chaieb *et al.*, 2017).

1.4.2 Techniques d'accès multiple dans un réseau à bandes multiples

L'objectif de la deuxième contribution est de maximiser le débit de données en tenant compte d'un débit minimum par utilisateur et des contraintes de puissance. Le modèle du système étudié est conçu pour être le plus réaliste possible. Nous considérons un réseau composé des stations de base utilisant la bande de fréquences sub-6 GHz et d'autres utilisant la bande de fréquences mmWave. Deux techniques d'accès multiple sont considérées : l'accès orthogonal et l'accès non orthogonal. Chaque station de base peut choisir entre l'utilisation de l'une ou l'autre de ces techniques pour chaque canal. La technique OMA consiste à allouer un canal à un utilisateur. En revanche, la technique MSC-NOMA permet à plusieurs utilisateurs appartenant au même groupe de partager simultanément le même canal, tandis que les utilisateurs appartenant à des groupes différents utilisent des canaux différents. Ceci permet d'augmenter le nombre d'utilisateurs satisfaits et d'optimiser l'utilisation des canaux disponibles et le débit du système. Cette contribution aborde le problème d'allocation des ressources, y compris l'association des utilisateurs aux stations de base, l'allocation des canaux et de la puissance, ainsi que la sélection de la technique d'accès. La sélection de la technique d'accès doit prendre en considération divers facteurs tels que le nombre d'utilisateurs, la qualité du canal de transmission et les contraintes de puissance et de qualité de service. Le problème formulé est un programme non linéaire et non convexe en variables partiellement en nombres entiers. Deux ensembles de variables d'optimisation sont considérées : les variables binaires d'association et les

niveaux de puissance de transmission alloués. Nous étudions la complexité du problème et nous proposons des solutions algorithmiques. Nous commençons par un algorithme glouton de complexité réduite et une métaheuristique basée sur la technique génétique. Puis, nous proposons un algorithme centralisé et un autre multiagent. Ces deux algorithmes reposent sur la technique d'apprentissage par renforcement profond. Les solutions algorithmiques proposées sont évaluées et comparées par simulation utilisant la méthode de Monte Carlo. Les résultats montrent les performances de nos solutions et les avantages de l'utilisation de deux techniques d'accès dans les réseaux cellulaires à bandes multiples, notamment en termes de nombre d'utilisateurs associés, de débit de données et d'équité entre les utilisateurs.

Cette contribution a été publiée dans le journal *IEEE Transactions on Communications, 2023* (Chaieb et al., 2023a). Des résultats préliminaires ont été publiés dans les conférences *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, London, UK, 2020* (Chaieb et al., 2020a) et *IEEE International Conference on Communications, Seoul, Republic of Korea, 2022* (Chaieb et al., 2022).

1.4.3 Efficacité énergétique dans un réseau à bandes multiples

L'objectif de la troisième contribution est de maximiser l'efficacité énergétique dans un réseau à bandes multiples utilisant les bandes de fréquences sub-6 GHz, mmWave et THz. Nous introduisons une nouvelle définition d'efficacité énergétique qui mesure le nombre d'utilisateurs qu'un watt peut satisfaire. Plus précisément, cette mesure est égale à la somme des associations par niveau de puissance alloué, sachant que chaque utilisateur ne peut s'associer qu'à une seule station de base utilisant la bande de fréquences sub-6 GHz, mmWave ou THz. Ceci rend le problème d'association plus difficile car il faut choisir la station de base et la bande passante qui maximisent l'efficacité énergétique du système. Nous considérons deux contraintes de qualité de service par utilisateur. La première est un débit de données minimum et la deuxième est un rapport signal sur interférence et bruit minimum (en anglais : *signal to interference plus noise ratio*, SINR). Le problème conjoint d'association des utilisateurs aux stations de base ainsi que d'allocation des canaux et de puissance est formulé comme un programme non linéaire et non convexe partiellement en nombres entiers. Le problème appartient à la classe des problèmes NP-difficile. Pour le résoudre en temps raisonnable, nous le divisons en deux sous-problèmes que nous résolvons de manière itérative. Tout d'abord, nous

étudions le problème conjoint d’association des utilisateurs aux stations de base et d’allocation des canaux en respectant les contraintes de puissance et de qualité de service. Ensuite, nous étudions le problème d’allocation de puissance en fonction des exigences de qualité de service minimales pour maximiser l’efficacité énergétique du réseau. En outre, nous proposons une autre solution d’allocation de ressources basée sur l’apprentissage par renforcement profond. Nous utilisons la méthode QL avec deux différentes fonctions de récompense. La première est calculée de manière distribuée et se base uniquement sur le score de l’utilisateur, tandis que la deuxième est calculée de manière centralisée et dépend du score de tous les utilisateurs. Les résultats de simulation utilisant la méthode de Monte-Carlo démontrent l’efficacité des solutions proposées, notamment la solution itérative, la solution d’apprentissage par renforcement profond et la solution gloutonne. Ils ont également mis en évidence l’impact des communications à bandes multiples sur l’amélioration de l’efficacité énergétique du système.

Cette contribution fait l’objet d’un papier de revue soumis au journal *IEEE Transactions on Wireless Communications*. Des résultats préliminaires de ce travail ont été présentés à la conférence *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Toronto, Canada, 2023* (Chaieb *et al.*, 2023b).

1.5 Organisation

Dans la suite, nous présentons les différentes contributions de la thèse. Chaque chapitre de ce rapport détaille une contribution du travail de la thèse et il est divisé en huit parties portant respectivement sur (i) le résumé de la contribution, (ii) l’introduction présentant l’état de l’art, les contributions et l’organisation, (iii) le modèle du système et les hypothèses utilisées, (iv) la formulation mathématique du problème et les contraintes, (v) l’analyse de sa complexité, (vi) les différentes solutions proposées, (vii) les résultats de simulation et enfin (viii) les conclusions. Le chapitre 2 discute notre première contribution qui vise à maximiser le nombre d’utilisateurs satisfaits dans un réseau cellulaire à bandes multiples utilisant les deux bandes de fréquences sub-6 GHz et mmWave. Le chapitre 3 présente notre deuxième contribution qui porte sur le problème d’association des utilisateurs et d’allocation des ressources dans un réseau cellulaire à bandes multiples utilisant les deux techniques d’accès multiple orthogonal et non orthogonal. Ensuite, le chapitre 4 propose des solutions algorithmiques d’association des utilisateurs et d’allocation de ressources dont le but est de maximiser l’efficacité

énergétique dans un réseau cellulaire à bandes multiples utilisant les bandes de fréquences sub-6 GHz, mmWave et THz. Finalement, le dernier chapitre conclut les contributions et les résultats de la thèse et liste des perspectives de recherche futures.

1.6 Conclusion

Dans ce chapitre, nous avons commencé par la mise en contexte, les motivations et les problématiques abordées. Ensuite, nous avons résumé les contributions qui seront minutieusement détaillées dans les chapitres à venir ainsi que la méthodologie suivie. Enfin, nous avons présenté l'organisation des chapitres de cette thèse.

CHAPITRE 2

MAXIMISATION DU NOMBRE D'UTILISATEURS ASSOCIÉS DANS UN RÉSEAU À BANDES MULTIPLES

Dans ce chapitre, nous étudions le problème d'association des utilisateurs aux stations de base et d'allocation des ressources temporelles dans les réseaux cellulaires à bandes multiples. Deux bandes de fréquences sont utilisées : la bande de fréquences sub-6 GHz et la bande de fréquences extrêmement hautes qui s'étend de 30 à 300 GHz (également connue sous le nom de la bande de fréquences mmWave).

Ce travail a été publié dans le journal *IEEE Systems Journal*, 2020 (Chaieb *et al.*, 2020b) et quelques résultats ont été préalablement publiés dans les actes de la conférence *IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, Montreal, QC, Canada, 2017 (Chaieb *et al.*, 2017).

2.1 Résumé

Ce chapitre étudie le problème conjoint d'association des utilisateurs aux stations de base et d'allocation des ressources dans les réseaux cellulaires. Ces ressources sont définies par le nombre de créneaux temporels alloués par des stations utilisant des fréquences sub-6 GHz ou mmWave. L'objectif est de maximiser le nombre d'utilisateurs associés tout en garantissant le débit de données requis par chaque utilisateur. Nous commençons par la formulation mathématique du problème puis nous prouvons son appartenance à la classe NP-difficile. Nous proposons deux algorithmes. Le premier est un algorithme basé sur l'approche gloutonne centralisé implémenté au niveau d'un point central et le deuxième est un algorithme distribué basé sur l'apprentissage automatique par renforcement. L'objectif de l'apprentissage est de trouver la meilleure association de chaque utilisateur en fonction de son expérience passée indépendamment des autres. Les résultats de simulation Monte-Carlo montrent que les algorithmes proposés approchent la solution optimale avec une complexité de calcul considérablement réduite.

2.2 Introduction

Les réseaux cellulaires à bandes multiples sont identifiés comme une solution prometteuse afin d'améliorer à la fois la capacité et la couverture (Alsaedi *et al.*, 2023). L'association des utilisateurs, c'est-à-dire l'association d'un utilisateur à une station de base particulière, est un problème bien étudié en général. Cependant, le déploiement des réseaux cellulaires à bandes multiples rend le problème plus complexe, car il implique non seulement la décision d'association, mais aussi le choix de la bande de fréquences de communication.

Dans ce chapitre, nous étudions le problème d'association des utilisateurs aux stations de base et d'allocation des ressources dans un réseau cellulaire hétérogène à bandes multiples. Sans perte de généralité, dans ce qui suit les ressources désignent les créneaux temporels, mais tous les autres types de ressources peuvent être considérés. Le problème étudié consiste à maximiser le nombre d'utilisateurs associés tout en respectant leurs exigences en termes de qualité de service. Puisque plusieurs solutions donnent le même nombre d'utilisateurs associés, celle utilisant le moins de créneaux temporels est choisie afin d'assurer une utilisation efficace de ressources.

2.2.1 Revue de littérature

Les principaux travaux de recherche dans la thématique d'allocation des ressources dans les réseaux à bandes multiples sont résumés ci-après. Les auteurs de (Semiari *et al.*, 2017) ont proposé un protocole de contrôle d'accès au média pour gérer dynamiquement le trafic des réseaux locaux sans fil sur les bandes de fréquences mmWave non licenciées (2.4 GHz et 5 GHz) et celles sub-6 GHz. Les résultats obtenus ont montré que les communications à bandes multiples améliorent significativement le débit et le délai. Dans (Liu, 2019), les auteurs ont proposé un schéma généralisé d'association des utilisateurs dans un réseau hétérogène à bandes multiples utilisant la bande de fréquences qui s'étend de 300 à 3000 MHz (également connue sous le nom de la bande d'ultra-hautes fréquences) et la bande de fréquences mmWave. Les auteurs ont développé une approche de modélisation et d'analyse basée sur la géométrie stochastique dans le but d'étudier l'association, le débit et la probabilité de couverture pour plusieurs modèles d'affaiblissement de propagation. Les auteurs ont démontré un compromis entre la couverture et le débit dans les réseaux cellulaires à bandes multiples. Autrement dit, lorsque deux différentes bandes de fréquences sont utilisées, l'optimisation simultanée de la couverture et du débit en utilisant le même schéma d'association devient une tâche difficile.

Dans (Shimodaira *et al.*, 2014), les auteurs ont étudié le problème d’association des utilisateurs dans un réseau hétérogène à bandes multiples composé d’une macro et des petites stations de base opérant respectivement sur les bandes de fréquences des 60 GHz et 2 GHz. Les auteurs ont proposé une méthode d’association des utilisateurs aux stations de base qui tient compte du débit de données, du trafic et du nombre d’utilisateurs associés à chaque station de base. Les résultats de simulation ont montré que la méthode d’association proposée permet d’obtenir un débit deux fois plus élevé que la méthode conventionnelle, où chaque utilisateur est associé à la station de base offrant le meilleur SINR. Nous soulignons que les auteurs dans (Liu, 2019; Semiari *et al.*, 2017; Shimodaira *et al.*, 2014) n’ont pas étudié la complexité des problèmes d’optimisation et n’ont pas considéré la même fonction objectif que celle étudiée dans ce chapitre. Des stratégies d’association des utilisateurs en liaison montante et descendante aux différentes stations de base dans un réseau hétérogène dense à bandes multiples ont été étudiées dans (Elshaer *et al.*, 2016). Dans le but de maximiser à la fois le SINR et la probabilité de couverture, les auteurs ont proposé un modèle analytique d’association des utilisateurs basé sur la technique d’accès découplé en liaison montante et descendante (en anglais : *downlink-uplink decoupling technique*). En utilisant cette technique, chaque utilisateur peut être associé à une seule station de base en liaison montante et descendante, ou à la fois à une macro-station de base opérant sur la bande de fréquences sub-6 GHz et à une petite station de base opérant sur la bande de fréquences mmWave. L’étude dans (Bai et Heath, 2015) a montré que les réseaux cellulaires mmWave peuvent atteindre des couvertures comparables et des débits de données beaucoup plus élevés que les réseaux sub-6 GHz. De plus, les auteurs de (Zhang *et al.*, 2017) ont proposé un algorithme basé sur la technique du gradient itératif pour résoudre le problème d’association des utilisateurs et d’allocation de puissance dans les réseaux mmWave ultra-denses sous des contraintes de qualité de service et de puissance. Ils ont montré que l’efficacité énergétique s’améliore lorsque la bande de fréquences mmWave et les technologies d’énergie renouvelable sont utilisées dans des réseaux cellulaires ultra-denses. Cependant, ils n’ont pas pris en compte les communications à bandes multiples. Dans (Zalghout *et al.*, 2017), les auteurs ont étudié le problème conjoint d’association des utilisateurs et d’allocation des ressources dans les réseaux cellulaires hétérogènes utilisant seulement la bande de fréquences sub-6 GHz. Ils ont introduit une fonction de profit des utilisateurs, qui est calculée en tenant compte de la consommation d’énergie et de la qualité du signal reçu. Pour maximiser ce profit, ils ont développé un algorithme heuristique centralisé de complexité polynomiale. Dans cet algorithme, l’association des utilisateurs est déterminée en fonction de la distance entre les utilisateurs et les stations de base. Néanmoins, cette approche n’est pas toujours efficace,

car le gain du canal dépend d'autres paramètres que la distance. Dans (Alizadeh et Vu, 2019), un schéma d'association des utilisateurs et de répartition de charge dans les réseaux cellulaires a été proposé afin d'améliorer l'équité et les débits de transmission des utilisateurs au bord des cellules. Cependant, (Alizadeh et Vu, 2019) ne tient pas compte des communications sub-6 GHz. Dans les travaux mentionnés ci-dessus, aucune preuve robuste n'a été fournie quant à la complexité des problèmes étudiés. De plus, les solutions algorithmiques proposées ne reposent pas sur l'apprentissage automatique et requièrent des informations complètes sur le réseau.

L'apprentissage par renforcement est une technique puissante pour résoudre des problèmes d'optimisation complexes, notamment les problèmes d'allocation des ressources dans les réseaux cellulaires. Les auteurs de (Hu *et al.*, 2017) ont proposé un modèle basé sur la théorie des jeux pour résoudre le problème conjoint d'association des utilisateurs, d'allocation des fréquences et d'équilibrage de charge dans les réseaux LTE utilisant le spectre sans licence. La technique d'accès découplé en liaison montante et descendante a été utilisée et un algorithme décentralisé basé sur l'apprentissage par renforcement QL a été proposé. L'apprentissage permet aux stations de base d'apprendre à sélectionner d'une manière autonome leurs schémas optimaux d'allocation du spectre. Les résultats obtenus ont démontré que l'algorithme QL proposé permet d'obtenir des gains de performance significatifs en termes de débit et d'équilibrage de charge par rapport à d'autres approches telles que l'algorithme du plus proche voisin. Cependant, les auteurs n'ont pas présenté les performances de l'algorithme en termes du nombre d'utilisateurs associés. Le problème conjoint d'allocation de puissance et d'association des utilisateurs dans les réseaux hétérogènes a été étudié dans (Fan *et al.*, 2019) dans le but de maximiser le débit du système. Basé sur la théorie des graphes, le problème a été décomposé en plusieurs sous-problèmes selon la topologie du réseau. Les auteurs ont proposé un algorithme d'apprentissage par renforcement distribué et évolutif pour résoudre les sous-problèmes tout en tenant compte du coût des messages échangés entre les utilisateurs et les stations de base. Les auteurs ont aussi proposé une méthode d'accélération du processus d'apprentissage en utilisant les connaissances préalables sur les variations du canal mmWave. L'analyse des performances des résultats obtenus à partir de mesures réelles a démontré l'efficacité de l'apprentissage comparativement à d'autres références. Les auteurs de (Malila *et al.*, 2018) ont aussi proposé un algorithme d'apprentissage par renforcement dans le but d'améliorer la fiabilité des liens de communication mm-Wave dans les réseaux d'amenés (en anglais : *backhaul networks*). L'algorithme permet au système de détecter un dysfonctionnement sur un chemin dans un délai prédéterminé et de le changer, tout

en adaptant la puissance de transmission et le mode de fonctionnement. Les auteurs ont démontré que l'apprentissage par renforcement est une technique solide pour s'adapter aux changements de l'environnement. Nous notons que les auteurs de (Fan *et al.*, 2019) et (Malila *et al.*, 2018) n'ont pas pris en considération les communications sur différentes bandes de fréquences.

L'étude de la littérature révèle que la majorité des articles se sont focalisés sur l'allocation des ressources dans les réseaux hétérogènes utilisant une seule bande de fréquences. En conséquence, de nouveaux algorithmes d'allocation de ressources et d'association des utilisateurs aux stations de base doivent être développés pour assurer une utilisation efficace de ressources lorsque les deux bandes sub-6 GHz et mmWave sont disponibles à utiliser.

2.2.2 Contributions

Motivé par les travaux mentionnés dans la section précédente, ce chapitre propose de nouvelles stratégies d'allocation des ressources et d'association des utilisateurs pour les réseaux cellulaires hétérogènes à bandes multiples. L'association des utilisateurs aux stations de base implique non seulement le choix de la station de base, mais aussi le choix de la bande de fréquences (sub-6 GHz ou mmWave). En effet, le choix de la bande de fréquences est important, car il peut affecter la qualité de la communication et les performances du réseau. Les contributions de ce chapitre sont résumées comme suit :

1. Nous considérons un réseau cellulaire utilisant les deux bandes de fréquences sub-6 GHz et mmWave et nous formulons le problème conjoint d'association et d'allocation des ressources temporelles, appelé UARAP, comme un programme linéaire en nombres entiers (en anglais : *integer linear programming*). Contrairement aux fonctions objectif étudiées dans (Fan *et al.*, 2019; Liu, 2019; Semiari *et al.*, 2017; Elshaer *et al.*, 2016; Shimodaira *et al.*, 2014), UARAP maximise le nombre d'utilisateurs associés tout en garantissant un débit minimal de données ainsi qu'une utilisation efficace de ressources.
2. Nous montrons que UARAP est NP-difficile en réduisant le problème d'affectation généralisé, qui est un problème connu NP-difficile, à un cas particulier de UARAP.
3. Nous proposons un algorithme glouton centralisé de complexité polynomiale.
4. Nous proposons un algorithme distribué basé sur la technique d'apprentissage par renforcement. Cet algorithme vise à trouver la meilleure association de chaque utilisateur automati-

quement et indépendamment des autres utilisateurs en se basant seulement sur son expérience passée.

5. Nous démontrons les performances des solutions proposées par simulation en les comparant avec la solution optimale.

2.2.3 Organisation

La suite de ce chapitre est organisé comme suit. La section 2.3 présente le modèle du système. La section 2.4 formule le problème mathématiquement et la section 2.5 étudie sa complexité. Les solutions algorithmiques proposées sont présentées dans la section 2.6. Les résultats de simulation sont fournis dans la section 2.7 et les conclusions sont explicitées dans la dernière section 2.8.

2.3 Modèle du système

Le modèle du système considéré est un réseau cellulaire hétérogène composé d'une station de base macro-cellule utilisant les fréquences sub-6 GHz (sMBS), des petites stations de base (SBSs) opérant sur des fréquences mmWave (mSBSs) et d'autres SBSs partageant les mêmes fréquences que la sMBS (sSBSs). Pour éviter toute confusion, nous utilisons sBS pour désigner une sMBS ou une sSBS. Les notations utilisées dans ce chapitre sont énumérées dans le tableau 2.1. Un exemple du modèle est illustré sur la figure 2.1, où la transmission sur le lien descendant est considérée. Nous notons respectivement par $\mathcal{M} = \{1, 2, \dots, M\}$, $\mathcal{N} = \{0, 1, 2, \dots, N\}$ et $\mathcal{K} = \{1, 2, \dots, K\}$ les ensembles des mSBSs, des sBSs et des utilisateurs. Les stations de base utilisent une technique d'accès multiple orthogonal. Nous désignons par 0 la sMBS et par $n = 1, 2, \dots, N$ une sSBS. Nous supposons que la sMBS est équipée de A_0 antennes et que chaque mSBS est équipée de A_m antennes. Nous supposons que chaque utilisateur ne peut être associé qu'à une sBS ou à une mSBS et l'état du canal sans fil est invariant pendant la période d'association.

Pour modéliser UARAP mathématiquement, nous introduisons les deux matrices binaires $\mathbf{X} = [x_{km}]$ et $\mathbf{X}' = [x'_{kn}]$. Les variables d'association x_{km} et x'_{kn} sont définies comme suit :

$$x_{km} = \begin{cases} 1, & \text{si l'utilisateur } k \in \mathcal{K} \text{ est associé à la mSBS } m \in \mathcal{M} \\ 0, & \text{sinon} \end{cases}, \quad (2.1)$$

TABLE 2.1 – Notations et paramètres de simulation.

Notation	Paramètre	Valeur
sMBS	Station de base macro-cellule opérant sur des fréquences sub-6 GHz	
sSBS	Petite station de base opérant sur des fréquences sub-6 GHz	
mSBS	Petite station de base opérant sur des fréquences mmWave	
sBS	Station de base opérant sur des fréquences sub-6 GHz	
β	Constante capturant des effets du système et de la transmission	10^3
M	Nombre de mSBSs	2
N	Nombre de sSBSs	2
γ_0	SNR de chaque sMBS	40 dBm
γ_n	SNR de chaque sSBS	30 dBm
p_m	Puissance de transmission de la mSBS m	30 dBm
W'	Bande passante d'un créneau temporel alloué par la sBS	180 kHz
W	Bande passante d'un créneau temporel alloué par la mSBS	$2 * W'$
W_m	Bande passante mmWave	1 GHz
A_m	Nombre d'antennes de la mSBS	4
A_0	Nombre d'antennes de la sMBS	20
Q_k	Débit de données demandé par l'utilisateur k	1 Mbps
s_m	Nombre maximal d'utilisateurs associés à la mSBS m	8
s'_n	Nombre maximal d'utilisateurs associés à la sSBS n	8
s'_0	Nombre maximal d'utilisateurs associés à la sMBS	16
t_m	Nombre de créneaux temporels disponibles pour la mSBS m	50
t'_n	Nombre de créneaux temporels disponibles pour la sSBS n	50
t'_0	Nombre de créneaux temporels disponibles pour la sMBS	50
d_0, d'_0	Distances de référence	1 m, 40 m
α	Exposant d'affaiblissement	4
δ	Facteur d'apprentissage	0.5
Y	Nombre d'itérations	10^3
σ^2	Puissance du bruit	$-174 \text{ dBm/Hz} + 10 \log_{10} W_m + 10 \text{ dB}$
ψ, κ, ν	Paramètres du modèle d'affaiblissement pour les transmissions NLoS (Akdeniz <i>et al.</i> , 2014)	$\psi = 86.6, \kappa = 2.45$ $\nu = 8 \text{ dB}$
ψ, κ, ν	Paramètres du modèle d'affaiblissement pour les transmissions LoS (Akdeniz <i>et al.</i> , 2014)	$\psi = 69.8, \kappa = 2$ $\nu = 5.8 \text{ dB}$

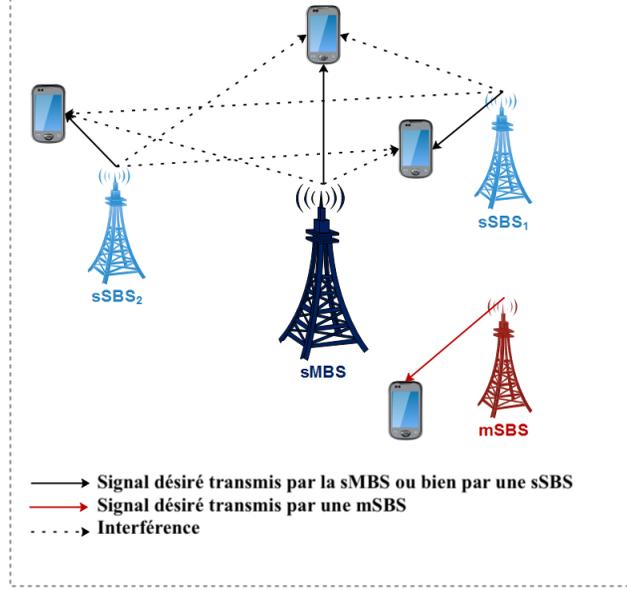


FIGURE 2.1 – Une illustration d'un réseau hétérogène à bandes multiples.

$$x'_{kn} = \begin{cases} 1, & \text{si l'utilisateur } k \in \mathcal{K} \text{ est associé à la sBS } n \in \mathcal{N} \\ 0, & \text{sinon} \end{cases}. \quad (2.2)$$

Puisque la sMBS et les sSBS partagent les mêmes fréquences sub-6 GHz, leurs communications peuvent interférer. Supposons que le nombre d'utilisateurs associés à la sMBS est inférieur au nombre d'antennes de transmission de la sMBS A_0 , le SINR reçu par l'utilisateur k à partir de la sSBS n et celui à partir de la sMBS peuvent être exprimés respectivement comme suit (Bethanabhotla *et al.*, 2014), pour tout $k \in \mathcal{K}$ et $n \in \mathcal{N}$:

$$\text{SINR}_{kn} = \frac{\gamma_n g'_{kn}}{1 + \sum_{\substack{n' \in \mathcal{N} \setminus \{0\}, \\ n' \neq n}} \gamma_{n'} g'_{kn'} + (A_0 - s'_0 + 1) \gamma_0 g'_{k0} / A_0}, \quad (2.3)$$

$$\text{SINR}_{k0} = \frac{A_0 - s'_0 + 1}{A_0} \frac{\gamma_0 g'_{k0}}{1 + \sum_{n \in \mathcal{N}, n \neq 0} \gamma_n g'_{kn}}, \quad (2.4)$$

où g'_{kn} est le gain du canal entre l'utilisateur k et la sSBS n et g'_{k0} est le gain entre l'utilisateur k et la sMBS, γ_n est le rapport signal sur bruit (en anglais : *signal-to-noise ratio*, SNR) reçu par l'utilisateur k de la sSBS et γ_0 de la sMBS. Le rapport $\frac{A_0 - s'_0 + 1}{A_0}$ est une approximation du gain

des antennes utilisant la technologie MIMO. Nous notons que, dans les équations (2.3) et (2.4), le calcul de l'interférence est basé sur le scénario le plus défavorable, c'est-à-dire lorsque toutes les sBSs transmettent simultanément, comme dans (Pollakis *et al.*, 2012; Parsaeefard et Sharafat, 2012). De plus, aucune interférence n'est considérée entre les communications mSBSs et sBSs car les deux types de stations de base utilisent deux bandes de fréquences différentes.

Dans le cadre des communications mmWave, l'interférence peut être négligée, ce qui permet de considérer le SNR à la place du SINR (Elshaer *et al.*, 2016). Le SNR reçu par l'utilisateur k à partir de la mSBS m est donné par, pour tout $k \in \mathcal{K}$ et $m \in \mathcal{M}$:

$$\text{SNR}_{km} = \frac{p_m g_{km} PL^{-1}(d_{km})}{\sigma^2}, \quad (2.5)$$

où σ^2 est la puissance du bruit, p_m est la puissance de transmission de la mSBS m , $PL^{-1}(d_{km})$ est la fonction d'affaiblissement de propagation entre l'utilisateur k et la mSBS m qui sont à une distance de d_{km} l'un de l'autre et g_{km} est le gain d'antenne de la mSBS m . Lorsque l'utilisateur k est associé, le débit de données possible (en bits par seconde par hertz) est calculé en se basant sur le théorème de Shannon :

$$R_{km} = \log_2(1 + \text{SNR}_{km}), \quad \forall k \in \mathcal{K}, m \in \mathcal{M}, \quad (2.6)$$

$$R'_{kn} = \log_2(1 + \text{SINR}_{kn}), \quad \forall k \in \mathcal{K}, n \in \mathcal{N}. \quad (2.7)$$

Chaque station de base alloue aux utilisateurs associés un certain nombre de ressources qui sont définies par des créneaux temporels. Le nombre de créneaux temporels alloués à chaque utilisateur dépend de ses exigences en termes de débit de données et de la station de base à laquelle il est associé. Nous notons, respectivement, par W et W' les largeurs de bande de fréquences sub-6 GHz et celles de mmWave allouées pour un créneau temporel. Les contraintes de qualité de service garantissant que le débit reçu par l'utilisateur k associé à la mSBS m ou à la sBS n soit supérieur ou égal au débit minimal Q_k sont exprimées comme suit, pour tout $k \in \mathcal{K}$, $m \in \mathcal{M}$ et $n \in \mathcal{N}$:

$$Q_k x_{km} \leq b_{km} W R_{km}, \quad (2.8)$$

$$Q_k x'_{kn} \leq b'_{kn} W' R'_{kn}. \quad (2.9)$$

Par conséquent, les nombres de créneaux temporels, b_{km} et b'_{kn} , requis par l'utilisateur k sont calculés comme suit, pour tout $k \in \mathcal{K}$, $m \in \mathcal{M}$ et $n \in \mathcal{N}$:

$$b_{km} = \left\lceil \frac{Q_k}{WR_{km}} \right\rceil, \quad (2.10)$$

$$b'_{kn} = \left\lceil \frac{Q_k}{W'R'_{kn}} \right\rceil, \quad (2.11)$$

où $\lceil \cdot \rceil$ fait référence à la fonction partie entière supérieure (ou la fonction plafond). Finalement, le nombre de créneaux temporels alloués par toutes les stations de base aux utilisateurs associés est égal à :

$$\sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{M}} b_{km} x_{km} + \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} b'_{kn} x'_{kn}. \quad (2.12)$$

2.4 Formulation du problème

Dans ce qui suit, UARAP est formulé dont l'objectif est de maximiser le nombre d'utilisateurs associés et d'assurer une utilisation efficace des ressources en minimisant le nombre de créneaux temporels alloués. Pour ce faire, nous introduisons respectivement les deux paramètres d'ajustement $\rho \in [0, 1]$ et $\rho' \in [0, 1]$ pour les communications sub-6 GHz et mmWave. Ces paramètres permettent un équilibre entre le nombre de créneaux temporels alloués et le nombre d'utilisateurs associés. Plus précisément, plus ces deux derniers paramètres sont proches de 1 (resp. 0), plus l'objectif est de maximiser le nombre d'utilisateurs associés (resp. de minimiser le nombre de créneaux temporels utilisés). Dans la plupart des cas, pour une instance donnée de UARAP, il peut y avoir plusieurs solutions qui atteignent le même nombre maximal d'utilisateurs associés. Ces solutions allouent différents nombres de créneaux temporels, en fonction de ρ et ρ' . Il existe donc des valeurs optimales de ρ^* et ρ'^* , si ρ (resp. ρ') est dans l'intervalle $[\rho^*, 1[$ (resp. dans l'intervalle $[\rho'^*, 1[$), le nombre optimal d'utilisateurs associés et de ressources utilisées peuvent être obtenus. Des problèmes similaires ont été étudiés dans (Sokun *et al.*, 2015; Matskani *et al.*, 2008) sans prendre en compte les communications à bandes multiples. UARAP est formulé comme suit :

$$\begin{aligned} \underset{\mathbf{X}, \mathbf{X}'}{\text{maximiser}} \quad & \sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{M}} (\rho^* x_{km} - (1 - \rho^*) b_{km} x_{km}) \\ & + \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} (\rho'^* x'_{kn} - (1 - \rho'^*) b'_{kn} x'_{kn}) \end{aligned} \quad (\text{P1a})$$

$$\text{sous les contraintes:} \quad x_{km}, x'_{kn} \in \{1, 0\}, \forall k \in \mathcal{K}, m \in \mathcal{M}, n \in \mathcal{N}, \quad (\text{P1b})$$

$$\sum_{k \in \mathcal{K}} b_{km} x_{km} \leq t_m, \forall m \in \mathcal{M}, \quad (\text{P1c})$$

$$\sum_{k \in \mathcal{K}} b'_{kn} x'_{kn} \leq t'_n, \forall n \in \mathcal{N}, \quad (\text{P1d})$$

$$\sum_{m \in \mathcal{M}} x_{km} + \sum_{n \in \mathcal{N}} x'_{kn} \leq 1, \forall k \in \mathcal{K}, \quad (\text{P1e})$$

$$\sum_{k \in \mathcal{K}} x_{km} \leq s_m, \forall m \in \mathcal{M}, \quad (\text{P1f})$$

$$\sum_{k \in \mathcal{K}} x'_{kn} \leq s'_n, \forall n \in \mathcal{N}. \quad (\text{P1g})$$

Les contraintes (P1b) garantissent que les variables d'association x_{km} et x'_{kn} soient binaires pour tout triplet (k, m, n) . Les contraintes (P1c) et (P1d) imposent que le nombre de créneaux temporels utilisés ne dépasse pas les seuils t_m et t'_n . Les contraintes (P1e) imposent que chaque utilisateur soit associé exactement à une seule station de base. Enfin, les contraintes (P1f) et (P1g) limitent le nombre d'utilisateurs associés à la mSBS m et à la sBS n à s_m et s'_n , respectivement. En pratique, cette limitation est liée au nombre d'antennes disponibles dans la station de base. Nous notons que lorsque $\rho = \rho' = 1$, la fonction objectif du problème (P1) se réduit à maximiser le nombre d'utilisateurs associés. Dans le reste de ce chapitre, nous utilisons $\underline{s} = [s_1, \dots, s_M]$, $\underline{s}' = [s'_0, s'_1, \dots, s'_N]$, $\underline{t} = [t_1, \dots, t_M]$ et $\underline{t}' = [t'_0, t'_1, \dots, t'_N]$ pour les notations vectorielles. Les matrices des créneaux temporels calculées par les mSBSs et les sBSs sont respectivement notées par $\mathbf{B} = [b_{km}]$ et $\mathbf{B}' = [b'_{kn}]$. En outre, la i -ème ligne (resp. la i -ème colonne) de la matrice \mathbf{A} est notée par \mathbf{A}_i (resp. \mathbf{A}^i) et la somme des éléments du vecteur \mathbf{X}^m est notée par $|\mathbf{X}^m|$.

2.5 NP-difficulté

Nous prouvons ci-dessous que UARAP est un problème NP-difficile en utilisant la technique de réduction. Par conséquent, nous démontrons que le problème du sac à dos multiple (en anglais : *multiple knapsack problem*, MKP) qui est un cas particulier du problème d'affectation généralisé (en anglais : *generalized assignment problem*) de complexité NP-difficile, peut se réduire en temps polynomial à UARAP (Maniezzo *et al.*, 2021; Chekuri et Khanna, 2005).

Théorème 2.5.1. *UARAP est NP-difficile.*

Démonstration : Considérons un ensemble d'objets et un ensemble de sacs. Chaque sac j a une

capacité maximale c_j . Pour chaque objet i et chaque sac j un poids w_{ij} est attribué. L'objectif est de répartir le plus grand nombre possible d'objets dans les sacs d'une manière que chaque sac ne dépasse pas sa capacité maximale (Chekuri et Khanna, 2005).

On peut créer une instance de UARAP à partir d'une instance de MKP en définissant $\mathcal{N} = \emptyset$, $\rho = \rho' = 1$, $s'_n = 0$, $s_m = M$, $\forall n \in \mathcal{N}$ et $m \in \mathcal{M}$. Avec ces restrictions, UARAP revient à maximiser le nombre d'utilisateurs associés aux mSBSs tout en respectant les contraintes des créneaux temporels. Par analogie, si les mSBSs représentent les sacs, les utilisateurs représentent les objets, le poids représente le nombre de créneaux temporels alloués et la capacité de chaque sac correspond au nombre de créneaux temporels disponibles, alors le problème (P1) devient équivalent à MKP. Puisque MKP est un problème NP-difficile, UARAP est également NP-difficile. ■

Étant donné que le problème UARAP est NP-difficile, il est impossible, sous l'hypothèse $P \neq NP$, de le résoudre en temps polynomial. Dans la section suivante, nous proposons des algorithmes peu complexes, centralisés et distribués. Pour des fins de comparaison, nous utilisons le solveur IBM ILOG CPLEX sous MATLAB pour résoudre le problème (P1) et obtenir des résultats optimaux. Ces résultats sont par la suite notés par OPT.

2.6 Solutions algorithmiques

Pour résoudre UARAP, nous proposons un algorithme glouton centralisé appelé C-GUA et deux algorithmes gloutons distribués appelés D-NBS et D-RLA.

2.6.1 Un algorithme centralisé (C-GUA)

Le pseudo-code de C-GUA est donné dans l'algorithme 1. Afin d'assurer une utilisation efficace de ressources, C-GUA commence par associer les utilisateurs demandant le moins de créneaux temporels. De plus, C-GUA privilégie les communications mmWave en raison des larges bandes passantes qu'elles offrent. Dans C-GUA, nous supposons l'existence d'un contrôleur central qui dispose des informations globales sur le nombre d'utilisateurs, le nombre de stations de base et l'état du canal (en anglais : *channel state information*, CSI) entre les utilisateurs et les stations de base. C-GUA initialise les matrices d'association \mathbf{X} et \mathbf{X}' à des matrices nulles, ainsi que le compteur k à 0. La boucle répéter au niveau de la ligne 3 commence le processus d'association. Ensuite, C-GUA

trie par ordre croissant le nombre de créneaux temporels requis pour associer l'utilisateur k . Les indices des mSBSs et sBSs sont stockés respectivement dans les vecteurs \underline{w} et \underline{w}' . Ensuite, C-GUA vérifie si les créneaux temporels disponibles sont suffisants et si le nombre maximal d'utilisateurs associés n'est pas encore atteint. Après chaque association, les ressources restantes et le nombre d'utilisateurs associés sont mis à jour. La boucle tant que se termine lorsque C-GUA trouve une association possible ou lorsqu'aucune station de base disponible n'est trouvée. C-GUA retourne les deux matrices d'association lorsque : (i) il n'y a plus de créneaux temporels disponibles, ou (ii) tous les utilisateurs sont associés, ou (iii) le nombre maximal d'utilisateurs associés à chaque station de base est atteint.

Algorithme 1 Pseudo-code de l'algorithme C-GUA

Initialiser : $K, M, N, \underline{s}, \underline{s}', \underline{t}, \underline{t}', \mathbf{B}, \mathbf{B}'$.

```

1:  $\mathbf{X} \leftarrow \mathbf{0}, \mathbf{X}' \leftarrow \mathbf{0}$ .
2:  $k \leftarrow 0$ .
3: répéter
4:    $k \leftarrow k + 1$ .
5:    $\underline{r} \leftarrow$  trier  $\mathbf{B}_k$  en ordre croissant.
6:    $\underline{r}' \leftarrow$  trier  $\mathbf{B}'_k$  en ordre croissant.
7:    $m \leftarrow 1, n \leftarrow 0$ .
8:    $\zeta \leftarrow 0$ .
9:   Tant que  $n \leq N$  et  $m \leq M$  et  $\zeta = 0$  faire
10:     $\underline{w} \leftarrow$  trouver  $(\mathbf{B}_k = r_m)$ .
11:     $\underline{w}' \leftarrow$  trouver  $(\mathbf{B}'_k = r'_n)$ .
12:    répéter
13:       $m_0 \leftarrow w_1, n_0 \leftarrow w'_1$ .
14:      Si  $r_m \leq r'_n$  et  $t_{m_0} \geq b_{km_0}$  et  $|\mathbf{X}^{m_0}| < s_{m_0}$  alors
15:         $x_{km_0} \leftarrow 1$ .
16:         $t_{m_0} \leftarrow t_{m_0} - b_{km_0}$ .
17:         $\zeta \leftarrow 1$ , sortir.
18:      Sinon Si  $r_m \geq r'_n$  et  $t'_{n_0} \geq b'_{kn_0}$  et  $|\mathbf{X}'^{n_0}| < s'_{n_0}$  alors
19:         $x'_{kn_0} \leftarrow 1$ .
20:         $t'_{n_0} \leftarrow t'_{n_0} - b'_{kn_0}$ .
21:         $\zeta \leftarrow 1$ , sortir.
22:      Sinon
23:        Supprimer  $w_1$  de  $\underline{w}$  et  $w'_1$  de  $\underline{w}'$ .
24:      Fin Si
25:    jusqu'à  $\underline{w} = [ ]$  et  $\underline{w}' = [ ]$ 
26:     $n \leftarrow n + 1, m \leftarrow m + 1$ .
27:  Fin Tant que
28: jusqu'à  $(\underline{t} = \underline{t}' = \underline{0})$  ou  $(k = K)$  ou  $(|\mathbf{X}^m| = s_m, \forall m \in \mathcal{M}$  et  $|\mathbf{X}'^n| = s'_n, \forall n \in \mathcal{N})$ 
29: Retourner  $\mathbf{X}, \mathbf{X}'$ .

```

Le calcul de la complexité asymptotique de C-GUA est le suivant. La boucle répéter de la ligne 3 nécessite $\mathcal{O}(K)$ itérations. La boucle répéter de la ligne 12 nécessite $\mathcal{O}(L_1)$ itérations avec $L_1 = \mathbf{max}(M, N)$. Les deux opérations trouver nécessitent $\mathcal{O}(M + N)$ itérations. Par conséquent, la complexité de la boucle tant que de la ligne 9 est $\mathcal{O}(L_2(M + N + L_1))$ avec $L_2 = \mathbf{min}(M, N)$. Les deux opérations trier nécessitent $\mathcal{O}(M \log M + N \log N)$ itérations. Par conséquent, la complexité dans le pire des cas de C-GUA est $\mathcal{O}(K(M \log M + N \log N + L_2(M + N + L_1)))$.

C-GUA nécessite des informations globales sur le réseau, ce qui rend indispensable la présence d'un contrôleur central pour coordonner les associations des stations de base. Afin de limiter la dépendance à un tel point central, qui est généralement un goulot d'étranglement, des algorithmes distribués sont proposés dans la section suivante.

2.6.2 Des algorithmes distribués (D-NBS et D-RLA)

Tout au long de cette section, nous supposons que chaque station de base connaît seulement le CSI entre elle-même et tous les utilisateurs (Xu et Mao, 2017). Autrement dit, elle ne connaît pas le CSI entre les autres stations de base et les utilisateurs. De plus, chaque station de base peut calculer et diffuser le nombre de créneaux temporels nécessaires pour satisfaire les exigences de débit de données de chaque utilisateur. Enfin, un scénario synchronisé, où tous les utilisateurs envoient leurs demandes d'association en même temps, est considéré.

Par la suite, nous proposons deux algorithmes distribués au niveau des utilisateurs. Le premier, appelé D-NBS (en anglais : *distributed nearest base station*), est basé sur une approche gloutonne et proposé pour des raisons de comparaison. Le deuxième, appelé D-RLA (en anglais : *distributed reinforcement learning algorithm*), est basé sur une technique d'apprentissage par renforcement.

D-NBS

Dans D-NBS, chaque utilisateur suit les étapes suivantes :

1. Il diffuse sa demande en termes de qualité de service.
2. Il choisit et envoie une demande d'association à la station de base ayant la meilleure condition de canal, c'est-à-dire à la station de base qui offre le nombre minimal de créneaux temporels.

La complexité asymptotique de D-NBS dans le pire des cas est $\mathcal{O}(N + M)$. Comme mentionné

précédemment, D-NBS n'est utilisé que pour des fins de comparaison. D-RLA est détaillé ci-dessous.

D-RLA

L'apprentissage par renforcement est une technique d'apprentissage automatique inspirée du comportement d'apprentissage des êtres vivants. Cette technique ne nécessite aucun modèle préalable de l'environnement (Watkins et Dayan, 1992). Elle permet à un agent d'apprendre à choisir des actions en fonction des récompenses ou des pénalités qu'il reçoit. L'algorithme D-RLA est basé sur la technique d'apprentissage par renforcement Q-learning (QL). L'idée de D-RLA est d'apprendre à chaque utilisateur à partir de son expérience passée la meilleure association de manière automatique et indépendamment des autres utilisateurs. Le modèle d'apprentissage proposé est défini par :

- Agents : L'ensemble d'utilisateurs \mathcal{K} , où chaque utilisateur est un agent indépendant.
- États : Soit \mathcal{V} l'ensemble d'états. Chaque état $v \in \mathcal{V}$ est un 2-uplet (i, i') , où i (resp. i') désigne l'intervalle des créneaux temporels disponibles pour les mSBSs (resp. les sBSs). Autrement dit, le nombre de créneaux temporels disponibles pour chaque mSBS (resp. chaque sBS) est divisé en un nombre d'intervalles I (resp. I'). Les ensembles des intervalles sont notés par \mathcal{I} et \mathcal{I}' . L'ensemble d'états \mathcal{V} peut être écrit sous la forme du produit cartésien des ensembles \mathcal{I} et \mathcal{I}' , c'est-à-dire $\mathcal{V} = \mathcal{I} \times \mathcal{I}'$. Par exemple, si $I = I' = 3$ et $t_m = t'_n = 6$ pour tout $m \in \mathcal{M}$ et $n \in \mathcal{N}$, alors $\mathcal{I} = \mathcal{I}' = \{[0; 0], [1; 3], [4; 6]\}$ et $\mathcal{V} = \{([0; 0], [0; 0]), ([0; 0], [1; 3]), ([0; 0], [4; 6]), ([1; 3], [0; 0]), ([1; 3], [1; 3]), ([1; 3], [4; 6]), ([4; 6], [0; 0]), ([4; 6], [1; 3]), ([4; 6], [4; 6])\}$. Nous notons que l'intervalle $[0; 0]$ est utilisé quand le nombre de créneaux temporels nécessaires à allouer dépasse le seuil t_m ou le seuil t_n .
- Actions : Soit $\mathcal{A} = \{a_1, a_2, a_3\}$ l'ensemble d'actions, où a_1 (resp. a_2) désigne une demande d'association à une mBS (resp. à une sBS) et a_3 désigne une action de silence. Nous notons que si les nombres minimaux de créneaux temporels requis pour les communications sub-6 GHz et mmWave dépassent les seuils, qui sont supposés connus par l'utilisateur, ce dernier choisit a_3 et aucune demande d'association n'est envoyée.
- Fonction de récompense : Soit $f_k : \mathcal{A} \mapsto \mathbb{R}$ la fonction de récompense de l'utilisateur k lorsqu'il choisit une action à partir de l'ensemble $\mathcal{A} \setminus \{a_3\}$. Il est à noter que la fonction de récompense proposée vise à trouver une politique qui maximise le nombre d'utilisateurs associés et minimise simultanément le nombre de créneaux temporels utilisés. Si l'action

choisie conduit à l'acceptation de la demande d'association, une récompense positive est considérée, tandis qu'une pénalité est accordée si l'action est refusée. La mSBS et la sBS choisies par l'utilisateur sont notées par la suite m^* et n^* , respectivement. Soit $\epsilon \in \{-1, 1\}$, la fonction de récompense lorsque l'utilisateur $k \in \mathcal{K}$ choisit une action $a \in \mathcal{A} \setminus \{a_3\}$ est comme suit :

$$f_k(a) = \begin{cases} \epsilon \tau_{m^*} / s_{m^*} - \epsilon b_{km^*} / t_{m^*} + \epsilon \Omega, & \text{si } a = a_1 \\ \epsilon \tau'_{n^*} / s'_{n^*} - \epsilon b'_{kn^*} / t'_{n^*} + \epsilon \Omega, & \text{si } a = a_2 \end{cases}, \quad (2.13)$$

où τ_{m^*} et τ'_{n^*} sont respectivement le nombre d'utilisateurs associés à m^* et n^* et Ω est une constante qui est ajoutée pour assurer des récompenses positives. Dans (2.13), $\epsilon = 1$ si et seulement si la demande d'association est acceptée et $\epsilon = -1$ sinon.

- Fonction action-valeur Q : La fonction action-valeur $Q(k, v, a_j)$ est calculée comme suit (Watkins et Dayan, 1992), pour tout $v \in \mathcal{V}$, $k \in \mathcal{K}$ et $j \in \{1, 2\}$:

$$Q(k, v, a_j) = (1 - \delta)Q(k, v, a_j) + \delta f_k(a_j), \quad (2.14)$$

où δ est le facteur d'apprentissage. Les valeurs état-action sont mémorisées dans une table, appelée la table Q. Dans ce chapitre, les utilisateurs sont considérés myopes, car seules les récompenses immédiates sont considérées.

L'objectif de D-RLA est d'associer chaque utilisateur en utilisant les valeurs de la table Q, à une mSBS si $Q(k, v, a_1) \geq Q(k, v, a_2)$ et à une sBS sinon. La taille de la table Q est $\overline{\mathcal{V}} \cdot \overline{\mathcal{A}}$, où $\overline{\mathcal{V}}$ et $\overline{\mathcal{A}}$ sont les cardinalités des ensembles \mathcal{V} et \mathcal{A} , respectivement. Le pseudo-code de D-RLA est donné dans l'algorithme 2.

À chaque itération, chaque utilisateur sélectionne son état actuel en associant le nombre minimal de créneaux temporels envoyés par les mSBSs et les sBSs aux intervalles correspondants. Ensuite, l'utilisateur sélectionne la meilleure action qui maximise sa récompense. L'action a_3 est choisie si les nombres de créneaux temporels envoyés par toutes les stations de base sont supérieurs aux seuils. Enfin, D-RLA met à jour la table Q. La boucle tant que est répétée jusqu'à ce que le nombre maximal d'itérations, noté par Y , soit atteint. La complexité asymptotique de D-RLA dans le pire des cas est $\mathcal{O}(Y(N + M))$.

Algorithme 2 Pseudo-code de l'algorithme D-RLA

Initialiser : δ , utilisateur k , Y .

- 1: $\mathbf{Q} \leftarrow \mathbf{0}$.
 - 2: $y \leftarrow 0$.
 - 3: **Tant que** $y \leq Y$ **faire**
 - 4: Recevoir le nombre de créneaux requis de toutes les stations de base.
 - 5: Trouvez les deux stations de base, m^* et n^* , qui offrent le nombre minimal de créneaux temporels.
 - 6: **Si** $b_{km^*} \leq t_{m^*}$ **ou** $b'_{kn^*} \leq t'_{n^*}$ **alors**
 - 7: Trouver l'état actuel $v \leftarrow (i, i') \in \mathcal{I} \times \mathcal{I}'$.
 - 8: Choisir une action $a^* \leftarrow \arg \max_a Q(k, v, a)$.
 - 9: Envoyer une demande d'association à la station de base m^* ou n^* choisie.
 - 10: Calculer $f_k(a^*)$.
 - 11: Mettre $Q(\cdot)$ à jour comme $Q(k, v, a^*) \leftarrow (1 - \delta)Q(k, v, a^*) + \delta f_k(a^*)$.
 - 12: **Sinon**
 - 13: $a^* \leftarrow a_3$.
 - 14: **Fin Si**
 - 15: $y \leftarrow y + 1$.
 - 16: **Fin Tant que**
-

À chaque fois que les algorithmes D-NBS et D-RLA sont exécutés, chaque station de base prend une décision en exécutant l'algorithme D-BSA. Elle peut accepter ou refuser les demandes d'association envoyées par les utilisateurs, en fonction de sa capacité. Le pseudo-code de D-BSA est donné dans l'algorithme 3, où les mêmes notations que dans l'algorithme 1 sont utilisées.

Algorithme 3 Pseudo-code de l'algorithme D-BSA

Initialiser : \mathbf{B}^m (ou \mathbf{B}'^n), t_m (ou t'_n), s_m (ou s'_n), $\forall m \in \mathcal{M}$ (ou $n \in \mathcal{N}$).

- 1: $k \leftarrow 0$.
 - 2: $somme \leftarrow 0$.
 - 3: $r \leftarrow$ trier \mathbf{B}^m (ou \mathbf{B}'^n) en ordre croissant.
 - 4: **Tant que** $k \leq s_m$ (ou s'_n) **et** $somme \leq t_m$ (ou t'_n) **faire**
 - 5: $somme \leftarrow somme + r_k$.
 - 6: Accepter la demande d'association.
 - 7: $k \leftarrow k + 1$.
 - 8: **Fin Tant que**
-

Afin de maximiser le nombre d'utilisateurs associés et d'optimiser les ressources allouées, D-BSA trie en ordre croissant le nombre de créneaux temporels requis, noté par \mathbf{B}^m pour tout $m \in \mathcal{M}$ (ou \mathbf{B}'^n pour tout $n \in \mathcal{N}$). Ensuite, la boucle tant que itère le processus d'association jusqu'à ce que le seuil des créneaux temporels (t_m pour la mSBS m ou t'_n pour la sBS n) ou le nombre maximal d'utilisateurs associés (s_m pour la mSBS m ou s'_n pour la sBS n) soient atteints.

Dans le pire des cas, la complexité asymptotique de D-BSA est $\mathcal{O}(\min(s_m, t_m))$ pour tout $m \in \mathcal{M}$ (resp. $\mathcal{O}(\min(s'_n, t'_n))$ pour tout $n \in \mathcal{N}$). La complexité des algorithmes proposés dans ce chapitre sont résumées dans le tableau 2.2.

TABLE 2.2 – Complexité asymptotique des algorithmes.

Algorithme	Complexité
C-GUA	$\mathcal{O}(K(M \log M + N \log N + L_2(M + N + L_1)))$
D-NBS	$\mathcal{O}(N + M)$
D-BSA	$\mathcal{O}(\min), \forall m \in \mathcal{M}, \mathcal{O}(\min(s'_n, t'_n)), \forall n \in \mathcal{N}$
D-RLA	$\mathcal{O}(Y(N + M))$

L'idée de D-RLA nécessite un échange d'informations entre les utilisateurs et les stations de base. Chaque station de base doit connaître le CSI de tous les utilisateurs, leurs exigences en termes de débit de données et leurs demandes d'association. Ainsi, les utilisateurs doivent connaître le nombre de créneaux temporels disponibles pour chaque station de base, le nombre de créneaux temporels calculés pour établir une association avec toutes les stations de base et la décision de chaque demande d'association. Les échanges de messages entre les utilisateurs et les stations de base sont envoyés sous forme de signaux pilotes et sont généralement transmis rapidement.

Dans D-RLA, les utilisateurs apprennent d'une manière autonome et indépendante à choisir la meilleure action dans chaque état. Ils envoient les demandes d'association et reçoivent des récompenses en conséquence. Il est important de noter que l'envoi d'une demande d'association à une station de base ne garantit pas l'acceptation de la demande. Ceci signifie qu'un utilisateur peut se trouver dans la zone de couverture de plusieurs stations de base. Par conséquent, cet utilisateur peut envoyer une demande d'association à plusieurs stations de base. La courbe de convergence de D-RLA est présentée dans la section suivante.

2.7 Résultats de simulation

Dans cette section, nous présentons les résultats de simulation Monte-Carlo pour illustrer les performances des algorithmes proposés. Nous considérons un réseau hétérogène composé d'une sMBS, de M mSBSs et de N sSBSs. La sMBS est située au centre d'une zone carrée de 300 m·300 m et les positions des utilisateurs et des stations de base sont distribuées d'une manière uniforme. Les gains du canal g'_{kn} et g'_{k0} sont modélisés comme dans (Bethanabhotla *et al.*, 2014; Xu et Mao, 2017) et

sont donnés par, pour tout $k \in \mathcal{K}$ et $n \in \mathcal{N}$:

$$g'_{kn} = |h_{kn}|^2 \beta \left(\frac{d_{kn}}{d_0} \right)^{-\alpha}, \quad (2.15)$$

$$g'_{k0} = \left(1 + \left(\frac{d_{k0}}{d'_0} \right)^\alpha \right)^{-\alpha}, \quad (2.16)$$

où $|h_{kn}|^2$ est le coefficient d'évanouissement de Rayleigh entre l'utilisateur k et la sSBS n , d_0 et d'_0 sont des distances de référence. Ce dernier est modélisé comme une variable aléatoire complexe gaussienne de moyenne nulle et de variance unitaire. Enfin, β est une constante capturant les effets du système et de la transmission et d_{kn} (resp. d_{k0}) est la distance euclidienne entre l'utilisateur k et la sSBS n (resp. la sMBS). L'affaiblissement de la puissance du signal est calculé comme dans (Akdeniz *et al.*, 2014) : $10 \log_{10}(PL(d_{km})) = \psi + 10\kappa \log_{10}(d_{km}) + \xi$, où ξ est l'ombrage représenté par une variable aléatoire qui suit la loi log-normale de moyenne nulle et de variance v^2 . Le gain d'antennes entre la mSBS m et les utilisateurs est égal au nombre d'antennes A_m . Le facteur d'apprentissage δ est égal à 0,5. Les valeurs des paramètres ainsi que les notations sont données dans le tableau 2.1. Nous comparons les performances des algorithmes C-GUA, D-NBS et D-RLA à celles de OPT et de l'algorithme proposé dans (Zalghout *et al.*, 2017). La solution de (Zalghout *et al.*, 2017) est adoptée à notre contexte afin de pouvoir l'utiliser comme une référence de comparaison.

Tout d'abord, nous nous concentrons sur la maximisation du nombre d'utilisateurs associés. Pour ce faire, nous fixons les valeurs des paramètres d'ajustement ρ et ρ' à 1 dans les figures 2.2, 2.3 et 2.4, ce qui signifie que le nombre de créneaux temporels alloués n'est pas optimisé.

La figure 2.2 illustre le pourcentage d'utilisateurs associés en fonction du nombre de SBSs déployées pour $K = 100$ et $N = M$. Il est clair que lorsque le nombre de SBSs augmente, plus d'utilisateurs sont associés. En outre, le pourcentage d'utilisateurs associés est presque égal à 16% pour $M = N = 0$, car seule l'association à la sMBS est considérée. Cette figure montre aussi que C-GUA offre des performances proches de celles de OPT avec une complexité de calcul largement inférieure. De plus, pour un petit nombre de SBSs, les résultats de D-RLA sont très proches de ceux de C-GUA. Cette différence est due au fait que C-GUA est centralisé alors que D-RLA est distribué.

La figure 2.3 montre les pourcentages optimaux d'utilisateurs associés et de créneaux temporels alloués pour $K = 100$ en fonction du nombre de créneaux temporels disponibles. Comme prévu,

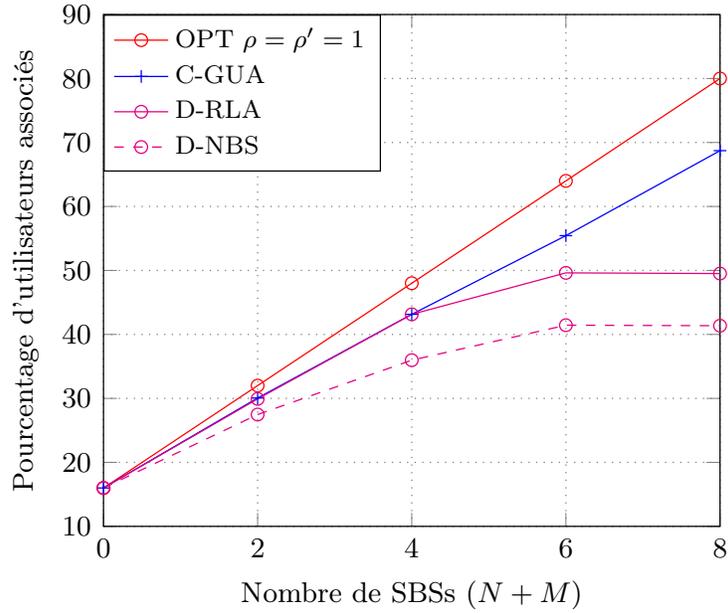


FIGURE 2.2 – Pourcentage d'utilisateurs associés en fonction du nombre de SBSs.

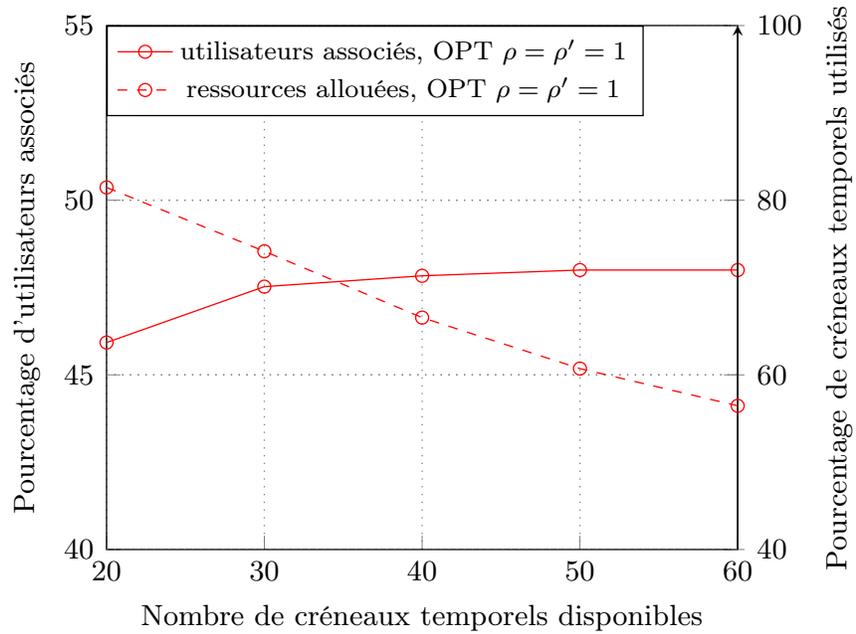


FIGURE 2.3 – Pourcentage d'utilisateurs associés et de créneaux temporels utilisés en fonction du nombre de créneaux temporels disponibles pour chaque station de base.

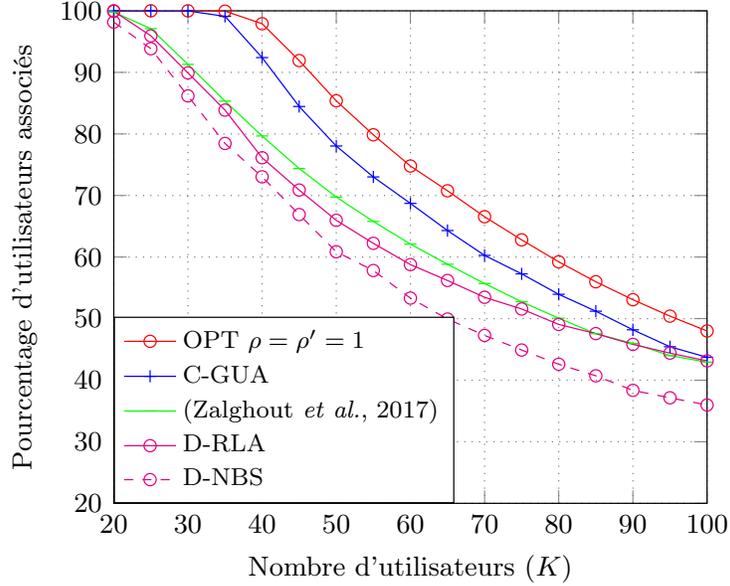


FIGURE 2.4 – Pourcentage d'utilisateurs associés en fonction du nombre d'utilisateurs.

si plus de créneaux sont disponibles, plus d'utilisateurs sont associés. De plus, lorsque le nombre de créneaux disponibles dépasse 50, le pourcentage d'utilisateurs associés demeure relativement constant à 48%, ce qui est dû aux contraintes (P1f) et (P1g) qui limitent le nombre d'utilisateurs associés à chaque station de base. À partir de la figure 2.3, nous constatons qu'il existe un nombre optimal de créneaux temporels disponibles pour chaque station de base, pour lequel OPT associe le plus grand nombre d'utilisateurs. Par conséquent, le nombre de créneaux temporels disponibles est fixé par la suite à $t_m = t_n = 50$.

La figure 2.4 illustre le pourcentage d'utilisateurs associés pour différentes valeurs de K . Il est clair que plus K augmente, plus le pourcentage d'utilisateurs associés diminue. Ceci est conforme aux contraintes (P1c) et (P1d) relatives au nombre de créneaux temporels disponibles et les contraintes (P1f) et (P1g) relatives au nombre d'utilisateurs associés à chaque station de base. De plus, C-GUA est plus performant que l'algorithme heuristique proposé dans (Zalghout *et al.*, 2017). Ceci est dû au fait que, contrairement à l'algorithme de (Zalghout *et al.*, 2017) qui commence par associer les utilisateurs les plus proches des stations de base, C-GUA commence par associer les utilisateurs en fonction du nombre de créneaux temporels demandés. La figure 2.4 montre que, pour de grandes valeurs de K , OPT est légèrement plus performant que C-GUA et D-RLA. Aussi, D-RLA est toujours plus performant que D-NBS. L'écart de performances entre D-RLA et C-GUA est très faible et

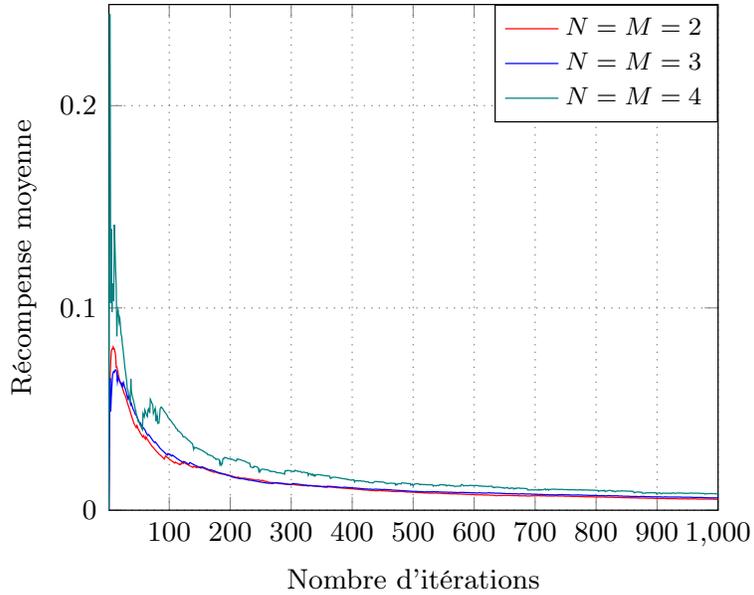


FIGURE 2.5 – Convergence de D-RLA.

ne dépasse pas 5%. Ces résultats illustrent l'efficacité de la technique d'apprentissage proposée et montrent le gain en termes de pourcentage d'utilisateurs associés. Même si les performances des algorithmes proposés ne sont pas analytiquement tractables, les résultats de simulation indiquent qu'elles se rapprochent de l'optimal.

La figure 2.5 illustre l'impact du nombre de SBSs sur le temps de convergence de D-RLA. On peut voir que, pour différents nombres de SBSs, D-RLA converge à partir de 400 itérations. Nous pouvons conclure que le nombre de SBSs n'a pas d'impact significatif sur le temps de convergence de D-RLA.

La figure 2.6 et la figure 2.7 illustrent l'impact des paramètres d'ajustement, ρ et ρ' , sur le pourcentage d'utilisateurs associés et le pourcentage de créneaux temporels alloués en fonction du nombre d'utilisateurs.

La figure 2.6 montre que, plus ρ et ρ' se rapprochent de 1, plus le pourcentage d'utilisateurs associés augmente. De plus, nous constatons que même lorsque $\rho = \rho' = 0.99$, le pourcentage d'utilisateurs associés par OPT est pratiquement le même que lorsque $\rho = \rho' = 1$, quelle que soit la valeur de K .

La figure 2.7 montre le pourcentage de créneaux temporels alloués en fonction du nombre d'utilisateurs. Lorsque ρ et ρ' sont très proches de 1 (c'est-à-dire 0,99), le pourcentage de créneaux temporels

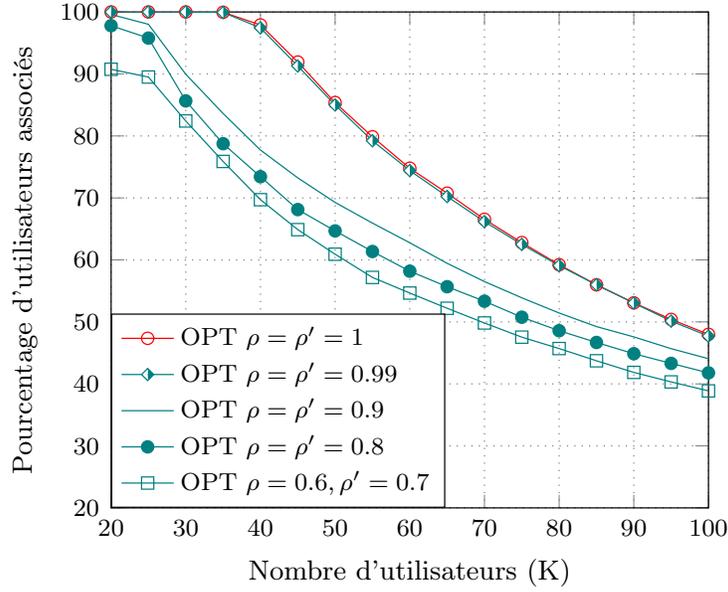


FIGURE 2.6 – Pourcentage d'utilisateurs associés en fonction du nombre d'utilisateurs.

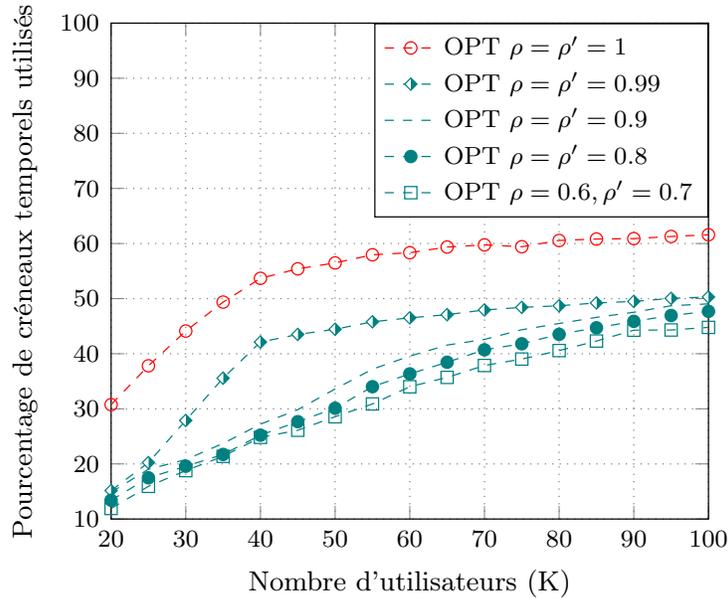


FIGURE 2.7 – Pourcentage de créneaux temporels utilisés en fonction du nombre d'utilisateurs.

alloués par OPT pour un grand nombre d'utilisateurs ne dépasse pas 51% alors que lorsqu'ils sont égaux à 1, il dépasse 60%. Nous pouvons conclure que, lorsque ρ et ρ' sont égaux à 0.99, OPT associe presque le nombre maximal d'utilisateurs tout en allouant moins de créneaux temporels que lorsque ceux-ci sont égaux à 1. Par conséquent, les valeurs optimales de ρ et ρ' , notées respectivement par ρ^* et ρ'^* , sont choisies égales à 0.99.

2.8 Conclusions

Dans ce chapitre, nous avons étudié le problème conjoint d'association des utilisateurs aux stations de base et d'allocation des ressources dans un réseau cellulaire hétérogène à bandes multiples. Deux bandes de fréquences de communication sont utilisées : la bande de fréquences sub-6 GHz et la bande de fréquences mmWave. Un problème d'optimisation linéaire en nombres entiers est formulé afin de maximiser le nombre d'utilisateurs associés tout en assurant une utilisation efficace de ressources. Nous avons étudié la NP-difficulté de ce problème et nous avons proposé deux algorithmes. Le premier est un algorithme glouton centralisé de complexité polynomiale, alors que le deuxième est distribué basé sur la technique d'apprentissage par renforcement QL. L'algorithme d'apprentissage permet à chaque utilisateur d'apprendre la meilleure association automatiquement et indépendamment des autres utilisateurs. Puisque deux différentes bandes de fréquences sont utilisées, l'association des utilisateurs englobe non seulement le choix de la station de base, mais aussi la bande de fréquences de communication (sub-6 GHz ou mmWave). Les résultats de simulation Monte-Carlo ont montré que les résultats des algorithmes proposés sont proches de la solution optimale, tout en réduisant considérablement la complexité de calcul. De plus, ces résultats ont démontré l'importance de l'optimisation des ressources dans les problèmes d'association des utilisateurs aux stations de base.

CHAPITRE 3

TECHNIQUES D'ACCÈS MULTIPLE DANS UN RÉSEAU À BANDES MULTIPLES

Dans ce chapitre, nous étudions le problème d'optimisation conjointe d'allocation des ressources dans un réseau cellulaire utilisant les bandes de fréquences sub-6 GHz et mmWave ainsi les deux techniques d'accès multiple orthogonal et non orthogonal (OMA et NOMA).

Le travail présenté dans ce chapitre a été publié dans l'article de revue *IEEE Transactions on Communications*, 2023 (Chaieb *et al.*, 2023a). Des résultats préliminaires de ce travail ont été publiés dans les conférences *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, London, UK, 2020* (Chaieb *et al.*, 2020a) et *IEEE International Conference on Communications, Seoul, Republic of Korea, 2022* (Chaieb *et al.*, 2022). Cependant, pour des raisons de simplicité, les résultats publiés dans (Chaieb *et al.*, 2020a) sont omis dans ce chapitre.

3.1 Résumé

Ce chapitre étudie le problème d'association des utilisateurs aux stations de base, d'allocation de puissance, d'affectation des sous-canaux et de la technique d'accès multiple. L'objectif est d'optimiser le débit de données tout en respectant les exigences d'un débit minimal par utilisateur et les contraintes de puissance. Nous considérons un réseau cellulaire utilisant à la fois la bande de fréquences sub-6 GHz et la bande de fréquences mmWave. Ainsi, les deux techniques d'accès OMA et NOMA sont utilisées. Le problème d'optimisation est formulé comme un programme non convexe partiellement en nombres entiers. Nous étudions sa complexité et nous proposons des solutions algorithmiques. Tout d'abord, un algorithme glouton ainsi qu'une métaheuristique sont détaillés. Puis, nous proposons deux autres algorithmes, un centralisé et un autre multiagent, qui reposent sur la technique d'apprentissage automatique par renforcement. Ces algorithmes sont ensuite évalués et comparés. Les résultats de simulation montrent les performances de nos solutions et révèlent également les avantages de l'utilisation de deux techniques d'accès dans les réseaux cellulaires à bandes multiples, en termes de débit de données et d'équité entre les utilisateurs.

3.2 Introduction

Avec la croissance explosive du trafic de données mobiles et du nombre d'appareils connectés sans fil, les techniques d'accès OMA et l'utilisation des fréquences de communication sub-6 GHz ne sont plus suffisantes pour répondre aux nouvelles exigences des applications mobiles. En tant que solutions prometteuses, la technique d'accès NOMA et les communications mmWave ont suscité un grand intérêt dans plusieurs travaux de recherche pour les raisons suivantes (Theis *et al.*, 2021; Maraqa *et al.*, 2020). Contrairement aux techniques d'accès OMA, telles que la technique d'accès multiple par répartition en fréquence orthogonale (en anglais : *orthogonal frequency-division multiple access*), la technique d'accès NOMA permet d'associer un grand nombre d'utilisateurs ayant différentes exigences de qualité de service utilisant la même ressource (fréquentielle et/ou temporelle). Ceci permet une utilisation plus efficace de ressources et donc une efficacité spectrale plus élevée. De plus, aucun temps d'attente pour une ressource orthogonale n'est nécessaire pour NOMA (Ding *et al.*, 2017).

La technique d'accès NOMA peut être utilisée dans le domaine de puissance ou dans le domaine de code (en anglais : *power domain or code domain*). Les deux domaines sont distingués par la nature du multiplexage des utilisateurs. Dans le domaine de code, les utilisateurs sont différenciés par différents codes de signature non orthogonaux. Dans le domaine de puissance, les utilisateurs sont différenciés par le niveau de puissance de transmission alloué à chaque utilisateur et qui dépend généralement de la qualité de service demandée, de la bande passante disponible et de la qualité des canaux de transmission. Du côté émetteur, le codage par superposition (en anglais : *superposition coding*) est utilisé pour combiner les signaux des utilisateurs partageant le même canal de transmission. Du côté récepteur, la technique d'annulation successive d'interférences (en anglais : *successive interference cancellation*, SIC) est utilisée pour séparer les différents signaux.

Outre la communication mmWave et la technique d'accès NOMA, la technique d'accès OMA sera probablement incorporée dans les futurs systèmes de communication. En fait, NOMA utilisant plusieurs sous-canaux (en anglais : *multiple sub-channels NOMA*, MSC-NOMA) est une combinaison des techniques d'accès OMA et NOMA. L'idée de MSC-NOMA est de diviser les utilisateurs en groupes. Les utilisateurs du même groupe transmettent sur le même canal avec des niveaux de puissance différents, tandis que les utilisateurs de différents groupes transmettent sur des canaux orthogonaux pour éviter les interférences entre les groupes. Ceci permet d'augmenter considérable-

ment la capacité du réseau, tout en réduisant la complexité du traitement des signaux au niveau du récepteur (Islam *et al.*, 2017; Ding *et al.*, 2017). Ce chapitre considère les deux techniques d'accès OMA et MSC-NOMA dans le domaine de puissance.

3.2.1 Revue de littérature

Les travaux de recherche en rapport avec notre travail sont discutés ci-après. Dans (Lamba *et al.*, 2020), les auteurs ont étudié le problème conjoint de regroupement des utilisateurs, d'allocation des sous-canaux et de puissance dans un réseau cellulaire utilisant seulement la technique d'accès NOMA coopératif. Dans un système NOMA coopératif, les utilisateurs ayant les meilleures conditions de canal de communication sont des relais pour les autres utilisateurs. Afin de maximiser le débit de données, tout d'abord, les auteurs ont proposé un algorithme d'allocation des sous-canaux et de regroupement des utilisateurs en fonction de leurs conditions de canal. Ensuite, pour optimiser la puissance, le problème a été modélisé comme un jeu de Stackelberg. Dans ce jeu, les stations de base jouent le rôle du meneur tandis que les utilisateurs sont les suiveurs. Les auteurs ont démontré que les performances de la technique d'accès NOMA dépassent largement celles de la technique d'accès OMA en termes de débit. De plus, l'optimisation conjointe d'association des utilisateurs et d'allocation de puissance dans un réseau cellulaire utilisant la technique d'accès NOMA a été étudiée dans (Wang *et al.*, 2019). Les auteurs ont proposé deux algorithmes basés sur la théorie des jeux. Dans (Baghani *et al.*, 2019; Zheng *et al.*, 2021), les auteurs ont abordé le problème d'allocation des ressources et de sélection de la technique d'accès (OMA ou NOMA) dans les réseaux cellulaires traditionnels utilisant seulement la bande de fréquences sub-6 GHz. Dans (Chaieb *et al.*, 2020a), nous avons étudié le problème conjoint d'association des utilisateurs aux stations de base, d'affectation des sous-canaux et de choix de la technique d'accès multiple dans un réseau cellulaire à bandes multiples (sub-6 GHz et mmWave) utilisant les deux techniques d'accès OMA et MSC-NOMA. Nous avons commencé par la formulation mathématique du problème comme un programme non convexe en nombres entiers dont l'objectif est de maximiser le nombre d'utilisateurs associés. Puis, nous avons démontré que le problème est NP-difficile et nous avons proposé ensuite un algorithme heuristique de complexité polynomiale. Nous avons démontré que la combinaison des techniques d'accès OMA et MSC-NOMA améliore la capacité du réseau en termes de nombre d'associations. Dans (Liu *et al.*, 2019), les auteurs ont étudié le problème d'association des utilisateurs et d'allocation de puissance dans un réseau hétérogène mmWave composé d'une station de base macro-cellule et des petites

stations de base. Différentes bandes de fréquences mmWave ont été utilisées : 28 GHz, 60 GHz et 73 GHz. Ainsi, deux modes d'accès sont considérés : le mode d'accès à bande unique et le mode d'accès à bandes multiples. Cependant, seulement la technique d'accès OMA a été utilisée. Les auteurs de (Baghani *et al.*, 2019) ont formulé le problème d'allocation de ressources comme un programme linéaire en nombres entiers et ils ont proposé un algorithme itératif utilisant la méthode de Lagrange. Les auteurs ont mentionné que le problème est NP-difficile sans fournir une preuve formelle ainsi ils n'ont pas considéré un réseau à bandes multiples et la mobilité des utilisateurs. Les résultats de simulation ont démontré les avantages de l'utilisation de deux techniques d'accès OMA et NOMA en termes de probabilité d'interruption et de débit de données. Différentes solutions de regroupement des utilisateurs dans les réseaux cellulaires utilisant la technique d'accès NOMA ont été proposées dans (You *et al.*, 2020; Ali *et al.*, 2016).

L'apprentissage par renforcement profond (en anglais : *deep reinforcement learning*) est une technique qui combine l'apprentissage par renforcement et l'apprentissage en profondeur. Cette technique s'appuie sur l'utilisation de réseaux de neurones artificiels à plusieurs couches, permettant aux agents de choisir des actions afin de maximiser leurs récompenses. Dans (Li *et al.*, 2020), le problème d'association des utilisateurs aux stations de base et d'allocation de la bande passante dans un réseau hétérogène a été résolu en utilisant la technique d'apprentissage par renforcement profond. Les résultats de simulation ont montré les performances de la technique d'apprentissage par renforcement profond en comparaison avec des algorithmes de référence. Cependant, les auteurs n'ont pas étudié la complexité computationnelle, n'ont pas considéré la communication sur plusieurs bandes de fréquences et n'ont pas proposé une solution multiagent. Dans (Tang *et al.*, 2020), les auteurs ont étudié les techniques d'allocation des ressources dans un réseau hétérogène utilisant seulement la bande de fréquences mmWave. Le duplexage par séparation temporelle (en anglais : *time-division duplex*) a été considéré, où les transmissions sur les deux liens montant et descendant sont effectuées sur la même fréquence. Les auteurs ont proposé un algorithme d'apprentissage QL profond (en anglais : *deep Q-learning*, DQL) qui tient compte de la forte mobilité des utilisateurs pour ajuster dynamiquement la configuration de la méthode de duplexage en vue d'optimiser le débit de données et le taux de perte de paquets. De plus, les auteurs de (Zhang *et al.*, 2020) ont étudié le problème d'optimisation combinatoire d'association des utilisateurs aux stations de base dans des réseaux ultra-denses utilisant seulement la technique d'accès OMA et la bande de fréquences sub-6 GHz. Les auteurs ont proposé une solution d'apprentissage par renforcement profond pour associer

intelligemment les utilisateurs à la station de base macro-cellule ou aux petites stations de base. Les auteurs de (Zhao *et al.*, 2019) ont proposé un algorithme distribué multiagent basé sur l’approche d’apprentissage par renforcement profond pour obtenir des solutions quasi-optimales d’allocation de ressources. Ils ont démontré l’efficacité de la solution proposée pour un grand espace d’actions. Il est à noter que seulement la technique d’accès OMA a été considérée dans ces travaux (Tang *et al.*, 2020; Li *et al.*, 2020; Zhang *et al.*, 2020; Liao *et al.*, 2019; Zhao *et al.*, 2019). Les principaux travaux de recherche sont résumés et comparés à notre travail dans le tableau 3.1.

TABLE 3.1 – Principaux travaux de recherche.

Référence	Fonction objectif				Technique		Bande de fréquences		Mobilité
	Nombre d’associations	Débit	Énergie	Délais	OMA	NOMA	mmWave	Sub-6 GHz	
(Lamba <i>et al.</i> , 2020) (Liu <i>et al.</i> , 2020b) (Tang <i>et al.</i> , 2018) (You <i>et al.</i> , 2020)		✓				✓		✓	
(Liu <i>et al.</i> , 2019)					✓		✓		
(Chaieb <i>et al.</i> , 2018)	✓				✓		✓	✓	✓
(Zhu <i>et al.</i> , 2020)			✓	✓		✓		✓	
(Ding <i>et al.</i> , 2022)			✓	✓	✓	✓		✓	
(Liu <i>et al.</i> , 2020a)			✓		✓			✓	
(Chaieb <i>et al.</i> , 2020a)	✓				✓	✓	✓	✓	
(Chaieb <i>et al.</i> , 2020b)	✓				✓		✓	✓	
(Muhammed <i>et al.</i> , 2020)			✓			✓		✓	
(Baghani <i>et al.</i> , 2019) (Zheng <i>et al.</i> , 2021) (Wu <i>et al.</i> , 2022)		✓			✓	✓		✓	
(Cui <i>et al.</i> , 2018) (Liu <i>et al.</i> , 2020b)		✓				✓	✓		
(Li <i>et al.</i> , 2020) (Zhang <i>et al.</i> , 2020)		✓			✓			✓	
(Tang <i>et al.</i> , 2020) (Zhao <i>et al.</i> , 2019) (Sritharan <i>et al.</i> , 2020)		✓			✓			✓	✓
(Liao <i>et al.</i> , 2019)			✓		✓			✓	✓
Solutions proposées		✓			✓	✓	✓	✓	✓

3.2.2 Contributions

Pour tirer pleinement parti des avantages de l’utilisation conjointe de différentes techniques d’accès multiple et de deux bandes de fréquences, une nouvelle approche d’allocation des ressources dans les réseaux cellulaires doit être étudiée. Cette approche implique l’association des utilisateurs aux stations de base, l’allocation de puissance, l’attribution des sous-canaux et la sélection de la technique d’accès multiple. À notre connaissance, notre travail est le premier à étudier un modèle du système complexe à bandes multiples utilisant les deux techniques d’accès OMA et MSC-NOMA.

Les contributions de ce chapitre sont résumées comme suit :

1. Nous formulons mathématiquement le problème conjoint d'allocation de ressources comme un programme mathématique non linéaire et non convexe partiellement en nombres entiers (en anglais : *mixed-integer programming problem*). Deux techniques d'accès multiple sont utilisées OMA et MSC-NOMA dans le domaine de puissance. L'objectif est de maximiser le débit de données tout en respectant les exigences de qualité de service des utilisateurs et les contraintes de puissance. Ce problème est appelé problème d'allocation de ressources RAP.
2. Nous étudions la NP-difficulté de RAP.
3. Nous proposons un algorithme glouton et une métaheuristique pour des raisons de comparaison.
4. Nous proposons un algorithme centralisé et un autre distribué basés sur l'approche d'apprentissage par renforcement profond.
5. Enfin, nous démontrons l'efficacité des solutions proposées en tenant compte de la mobilité des utilisateurs, ainsi les avantages de l'utilisation de deux bandes de fréquences et de deux techniques d'accès multiple OMA et MSC-NOMA en termes d'équité entre les utilisateurs et de débit du système.

3.2.3 Organisation

Dans la suite de ce chapitre, la section 3.3 introduit le modèle du système. La section 3.4 présente la formulation mathématique du problème et la section 3.5 étudie sa complexité. Les solutions algorithmiques sont proposées dans la section 3.6. Les résultats de simulation sont fournis dans la section 3.7 et les conclusions sont présentées dans la dernière section 3.8.

3.3 Modèle du système

Nous considérons un réseau à bandes multiples composé de U utilisateurs et de B stations de base. Nous notons par $\mathcal{U} = \{1, \dots, U\}$ l'ensemble d'utilisateurs et par $\mathcal{B} = \mathcal{B}_s \cup \mathcal{B}_m$ l'ensemble de stations de base, où $\mathcal{B}_s = \{1, \dots, B_s\}$ représente l'ensemble de stations de base utilisant la bande de fréquences sub-6 GHz (sBSs) et $\mathcal{B}_m = \{1, \dots, B_m\}$ représente l'ensemble de stations de base utilisant la bande de fréquences mmWave (mBSs). Nous notons par $\mathcal{C} = \mathcal{C}_s \cup \mathcal{C}_m$ l'ensemble de sous-canaux, où

$\mathcal{C}_s = \{1, \dots, C_s\}$ et $\mathcal{C}_m = \{1, \dots, C_m\}$ sont respectivement les ensembles de sous-canaux disponibles pour les sBSs et les mBSs. La transmission sur le lien descendant, où les stations de base envoient des messages aux utilisateurs associés, est considérée. Un exemple de modèle du système est illustré dans la figure 3.1, où $B_s = B_m = 2$, $U = 5$, et $C_s = C_m = 1$. Dans ce modèle, nous supposons que les stations de base peuvent basculer, pour chaque sous-canal alloué, entre les deux techniques d'accès OMA et MSC-NOMA. De plus, chaque utilisateur ne peut être associé qu'à une seule station de base et chaque station de base peut associer plusieurs utilisateurs. Le problème d'allocation de ressources étudié (RAP) répond à la question suivante : Comment associer les utilisateurs, choisir la technique d'accès et allouer les sous-canaux et la puissance de façon à satisfaire les exigences de qualité de service des utilisateurs et à optimiser le débit descendant du système ?

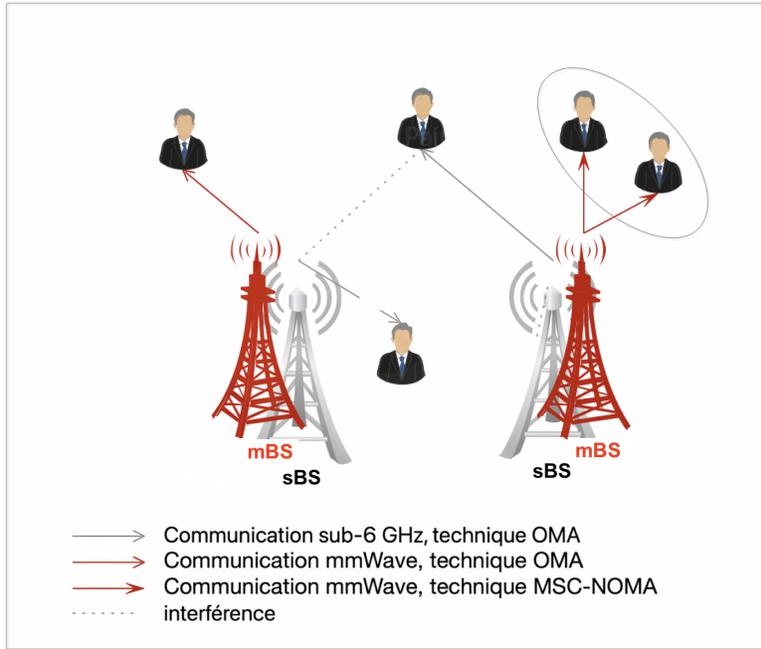


FIGURE 3.1 – Une illustration d'un réseau à bandes multiples utilisant les deux techniques OMA et MSC-NOMA.

Pour formuler mathématiquement RAP, nous utilisons une variable d'association binaire x_{ubc} qui prend les valeurs suivantes :

$$x_{ubc} = \begin{cases} 1, & \text{si l'utilisateur } u \text{ est associé à la station } b \text{ utilisant le sous-canal } c \\ 0, & \text{sinon} \end{cases} \quad (3.1)$$

Nous utilisons p_{ubc} pour représenter le niveau de puissance de transmission de la station de base b à

l'utilisateur u sur le sous-canal c . En notation matricielle, les matrices à trois dimensions associées à x_{ubc} et p_{ubc} sont notées par $\mathbf{X} = [x_{ubc}]$ et $\mathbf{P} = [p_{ubc}]$.

Comme plusieurs utilisateurs peuvent partager la même ressource utilisant la technique d'accès MSC-NOMA, le bruit NOMA résultant est calculé de la manière suivante, pour tout $u \in \mathcal{U}, b \in \mathcal{B}$ et $c \in \mathcal{C}$:

$$n_{ubc} = g_{ubc} \sum_{u' \in \mathcal{U}'} p_{u'bc}, x_{u'bc}, \quad (3.2)$$

où $\mathcal{U}' = \{u, u' \in \mathcal{U} \mid g_{u'bc} > g_{ubc}\}$ est l'ensemble d'utilisateurs dont le gain du canal est supérieur à celui de l'utilisateur u , g_{ubc} est le gain du canal entre l'utilisateur u et la station de base b et σ_b^2 est la puissance du bruit blanc additif Gaussien (en anglais : *additive white gaussian noise*, AGWN). Nous supposons que chaque mBS b dispose de n_b antennes capables d'envoyer des faisceaux directionnels. Le lobe principal est formé en direction de l'utilisateur associé. Par contre, l'utilisateur peut être aléatoirement placé dans le lobe principal ou secondaire des autres mBSs. Nous modélisons le gain d'antennes en utilisant le modèle sectoriel suivant (Xu *et al.*, 2016; Venugopal *et al.*, 2015) :

$$a_b = \begin{cases} n_b, & \text{si } x_{ubc} = 1, \\ n_b, & \text{si } x_{ubc} = 0 \text{ et } |\zeta| \leq \zeta^m / 2\pi, \\ 1 / \sin^2 \left(\frac{3\pi}{2\sqrt{n_b}} \right), & \text{sinon} \end{cases}, \quad (3.3)$$

où ζ désigne l'angle de la direction de visée qui est uniformément distribué entre $-\pi$ et π et $\zeta^m = 2\pi/\sqrt{n_b}$ est la largeur du faisceau d'antenne. Le gain du canal entre l'utilisateur u et la mBS b est donné par, pour tout $u \in \mathcal{U}, b \in \mathcal{B}_m$ et $c \in \mathcal{C}_m$:

$$g_{ubc} = a_b f_c(d_{ub})^{-1}, \quad (3.4)$$

où $f_c(d_{ub})$ est la fonction d'affaiblissement de propagation et d_{ub} est la distance entre l'utilisateur u et la mBS b . De plus, le gain du canal entre l'utilisateur u et la SBS b est donné par, pour tout $u \in \mathcal{U}, b \in \mathcal{B}_s$ et $c \in \mathcal{C}_s$:

$$g_{ubc} = |g'_{ubc}|^2 \beta (d_{ub}/d_0)^{-\delta}, \quad (3.5)$$

où $|g'_{ubc}|$ est le module d'une variable aléatoire gaussienne avec moyenne nulle et variance unitaire, δ est l'exposant d'affaiblissement, d_0 est une distance de référence et β est une constante qui capture

les effets du système et de transmission.

Dans ce chapitre, nous supposons une connaissance parfaite et instantanée de l'état du canal. Lorsque l'utilisateur u est associé à la station de base b , le SINR est donné par, pour tout $u \in \mathcal{U}$, $b \in \mathcal{B}$ et $c \in \mathcal{C}$:

$$\text{SINR}_{ubc} = \frac{p_{ubc}g_{ubc}}{\sigma_b^2 + n_{ubc} + i_{ubc}}, \quad (3.6)$$

où i_{ubc} est l'interférence sur le lien descendant et qui est égale à la somme du produit de la puissance allouée par d'autres stations de base à d'autres utilisateurs utilisant le même sous-canal c et du gain du canal entre l'utilisateur u et d'autres stations de base. Par conséquent, i_{ubc} s'exprime par, pour tout $u \in \mathcal{U}$, $b \in \mathcal{B}$ et $c \in \mathcal{C}$:

$$i_{ubc} = \sum_{u' \in \mathcal{U} \setminus \{u\}} \sum_{b' \in \mathcal{B} \setminus \{b\}} p_{u'b'c} g_{ub'c} x_{u'b'c}. \quad (3.7)$$

En se basant sur le théorème du Shannon, le débit binaire possible r_{ubc} reçu par l'utilisateur $u \in \mathcal{U}$ est donné par, pour tout $b \in \mathcal{B}$ et $c \in \mathcal{C}$:

$$r_{ubc} = w'_b \log_2(1 + \text{SINR}_{ubc}), \quad (3.8)$$

où w'_b est la bande passante du sous-canal allouée par la station de base b . Nous notons par q_{bc} le nombre d'utilisateurs associés à la station de base b partageant le même sous-canal c . Nous introduisons le paramètre q_{bc} tel que lorsque la technique d'accès OMA est sélectionnée, q_{bc} est égale à 1, tandis que lorsque la technique d'accès MSC-NOMA est choisie, q_{bc} prend une valeur supérieure à 1. La bande passante du canal est égale à la bande passante disponible pour chaque station de base b divisée par le nombre maximal d'associations utilisant des sous-canaux distincts. Par conséquent, w'_b est donnée par :

$$w'_b = \begin{cases} \frac{w_b}{\min(U, C_m)}, & \text{si } b \in \mathcal{B}_m \\ \frac{w_b}{\min(U, C_s)}, & \text{si } b \in \mathcal{B}_s \end{cases}. \quad (3.9)$$

Pour que le débit minimal D_u^{\min} requis par l'utilisateur u soit respecté, le niveau de puissance alloué

devrait satisfaire la contrainte suivante, pour tout $u \in \mathcal{U}$, $b \in \mathcal{B}$ et $c \in \mathcal{C}$:

$$r_{abc} \geq D_u^{min}. \quad (3.10)$$

Par conséquent, le niveau de puissance minimal alloué à l'utilisateur u est calculé comme suit, pour tout $u \in \mathcal{U}$, $b \in \mathcal{B}$ et $c \in \mathcal{C}$:

$$p_{abc} = \left(2^{\frac{D_u^{min}}{w_b}} - 1\right) \left(\frac{\sigma_b^2 + n_{abc} + i_{abc}}{g_{abc}}\right). \quad (3.11)$$

Enfin, le débit total descendant est la somme des débits reçus par tous les utilisateurs associés et il est calculé comme suit :

$$R = \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} r_{abc} x_{abc}. \quad (3.12)$$

3.4 Formulation du problème

La fonction objectif du problème d'optimisation est de maximiser le débit descendant du système tout en respectant les exigences de qualité de service des utilisateurs et les contraintes de puissance.

Par conséquent, RAP est formulé comme suit :

$$\underset{\mathbf{X}, \mathbf{P}}{\text{maximiser}} \quad R \quad (\text{P2a})$$

$$\text{sous les contraintes: } x_{abc} \in \{0, 1\}, \forall u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}, \quad (\text{P2b})$$

$$D_u^{min} x_{abc} \leq r_{abc} \leq D_u^{max} x_{abc}, \forall u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}, \quad (\text{P2c})$$

$$p_{abc} \geq 0, \forall u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}, \quad (\text{P2d})$$

$$\sum_{u \in \mathcal{U}} \sum_{c \in \mathcal{C}} p_{abc} x_{abc} \leq p_b, \forall b \in \mathcal{B}, \quad (\text{P2e})$$

$$\sum_{u \in \mathcal{U}} x_{abc} \leq q_{bc}, \forall b \in \mathcal{B}, c \in \mathcal{C}, \quad (\text{P2f})$$

$$\sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} x_{abc} \leq 1, \forall u \in \mathcal{U}, \quad (\text{P2g})$$

$$x_{abc} = 0, \forall u \in \mathcal{U}, b \in \mathcal{B}_m, c \notin \mathcal{C}_m, \quad (\text{P2h})$$

$$x_{abc} = 0, \forall u \in \mathcal{U}, b \notin \mathcal{B}_m, c \in \mathcal{C}_m. \quad (\text{P2i})$$

Les contraintes (P2b) assurent que la variable d'association x_{abc} soit binaire pour tout u, b , et c . Les contraintes (P2c) garantissent les valeurs seuils de débit D_u^{min} et D_u^{max} de chaque utilisateur associé. Il est important de noter que le débit maximal par utilisateur limite le niveau de puissance alloué à l'utilisateur ayant le meilleur canal, ce qui permet une répartition équitable de la puissance disponible entre les utilisateurs. Les contraintes (P2d) et (P2e) assurent que les niveaux de puissance alloués soient positifs et que la somme des niveaux de puissance alloués par la station de base b ne dépasse pas sa puissance totale disponible notée par p_b . Les contraintes (P2f) limitent le nombre d'associations par sous-canal à q_{bc} et les contraintes (P2g) garantissent que chaque utilisateur soit associé à une seule station de base en utilisant un seul sous-canal. Enfin, les contraintes (P2h) (resp. les contraintes (P2i)) garantissent que si un utilisateur est associé à une mBS (resp. à une sBS), le sous-canal alloué sera dans l'ensemble \mathcal{C}_m (resp. dans l'ensemble \mathcal{C}_s). Le problème (P2) est un programme mathématique non linéaire et non convexe partiellement en nombres entiers en raison de la fonction logarithmique et du terme d'interférence dans les contraintes (P2c).

3.5 NP-difficulté

Dans cette section, nous montrons que le problème d'optimisation (P2) est NP-difficile en utilisant la technique par réduction. Par conséquent, nous montrons que le problème d'appariement à 3 dimensions (*en anglais : 3-dimensional matching problem, 3-DMP*), qui est un problème connu NP-difficile, peut se réduire en un nombre polynomial d'opérations au problème (P2) (Karp, 2010).

Théorème 3.5.1. *RAP est NP-difficile.*

Démonstration : Soit \mathcal{Z}, \mathcal{Y} et \mathcal{V} trois ensembles finis et disjoints et \mathcal{T} un sous-ensemble de $\mathcal{Z} \times \mathcal{Y} \times \mathcal{V}$. L'objectif de 3-DMP est de trouver un appariement tridimensionnel $\mathcal{M} \subseteq \mathcal{T}$ qui maximise le nombre d'appariements, où pour tout les deux distincts triplets $(z_1, y_1, v_1) \in \mathcal{M}$ et $(z_2, y_2, v_2) \in \mathcal{M}$, nous avons $z_1 \neq z_2, y_1 \neq y_2$ et $v_1 \neq v_2$. À partir d'une instance de 3-DMP, une instance de (P2) peut être construite comme suit : soit $\mathcal{B}_s = \mathcal{C}_s = \emptyset, B_m = 1, \mathcal{U}_1 = \left\{1, \dots, \frac{U}{2}\right\}, q_{bc} = 2$ et $\mathcal{U}_2 = \mathcal{U} \setminus \mathcal{U}_1$. Nous considérons seulement la technique d'accès MSC-NOMA, c'est-à-dire que chaque sous-canal millimétrique est partagé par deux utilisateurs $u_1 \in \mathcal{U}_1$ et $u_2 \in \mathcal{U}_2$. De plus, nous supposons que pour tout $u \in \mathcal{U}, b \in \mathcal{B}_m$ et $c \in \mathcal{C}_m$, le niveau de puissance alloué est $p_{abc} = \frac{p_b}{C_m} \geq 0$. Enfin, (P2)

peut être reformulé comme suit :

$$\underset{\mathbf{x}}{\text{maximiser}} \quad \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} x_{ubc} \quad (\text{P3a})$$

$$\text{sous les contraintes: } x_{ubc} \in \{0, 1\}, \forall u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}, \quad (\text{P3b})$$

$$D_u^{\min} x_{ubc} \leq r_{ubc} \leq D_u^{\max} x_{ubc}, \forall u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}, \quad (\text{P3c})$$

$$\sum_{u_1 \in \mathcal{U}_1} \sum_{u_2 \in \mathcal{U}_2} x_{u_1 bc} + x_{u_2 bc} = 2, \forall b \in \mathcal{B}, c \in \mathcal{C}, \quad (\text{P3d})$$

$$\sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} x_{ubc} \leq 1, \forall u \in \mathcal{U}. \quad (\text{P3e})$$

Il est à noter que (P3) est équivalent à maximiser le nombre d'utilisateurs associés aux mBSs sous les contraintes de débit de données. Par analogie, une instance du problème (P2) peut être créée en un temps polynomial comme suit : $\mathcal{U}_1 = \mathcal{Z}$, $\mathcal{U}_2 = \mathcal{Y}$ et $\mathcal{C}_m = \mathcal{V}$. Étant donné que 3-DMP est NP-difficile, (P2) est également NP-difficile. ■

Résoudre le problème (P2) tout en utilisant la recherche exhaustive est coûteux en termes de calculs, car toutes les combinaisons possibles d'association des utilisateurs, d'allocation de puissance, d'affectation des sous-canaux et de la technique d'accès multiple doivent être évaluées. Par conséquent, des solutions algorithmiques efficaces sont proposées.

3.6 Solutions algorithmiques

Dans cette section, nous détaillons les algorithmes proposés, ensuite nous évaluons leurs complexités asymptotiques.

3.6.1 Un algorithme glouton (GRA)

L'idée principale de l'algorithme glouton proposé (appelé GRA) est de résoudre le problème (P2) en deux étapes. Dans la première étape, GRA associe les utilisateurs aux stations de base, attribue les sous-canaux et choisit la technique d'accès (OMA ou MSC-NOMA) en allouant aux utilisateurs des niveaux de puissance de transmission qui garantissent les exigences minimales en matière de débit de données (c'est-à-dire D_u^{\min}). Dans la deuxième étape, le calcul des niveaux de puissance est basé sur les exigences maximales (c'est-à-dire D_u^{\max}) afin de maximiser le débit total. Le pseudo-code de

l'algorithme GRA est donné dans l'algorithme 4 et il est détaillé comme suit.

GRA commence par initialiser la matrice d'association \mathbf{X} à une matrice nulle, le compteur u représentant l'indice de l'utilisateur à 1 et le débit total R à zéro. Ensuite, GRA itère les utilisateurs. Pour chaque sous-canal disponible c dans la station de base b , GRA trie les utilisateurs par ordre décroissant selon leurs gains du canal, regroupe l'utilisateur u à l'utilisateur $\left(\left\lceil \frac{U}{q_{bc}} \right\rceil + u\right)$ comme dans (Ali *et al.*, 2016). Dans la ligne 10, les niveaux de puissance de transmission p_{ubc}^{oma} , p_{ubc}^{noma} et $p_{u'bc}^{noma}$ qui garantissent leurs exigences minimales de débit de données sont calculés pour les deux techniques d'accès OMA et MSC-NOMA. Il est à noter que q_{bc} est égale à 2 pour tout $b \in \mathcal{B}$ et $c \in \mathcal{C}$ pour garantir les performances de la technique SIC. Ensuite, GRA sélectionne l'indice de la station de base et le sous-canal offrant le niveau minimal de puissance de transmission et met à jour la matrice d'association \mathbf{X} . Dans le pire des cas, la complexité asymptotique de GRA est $\mathcal{O}(U^3 B_s^3 C_s)$ pour $C_s = C_m$, $B_s = B_m$ et $U \gg B_s$. L'algorithme *water filling*, appelé WFA, calcule la puissance allouée en se basant sur le débit maximal de chaque utilisateur. Le pseudo-code de l'algorithme WFA est détaillé dans la sous-section 3.6.3.

Algorithme 4 Pseudo-Code de l'algorithme GRA

- 1: $\mathbf{X} \leftarrow \mathbf{0}$, $R \leftarrow 0$, $u \leftarrow 1$, $\mathbf{P} \leftarrow \mathbf{0}$, $q_{bc} \leftarrow 2$ pour tout $b \in \mathcal{B}$ et $c \in \mathcal{C}$.
 - 2: **Tant que** $u \leq U$ **et** somme(\mathbf{X}) $< q_{bc}(B_s C_s + B_m C_m)$ **faire**
 - 3: $\mathcal{P}^* \leftarrow \emptyset$.
 - 4: **Pour** b de 1 à B **faire**
 - 5: **Pour** c de 1 à C_s ou C_m **faire**
 - 6: Trier les utilisateurs non associés dans un ordre décroissant en fonction de leurs gains du canal.
 - 7: Regrouper l'utilisateur u avec l'utilisateur $u' = \left(\left\lceil \frac{U}{q_{bc}} \right\rceil + u\right)$.
 - 8: Calculer les niveaux de puissance requis p_{ubc}^{oma} , p_{ubc}^{noma} et $p_{u'bc}^{noma}$ en utilisant l'équation (3.11).
 - 9: **Si** les contraintes (P5c), (P5e) et (P5f) sont respectées **alors**
 - 10: Ajouter p_{ubc}^{oma} , p_{ubc}^{noma} et/ou $p_{u'bc}^{noma}$ dans le vecteur \mathcal{P}^* .
 - 11: **Fin Si**
 - 12: **Fin Pour**
 - 13: **Fin Pour**
 - 14: $b^* \leftarrow \arg \min_b(\mathcal{P}^*)$; $c^* \leftarrow \arg \min_c(\mathcal{P}^*)$.
 - 15: Mettre à jour \mathbf{X} et \mathbf{P} ; $u \leftarrow u + 1$.
 - 16: **Fin Tant que**
 - 17: **Retourner** \mathbf{X}
-

3.6.2 Une métaheuristique (GA)

L'approche génétique est une approche métaheuristique inspirée du processus de l'évolution naturelle des espèces (Simon, 2013). Elle reprend plusieurs opérations de la nature telles que la sélection, le croisement et la mutation pour générer de nouvelles solutions en explorant l'espace de recherche. Ces opérateurs permettent de converger vers une solution optimale ou quasi-optimale en un nombre limité d'itérations. Pour adapter l'approche génétique à notre problème, nous utilisons la méthode d'encodage binaire. Chaque solution d'association \mathbf{X} , appelée individu, est une matrice binaire à trois dimensions de taille $U \cdot B \cdot C$. La fonction d'évaluation adoptée pour évaluer la pertinence (en anglais : *fitness*) de chaque individu est donnée par l'équation (3.12). Les différentes opérations de l'approche génétique sont détaillées ci-dessous :

- **Sélection** : Nous choisissons aléatoirement deux individus parmi la population, nommés parents, en utilisant la sélection élitiste. La sélection élitiste est adoptée pour éviter la destruction des individus ayant de bonnes caractéristiques génétiques garantissant ainsi la présence des individus les plus performants de chaque génération.
- **Croisement** : Nous adoptons une stratégie de croisement par appariement partiel (en anglais : *partial matching crossover strategy*). En d'autres termes, nous divisons la matrice de chaque parent sélectionné en vecteurs binaires et nous croisons ensuite deux vecteurs pour obtenir de nouveaux individus. Enfin, nous corrigeons les nouveaux individus en remplaçant les bits 0 par 1 et vice-versa de sorte que toutes les contraintes du problème (P2) soient respectées. Ces individus sont considérés pour une exécution ultérieure de l'algorithme.
- **Mutation** : Pour maintenir la diversité dans la population, nous choisissons un bit aléatoirement des nouveaux individus et nous changeons sa valeur. Enfin, nous vérifions si les nouveaux individus respectent toutes les contraintes du problème (P2).

Le pseudo-code de l'algorithme génétique, nommé GA, est donné dans l'algorithme 5 et il est résumé comme suit. GA commence par générer l'ensemble d'individus, appelé population initiale, en utilisant notre algorithme heuristique proposé dans (Chaieb *et al.*, 2020a). Ce dernier maximise le nombre d'associations et attribue aux utilisateurs associés des niveaux de puissance de transmission qui garantissent leurs exigences en matière de qualité de service. GA itère ensuite jusqu'à ce que le nombre maximal d'itérations N_c soit atteint. À chaque itération, l'algorithme évalue chaque individu, sélectionne aléatoirement deux individus parents et les croise pour générer deux nouveaux individus. Ensuite, GA supprime les individus répétés jusqu'à ce qu'il n'y ait plus de répétition.

Afin de maintenir une complexité de calcul raisonnable, nous adoptons un nombre maximal d'itérations et une taille maximale de population. Dans le pire des cas, la complexité asymptotique de l'algorithme GA est $\mathcal{O}(N_c U^3 B_s^3 C_s)$ pour $C_s = C_m$, $B_s = B_m$ et $U \gg B_s$. Enfin, l'algorithme WFA est utilisé pour calculer efficacement les niveaux de puissance alloués aux utilisateurs associés.

Algorithme 5 Pseudo-code de l'algorithme GA

- 1: Générer la population initiale. Initialiser N_c .
 - 2: $i \leftarrow 1$.
 - 3: **Tant que** $i \leq N_c$ **faire**
 - 4: Évaluer chaque individu dans la population.
 - 5: Effectuer l'opération de sélection.
 - 6: Effectuer l'opération de croisement et l'opération de mutation.
 - 7: Vérifier les contraintes du problème (P2).
 - 8: $i \leftarrow i + 1$.
 - 9: **Fin Tant que**
 - 10: **Retourner** Le meilleur individu.
-

3.6.3 Un algorithme d'allocation de puissance (WFA)

L'approche itérative *water filling* est utilisée par les algorithmes proposés GRA et GA pour calculer la puissance allouée d'une manière à ce que les contraintes de puissance soient satisfaites avec égalité. Soit $\underline{p} = [p_b] \geq \underline{0}$ le vecteur de puissance disponible dans les stations de base, le problème RAP peut être écrit sous la forme suivante :

$$\underset{\mathbf{P}}{\text{maximiser}} \quad R \quad (\text{P4a})$$

$$\text{sous les contraintes: } \sum_{u \in \mathcal{U}} \sum_{c \in \mathcal{C}} p_{ubc} x_{ubc} = p_b, \forall b \in \mathcal{B}, \quad (\text{P4b})$$

$$p_{ubc} \geq 0, \forall u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}. \quad (\text{P4c})$$

La fonction objectif du problème (P4) est de maximiser le débit total du réseau. Le débit réalisable pour l'utilisateur u est une fonction logarithmique qui prend en compte le bruit NOMA et le bruit intercellulaire. La matrice hessienne de la fonction de débit, qui est la matrice des dérivées partielles secondes de R , est calculée et prouvée dans le lemme 3.6.1 comme étant définie négativement.

Lemme 3.6.1. *Le problème (P4) est convexe.*

Démonstration : Nous démontrons que la matrice hessienne de r_{ubc} est définie négative (Ali *et al.*, 2018). Nous considérons les deux cas : la communication utilisant la technique d'accès OMA et la communication utilisant la technique d'accès MSC-NOMA, où deux utilisateurs peuvent partager la même ressource.

Cas 1 : Lorsque la technique d'accès OMA est sélectionnée, le débit de données sur le sous-canal c est représenté par R_{bc} et il est donné par, pour tout $b \in \mathcal{B}$ et $c \in \mathcal{C}$:

$$R_{bc} = \sum_{u \in \mathcal{U}} r_{ubc} x_{ubc} = w'_b \log_2 \left(1 + \frac{p_{ubc} g_{ubc}}{\sigma_b^2 + i_{ubc}} \right). \quad (3.13)$$

La dérivée seconde de r_{ubc} , notée par $-\phi_1$, est négative et égale à :

$$\phi_1 = \frac{w'_b}{\ln 2} \left(\frac{g_{ubc}}{\tau_{ubc} + p_{ubc} g_{ubc}} \right)^2, \quad (3.14)$$

où $\tau_{ubc} = \sigma_b^2 + i_{ubc}$. Comme la fonction objectif du problème (P4) est de maximiser une somme de fonctions concaves, le problème (P4) est donc convexe (Cohen et Megiddo, 1993).

Cas 2 : Considérons le cas d'une communication MSC-NOMA à 2 utilisateurs u et u' (c'est-à-dire $x_{ubc} = 1$ et $x_{u'bc} = 1$). Le débit de données sur le sous-canal c est donné par, pour tout $b \in \mathcal{B}$ et $c \in \mathcal{C}$:

$$R_{bc} = \sum_{u \in \mathcal{U}} r_{ubc} x_{ubc} \quad (3.15)$$

$$= w'_b \log_2 \left(1 + \frac{p_{ubc} g_{ubc}}{\sigma_b^2 + n_{ubc} + i_{ubc}} \right) + w'_b \log_2 \left(1 + \frac{p_{u'bc} g_{u'bc}}{\sigma_b^2 + i_{u'bc}} \right). \quad (3.16)$$

La matrice hessienne de R_{bc} , notée par H , peut être écrite comme suit :

$$H = \begin{bmatrix} \frac{\partial^2 R_{bc}}{\partial^2 p_{ubc}} & \frac{\partial^2 R_{bc}}{\partial p_{ubc} \partial p_{u'bc}} \\ \frac{\partial^2 R_{bc}}{\partial p_{u'bc} \partial p_{ubc}} & \frac{\partial^2 R_{bc}}{\partial^2 p_{u'bc}} \end{bmatrix} = \begin{bmatrix} -\phi_2 & -\phi_2 \\ -\phi_2 & -\phi_2 + \phi'_2 - \phi_3 \end{bmatrix}, \quad (3.17)$$

où ϕ_2 , ϕ'_2 et ϕ_3 sont donnés par :

$$\phi_2 = \frac{w'_b}{\ln 2} \left(\frac{g_{ubc}}{\tau_{ubc} + g_{ubc}(p_{ubc} + p_{u'bc})} \right)^2, \quad (3.18)$$

$$\phi'_2 = \frac{w'_b}{\ln 2} \left(\frac{g_{ubc}}{\tau_{ubc} + p_{u'bc} g_{ubc}} \right)^2, \quad (3.19)$$

$$\phi_3 = \frac{w'_b}{\ln 2} \left(\frac{g_{u'bc}}{\tau_{u'bc} + p_{u'bc}g_{u'bc}} \right)^2. \quad (3.20)$$

Ci-dessous, nous calculons les mineurs de la matrice H qui correspondent aux déterminants de ses sous-matrices carrées. Le mineur du premier ordre de H est négatif et il est égal à $-\phi_2$. Le mineur d'ordre 2 de H peut être écrit comme suit :

$$\phi_2(\phi_3 - \phi'_2) = \frac{w'_b}{\ln 2} \phi_2 \cdot \left(\frac{(2p_{u'bc}g_{ubc}g_{u'bc} + g_{ubc}\tau_{u'bc} + g_{u'bc}\tau_{ubc})(g_{u'bc}\tau_{ubc} - g_{ubc}\tau_{u'bc})}{(\tau_{u'bc} + p_{u'bc}g_{u'bc})^2(\tau_{ubc} + p_{u'bc}g_{ubc})^2} \right). \quad (3.21)$$

Nous supposons que $\tau_{u'bc} \approx \tau_{ubc}$, où $\sigma_b^2 = \sigma'_b{}^2$, le mineur d'ordre 2 de H est positif puisque $g_{u'bc} > g_{ubc}$ et $\phi_2 > 0$. Les mineurs successifs de la matrice H alternent en signe, ce qui implique que la fonction R_{bc} est strictement concave. De plus, les contraintes de puissance de (P4) sont convexes. Par conséquent, le problème (P4) est convexe. Ceci prouve ce lemme. \blacksquare

Dans ce qui suit, nous utilisons la méthode de Lagrange pour résoudre (P4). Cette méthode consiste à transformer le problème d'optimisation en un problème sans contraintes en introduisant des multiplicateurs de Lagrange pour chaque contrainte. La fonction de Lagrange est donnée par :

$$\mathcal{L}(\mathbf{P}, \underline{\mu}, \underline{\iota}) = R + \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} \left(\iota_{ubc} p_{ubc} - \mu_b (p_{ubc} x_{ubc} - p_b) \right), \quad (3.22)$$

où $\underline{\mu} = [\mu_b]$ et $\underline{\iota} = [\iota_{ubc}]$ sont les multiplicateurs de Lagrange des contraintes de puissance. Les conditions de Karush-Kuhn-Tucker (KKT) assurant les solutions optimales de (P4) sont les suivantes :

$$\frac{\partial \mathcal{L}(\mathbf{P}, \underline{\mu}, \underline{\iota})}{\partial p_{ubc}} = 0, \quad (3.23)$$

où $\underline{\mu} \geq \underline{0}$ and $\underline{\iota} \geq \mathbf{0}$. La dérivée de l'équation (3.22) est calculée comme suit, pour tout $u \in \mathcal{U}$, $b \in \mathcal{B}$ et $c \in \mathcal{C}$:

$$\frac{\partial \mathcal{L}(\mathbf{P}, \underline{\mu}, \underline{\iota})}{\partial p_{ubc}} = \frac{\partial r_{ubc}}{\partial p_{ubc}} x_{ubc} - \mu_b x_{ubc} + \iota_{ubc}. \quad (3.24)$$

Nous commençons par le calcul de la dérivée de r_{ubc} :

$$\begin{aligned} \frac{\partial r_{ubc}}{\partial p_{ubc}} &= \frac{w'_b}{\ln(2)(1 + \text{SINR}_{ubc})} \frac{\partial \text{SINR}_{ubc}}{\partial p_{ubc}} \\ &= \frac{w'_b}{\ln(2)(1 + \text{SINR}_{ubc})} \frac{g_{ubc}}{\sigma_b^2 + n_{ubc} + i_{ubc}} \end{aligned}$$

$$= \frac{w'_b}{\ln(2)} \frac{1}{p_{ubc} + h_{ubc}},$$

où :

$$h_{ubc} = \frac{\sigma_b^2 + n_{ubc} + i_{ubc}}{g_{ubc}}. \quad (3.25)$$

Par conséquent, nous obtenons la dérivée de la fonction de Lagrange $\mathcal{L}(\mathbf{P}, \underline{\mu}, \underline{\iota})$:

$$\frac{\partial \mathcal{L}(\mathbf{P}, \underline{\mu}, \underline{\iota})}{\partial p_{ubc}} = \frac{w'_b}{\ln(2)(p_{ubc} + h_{ubc})} x_{ubc} - \mu_b x_{ubc} + \iota_{ubc}. \quad (3.26)$$

En utilisant les équations (3.23) et (3.26), et après quelques manipulations mathématiques, nous déduisons le niveau de puissance alloué à l'utilisateur $u \in \mathcal{U}$ qui est comme suit, pour tout $b \in \mathcal{B}$ et $c \in \mathcal{C}$:

$$p_{ubc} = \frac{w'_b}{\ln(2)(\mu_b - \iota_{ubc})} - h_{ubc}, \quad (3.27)$$

où $0 \leq p_{ubc} \leq p_b$. Nous désignons par p_{ubc}^{min} (resp. p_{ubc}^{max}) le niveau de puissance qui garantit D_u^{min} (resp. D_u^{max}) à l'utilisateur u . Le niveau de puissance minimal alloué à chaque utilisateur est, pour tout $u \in \mathcal{U}$, $b \in \mathcal{B}$ et $c \in \mathcal{C}$:

$$p_{ubc}^* = \mathbf{max} \left(p_{ubc}^{min}, \mathbf{min}(p_{ubc}, p_{ubc}^{max}) \right). \quad (3.28)$$

Il est à noter que les niveaux de puissance alloués sont positifs et les contraintes (P4c) sont respectées. Par conséquent, la matrice du multiplicateur de Lagrange $\underline{\iota}$ est nulle et le multiplicateur de Lagrange μ_b des contraintes (P4b) peut être écrit comme suit, pour tout $b \in \mathcal{B}$:

$$\mu_b = \frac{w'_b}{\ln(2)} \frac{UC}{p_b + \sum_{u \in \mathcal{U}} \sum_{c \in \mathcal{C}} h_{ubc}}, \quad (3.29)$$

où $\mathcal{C} = \mathcal{C}_m$ (resp. $\mathcal{C} = \mathcal{C}_s$) si $b \in \mathcal{B}_m$ (resp. $b \in \mathcal{B}_s$).

Le pseudo-code de l'algorithme WFA est donné dans l'algorithme 6. WFA commence par calculer la puissance requise pour satisfaire les exigences de qualité de service des utilisateurs associés. Ensuite, à chaque itération, WFA calcule le multiplicateur de Lagrange μ_b et met à jour le niveau de puissance alloué p_{ubc}^* jusqu'à la convergence du débit total.

Algorithme 6 Pseudo-code de l'algorithme WFA

- 1: Calculer p_{abc} selon l'équation (3.11), pour tout $u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}$.
 - 2: **répéter**
 - 3: **Pour** u de 1 à U **faire**
 - 4: **Pour** b de 1 à B **faire**
 - 5: **Pour** c de 1 à C **faire**
 - 6: Calculer μ_b et mettre à jour p_{abc}^* selon l'équation (3.28).
 - 7: **Fin Pour**
 - 8: **Fin Pour**
 - 9: **Fin Pour**
 - 10: Calculer le débit total des utilisateurs associés R selon l'équation (3.12).
 - 11: **jusqu'à** jusqu'à la convergence du débit.
-

Les algorithmes proposés GRA et GA nécessitent des informations complètes et ne peuvent pas généralement s'adapter aux réseaux à grande dimension. Des solutions efficaces basées sur l'approche d'apprentissage par renforcement profond sont proposées dans la section suivante.

3.6.4 Solutions d'apprentissage par renforcement profond (DRLA et MDRLA)

Dans ce qui suit, nous introduisons tout d'abord l'approche d'apprentissage profond DQL puis nous détaillons nos solutions DQL centralisée et multiagent.

Notions générales sur l'apprentissage profond QL

L'approche d'apprentissage QL permet à un agent d'apprendre une politique optimale (ou quasi-optimale) en interagissant avec son environnement, sans nécessiter de connaissances préalables sur les probabilités de transition entre les états. Bien que l'efficacité de l'approche QL soit avérée, elle est inadaptée généralement aux espaces d'actions et d'états de grande dimension. Ceci est principalement dû au fait que (i) de nombreuses actions sont rarement choisies et (ii) une grande mémoire pour la table Q est nécessaire pour stocker la valeur Q pour chaque état et chaque action. L'idée clé de l'apprentissage profond est d'utiliser un ou plusieurs réseaux de neurones pour approximer les valeurs état-action Q, où l'état actuel $s^{(t)}$ est donné en entrée et les valeurs Q de toutes les actions possibles dans $s^{(t)}$ sont générées en sortie. La valeur Q de DQL est notée par $Q(s^{(t)}, a^{(t)}; \underline{\theta})$, où $\underline{\theta}$ représente les paramètres de neurones.

Deux réseaux de neurones sont considérés : un réseau de neurones cible avec des paramètres $\underline{\theta}_{target}$

et un réseau de neurones d'entraînement avec des paramètres $\underline{\theta}_{train}$. Le réseau d'entraînement est responsable de l'apprentissage des valeurs de $\underline{\theta}_{train}$, tandis que le réseau cible est utilisé pour prendre des actions. La valeur de $\underline{\theta}_{target}$ est mise à jour à chaque itération t et elle est égale à $\underline{\theta}_{train}$ (Van Hasselt *et al.*, 2016). DQL entraîne chaque réseau de neurones en minimisant une fonction de perte L qui est définie comme l'erreur quadratique entre la valeur prédite et la valeur cible. Cette dernière est donnée par (Graesser et Keng, 2019) :

$$L(\underline{\theta}^{(t)}) = (Q_{target}^{(t)} - Q(s^{(t)}, a^{(t)}; \underline{\theta}_{train}^{(t)}))^2. \quad (3.30)$$

La fonction action-valeur Q du réseau de neurones cible à l'instant t , notée par $Q_{target}^{(t)}$, est définie comme suit (Graesser et Keng, 2019) :

$$Q_{target}^{(t)} = R^{(t)} + \gamma \mathbf{max}_{a'} Q(s^{(t+1)}, a'; \underline{\theta}_{target}^{(t)}). \quad (3.31)$$

Solution d'apprentissage centralisée (DRLA)

Nous modélisons le problème RAP par un processus de décision markovien sous forme d'un quadruplet $(\mathcal{S}, \mathcal{A}, R, \Pr_{ss'}^a)$. \mathcal{S} est l'ensemble d'états possibles, \mathcal{A} est l'ensemble d'actions possibles, $\Pr_{ss'}^a = \Pr(s^{(t+1)} = s' | s^{(t)} = s, a^{(t)} = a)$ est la probabilité de transition de l'état s à l'état s' en effectuant l'action a et R est la fonction de récompense immédiate reçue, équivalente au débit de données descendant (Bellman, 1957). Nous supposons la présence d'un agent de contrôle central, noté par $b_c \in \mathcal{B}$, qui peut être une mBS ou une sBS. À chaque itération t , l'agent b_c observe son état $s^{(t)} \in \mathcal{S}$ puis prend une action $a^{(t)} \in \mathcal{A}$ qui inclut (i) l'association des utilisateurs aux stations de base, (ii) l'attribution des sous-canaux et (iii) la sélection de la technique d'accès multiple pour chaque sous-canal. L'espace d'actions est la somme de toutes les combinaisons possibles, ce qui est égal au produit cartésien de $\sum_{i=1}^{\min(U, \sum_{c \in \mathcal{C}_s} q_{bc})} \binom{U}{i}$, $\forall b \in \mathcal{B}_s$ et $\sum_{i=1}^{\min(U, \sum_{c \in \mathcal{C}_m} q_{bc})} \binom{U}{i}$, $\forall b \in \mathcal{B}_m$. Ensuite, l'agent calcule les niveaux de puissance de transmission des utilisateurs de façon à maximiser le débit de données et à garantir les exigences de qualité de service des utilisateurs.

Deux états sont possibles pour l'agent b_c pour indiquer si la récompense moyenne est améliorée : $s^{(t)} = 1$ signifie que la récompense moyenne a augmenté ou a gardé la même valeur entre deux itérations et $s^{(t)} = 0$ sinon. Les probabilités de passer d'un état à un autre et de prendre certaines actions dépendent de l'environnement et sont difficiles à déterminer. Une fois que l'agent b_c prend

une action, il passe de l'état actuel $s^{(t)} \in \mathcal{S}$ à un autre état $s^{(t+1)} \in \mathcal{S}$. L'état $s^{(t+1)}$ dépend uniquement de l'état actuel $s^{(t)}$ et de l'action choisie $a^{(t)}$. Autrement dit, chaque état dans \mathcal{S} suit la propriété de Markov et donc chaque état est indépendant de toutes les actions et les états passés.

Pour chaque action choisie, l'agent b_c vérifie les contraintes de puissance et de qualité de service. Si toutes les contraintes sont respectées, alors l'action est réalisable. Sinon, l'action n'est pas réalisable et une récompense nulle est considérée. Comme RAP maximise le débit total de données, la récompense du système instantanée est égale à la somme de tous les débits des stations de base et dont l'expression est donnée par l'équation (3.12). Enfin, l'agent envoie l'action choisie et les niveaux de puissance aux autres stations de base. Nous supposons dans ce qui suit que chaque station de base ne connaît que le gain du canal entre elle-même et tous les utilisateurs, tandis que l'agent, autrement dit la station de base contrôleur, connaît le gain du canal entre toutes les stations de base et tous les utilisateurs.

Le pseudo-code de l'algorithme proposé, appelé DRLA, est donné dans l'algorithme 7. DRLA est basé sur un entraînement centralisé. Ceci améliore la stabilité de l'apprentissage et simplifie la mise en œuvre. L'agent b_c entraîne le réseau de neurones en exploitant les expériences acquises. Dans les lignes 4-6, l'agent b_c sélectionne une action selon la stratégie ϵ -gloutonne (c'est-à-dire l'agent choisit la meilleure action avec une probabilité de $1 - \epsilon$ et sélectionne une action aléatoire parmi les actions possibles avec une probabilité de ϵ), calcule la récompense et met à jour l'état actuel. Ensuite, DRLA stocke toutes les expériences passées dans la mémoire de relecture (en anglais : *replay buffer*) pour améliorer l'apprentissage. Enfin, DRLA calcule la fonction de perte et met à jour les paramètres des neurones. La complexité de DRLA est liée à celle du réseau profond. Étant donné que la structure considérée est entièrement connectée, la complexité est $\mathcal{O}(\sum_{i \in [1, L-1]} N_i N_{i+1})$, où L est le nombre de couches et N_i est le nombre de neurones dans la couche i . Puisque les solutions distribuées sont généralement plus évolutives que les solutions centralisées et plus adaptées aux réseaux cellulaires à grande échelle, une solution multiagent basée sur l'approche DQL est proposée ci-dessous.

Solution d'apprentissage multiagent (MDRLA)

La solution d'apprentissage multiagent proposée (appelée MDRLA) considère les stations de base comme des agents indépendants. Chaque agent apprend sa propre politique qui maximise son débit

Algorithme 7 Pseudo-code de l'algorithme DRLA

- 1: **Tant que** pas de convergence **faire**
 - 2: Commencer un nouvel épisode et réinitialiser l'état actuel.
 - 3: **Pour** chaque itération t **faire**
 - 4: Sélectionner l'action $a^{(t)}$ en utilisant la stratégie ϵ -gloutonne.
 - 5: Effectuer l'action choisie et calculer la récompense.
 - 6: Mettre à jour l'état actuel $s^{(t)}$.
 - 7: Stocker $(s^{(t)}, a^{(t)}, s^{(t+1)}, R^{(t)})$ dans la mémoire de relecture.
 - 8: Échantillonner un mini-lot aléatoire de transitions.
 - 9: Calculer la fonction de perte selon (3.30).
 - 10: Effectuer une descente de gradient pour minimiser cette perte et mettre à jour les poids du réseau.
 - 11: **Fin Pour**
 - 12: **Fin Tant que**
-

de données. L'action choisie comprend (i) la sélection des utilisateurs à associer, (ii) l'allocation des sous-canaux et (iii) le choix de la technique d'accès multiple pour chaque sous-canal. L'agent est représenté par un quadruplet $(\mathcal{S}, \mathcal{A}, R, \text{Pr}_{ss'}^a)$, où \mathcal{A}_b est l'ensemble d'actions, $\text{Pr}_{ss'}^a$ est la probabilité de transition, R est la fonction de récompense immédiate reçue quand il passe de l'état s à l'état s' et qui est égale au débit de l'agent $b \in \mathcal{B}$ ($\sum_{u \in \mathcal{U}} \sum_{c \in \mathcal{C}} r_{ubc} x_{ubc}$). Nous supposons que chaque station de base ne connaît que le gain du canal entre elle-même et tous les utilisateurs dans le réseau. Pour chaque action choisie, l'agent vérifie les contraintes de puissance et de qualité de service des utilisateurs à associer. Ensuite, il envoie des invitations d'association aux utilisateurs sélectionnés. Chaque utilisateur ne peut accepter qu'une seule association. Ainsi, si plusieurs invitations sont envoyées au même utilisateur, ce dernier choisit la station de base ayant les meilleures conditions du canal. Enfin, chaque agent utilise l'algorithme WFA pour calculer les niveaux de puissance et calcule sa récompense.

3.7 Résultats de simulation

Nous considérons une zone géographique de taille 200 m · 200 m dans laquelle les utilisateurs et les stations de base sont distribués d'une manière aléatoire et uniforme. Les sBSs et les mBSs déployées sont colocalisées. Les largeurs de bande disponibles pour les sBSs et les mBSs sont respectivement 10 MHz et 100 MHz. La puissance du bruit AWGN est de -174 dBm/Hz. La puissance disponible pour chaque station de base est de 1 W. Sauf indication contraire, le rapport des seuils de débit de données des utilisateurs D_u^{max}/D_u^{min} est fixé à 50. La fonction d'affaiblissement de propagation entre l'utilisateur u et la mBS b est étudiée dans (Akdeniz *et al.*, 2014) et elle est donnée par

$10 \log_{10}(f_c(d_{ub})) = \psi + 10\kappa \log_{10}(d_{ub}) + \xi_c$, où $\xi_c \sim \mathcal{N}(0, v^2)$. Les autres valeurs de simulation sont données dans le tableau 3.2.

Comme le montre la figure 3.2, l’architecture du modèle d’apprentissage adopté est composé d’une couche d’entrée, d’une couche de sortie et de trois couches cachées entièrement connectées contenant respectivement 512, 256 et 64 neurones. La fonction d’activation unité linéaire rectifiée (en anglais : *rectified linear unit*, ReLU) est utilisée par les couches cachées. Cette fonction permet de remplacer les résultats négatifs par des zéros. Pour la mise à jour des poids du réseau de neurones, la descente du gradient stochastique est obtenue en utilisant l’approche de propagation quadratique moyenne (en anglais : *Root Mean Square Propagation*, RMSProp). Le problème d’optimisation (P2) est modélisé et résolu avec la bibliothèque GEKKO de Python et le solveur APOPT (en anglais : *advanced Process OPTimizer*).

TABLE 3.2 – Notations et paramètres de simulation.

Notations	Paramètres	Valeurs
δ	Exposant d’affaiblissement	4
d_0	Distance de référence	1 m
β	Constante capturant des effets du système et de la transmission	10^3
ψ, κ, v	Paramètres du modèle d’affaiblissement pour les transmissions NLoS (Akdeniz <i>et al.</i> , 2014)	$\psi = 86.6, \kappa = 2.45$ $v = 8$ dB
ψ, κ, v	Paramètres du modèle d’affaiblissement pour les transmissions LoS (Akdeniz <i>et al.</i> , 2014)	$\psi = 69.8, \kappa = 2$ $v = 5.8$ dB
α	Taux d’apprentissage	0.009
γ	Facteur d’actualisation	0,95
L	Nombre de couches cachées	3
N_i	Nombre de neurones dans la couche i ($i \in 1, \dots, L$)	$N_1 = 512, N_2 = 256$ $N_3 = 64$

3.7.1 Ajustement des hyperparamètres du réseau de neurones et convergence

Dans la suite, nous ajustons les différents hyperparamètres, tels que le taux d’apprentissage et l’architecture du réseau de neurones, et dont les valeurs sont utilisées pour contrôler le processus d’apprentissage automatique. Ensuite, nous évaluons les performances des solutions proposées pour des utilisateurs stationnaires et mobiles. Nous notons que les simulations ont été effectuées sur un processeur Intel Core i7 avec 8 Go de mémoire.

Tout d’abord, nous évaluons la convergence de DRLA pour différents taux d’apprentissage. Comme

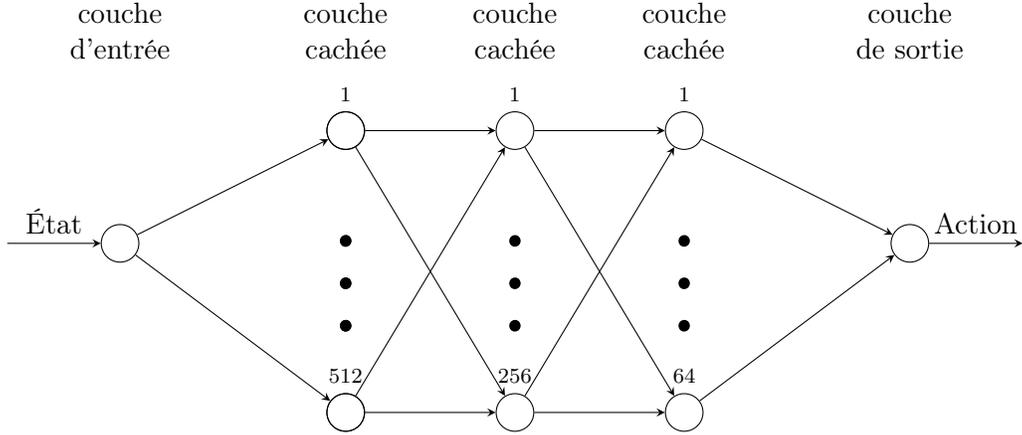


FIGURE 3.2 – Architecture du modèle d'apprentissage.

le montre la figure 3.3, nous remarquons que plus le taux d'apprentissage α diminue, plus il faut d'itérations d'entraînement pour atteindre la convergence. De plus, nous observons que les meilleures performances de DRLA sont obtenues avec un taux d'apprentissage de 0.009, et qu'une augmentation ou une diminution de cette valeur ne conduit pas à une amélioration des performances. Étant donné qu'un taux d'apprentissage élevé peut conduire à un optimum local au lieu d'un optimum global, dans ce chapitre, la valeur de α est choisie égale à 0.009.

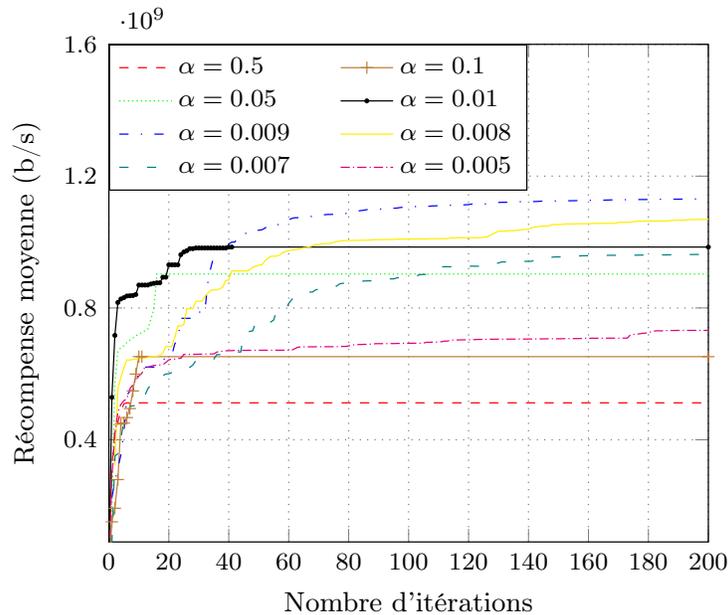


FIGURE 3.3 – L'impact du taux d'apprentissage sur les résultats de l'algorithme DRLA ($L = 3$, $N_1 = 512$, $N_2 = 256$ et $N_3 = 64$).

La figure 3.4 montre les récompenses moyennes pour différents nombres de couches cachées. Nous observons que DRLA avec 5 couches cachées converge plus rapidement qu’avec 4 et 3 couches cachées. Ceci est dû au fait que lorsque le nombre de couches est grand, le risque de surapprentissage augmente et donc le modèle peut être capable de mémoriser les données d’apprentissage au lieu de les généraliser. Cependant, lorsque le nombre de couches est trop petit, moins d’itérations d’entraînement sont nécessaires pour la convergence et, par conséquent, DRLA peut converger vers un optimum local plutôt qu’un optimum global. Dans le reste de ce chapitre, l’architecture de DRLA utilisée comporte 3 couches cachées.

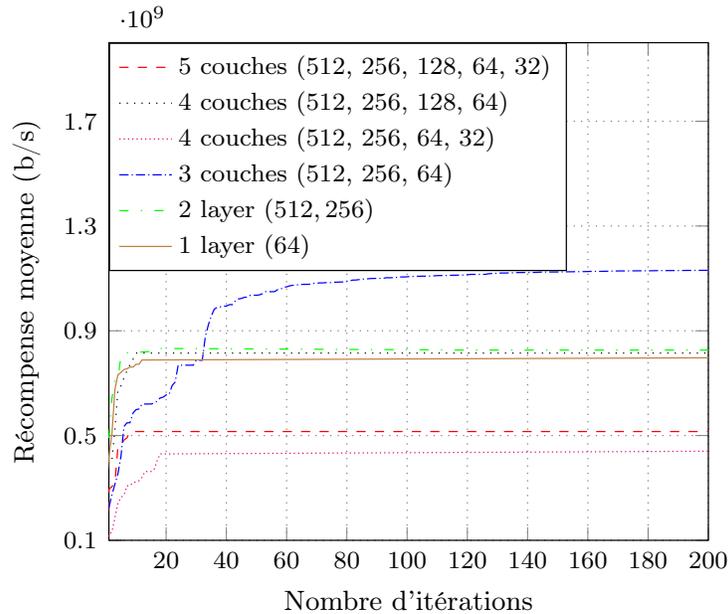


FIGURE 3.4 – L’impact du nombre de couches cachées sur les résultats de l’algorithme DRLA ($\alpha = 0.009$).

La figure 3.5 montre que l’augmentation du nombre de neurones augmente le nombre d’itérations nécessaires pour la convergence et, par conséquent, la complexité computationnelle. Ceci mène à l’apprentissage des représentations plus complexes. Nous pouvons également remarquer qu’une convergence plus rapide peut être obtenue lorsque $N_1 = 64$, $N_2 = 32$ et $N_3 = 8$. Cependant, afin d’éviter de tomber dans un optimum local et de prévenir le problème de surapprentissage, nous considérons un réseau de 3 couches cachées, où $N_1 = 512$, $N_2 = 256$ et $N_3 = 64$.

La figure 3.6 illustre la convergence de GA pour un réseau cellulaire composé de 2 stations de base et 10 utilisateurs. Nous pouvons voir que GA trouve la meilleure solution d’une manière itérative en

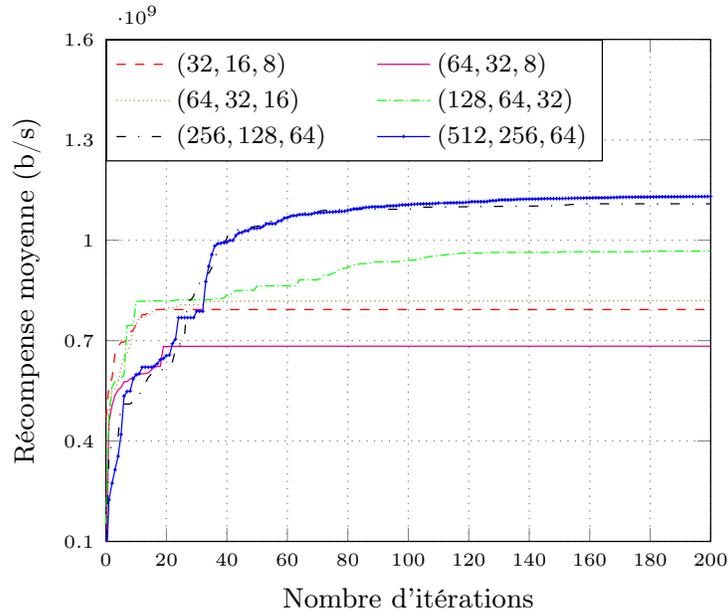


FIGURE 3.5 – L’impact du nombre de neurones sur les résultats de l’algorithme DRLA ($\alpha = 0.009$, $U = 10$, $B_s = B_m = 1$ et $C_s = C_m = 2$).

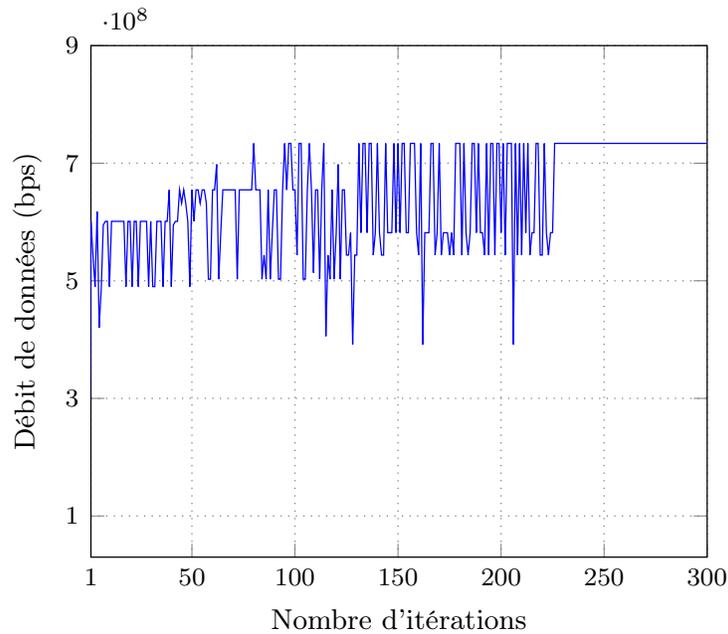


FIGURE 3.6 – Convergence de GA ($U = 10$, $B_s = B_m = 1$, $C_s = C_m = 2$ et $D_u^{max}/D_u^{min} = 20$).

explorant l’espace de recherche, ce qui correspond à son caractère évolutif. Nous pouvons également remarquer que GA converge après un nombre limité d’itérations (environ 220 itérations), ce qui prouve la rapidité de convergence de GA.

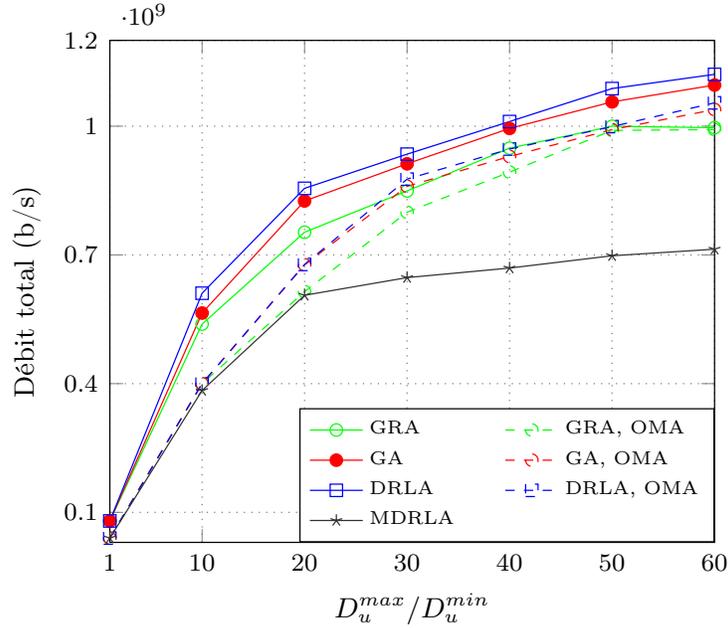


FIGURE 3.7 – Comparaison des performances des algorithmes proposés en termes de débit de données ($U = 10$, $B_s = B_m = 1$ et $C_s = C_m = 2$).

3.7.2 Comparaison des performances des algorithmes proposés

Les figures 3.7 et 3.8 comparent les performances des algorithmes proposés pour différentes valeurs de D_u^{max}/D_u^{min} . Il est clairement visible que DRLA atteint le débit maximal pour les différentes techniques d'accès (uniquement OMA et OMA avec MSC-NOMA). De plus, le débit maximal de DRLA reste proche de 11 Gb/s quand D_u^{max}/D_u^{min} est supérieur à 50. Ceci est dû à la limitation de puissance de transmission disponible pour les stations de base, comme indiqué dans les contraintes (P5f). Comparé à MDRLA, DRLA offre de meilleures performances en termes de débit, car MDRLA est une approche complètement distribuée et donc il n'y a aucune coordination entre les stations de base. Cependant, il est important de noter que bien que les performances de MDRLA soient inférieures aux autres solutions centralisées (GA, GRA, DRLA), MDRLA est plus évolutif et s'adapte mieux à des réseaux à grande échelle. En outre, GA est plus performant que GRA, avec un écart de performance inférieur à 12% lorsque les deux techniques d'accès OMA et MSC-NOMA sont considérées. Nous concluons que la technique d'accès MSC-NOMA permet d'obtenir un débit plus élevé surtout pour de petites valeurs de D_u^{max}/D_u^{min} car elle permet de satisfaire plus d'utilisateurs que la technique d'accès OMA.

La figure 3.8 montre que lorsque les deux techniques d'accès OMA et MSC-NOMA sont utilisées,

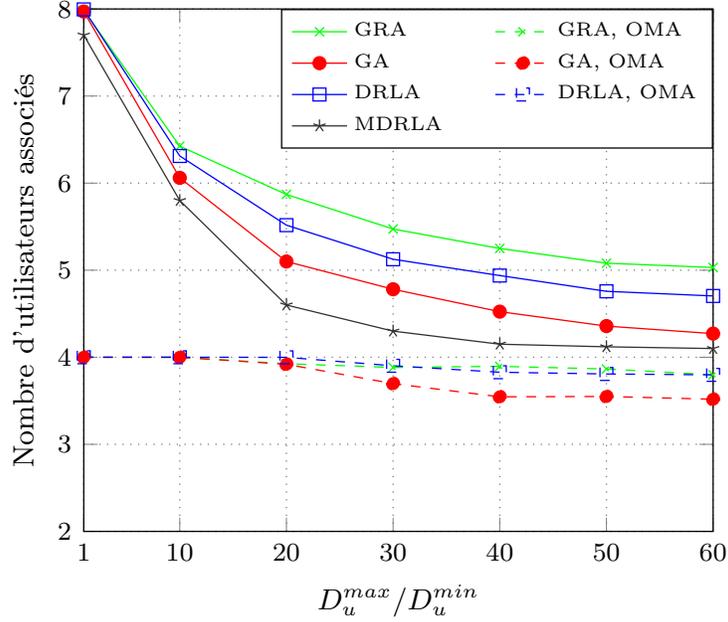


FIGURE 3.8 – Comparaison des performances des algorithmes proposés en termes de nombre d'utilisateurs associés ($U = 10$, $B_m = B_s = 1$ et $C_m = C_s = 2$).

plus d'utilisateurs sont associés puisque plusieurs utilisateurs peuvent partager le même canal. De plus, le nombre d'utilisateurs associés diminue lorsque D_u^{max}/D_u^{min} augmente. Nous observons aussi que l'écart de performance entre DRLA et GA est inférieur à 10% et que le nombre d'utilisateurs associés de GRA est presque constant, 4 utilisateurs quand la technique d'accès OMA est utilisée. Cela est dû au fait que GRA commence d'abord par maximiser le nombre d'utilisateurs associés afin d'optimiser le débit total de données. En outre, MDRLA associe moins d'utilisateurs que DRLA lorsque les deux techniques d'accès OMA et MSC-NOMA sont considérées, et plus d'utilisateurs que DRLA avec OMA. En fait, dans MDRLA, la technique d'accès MSC-NOMA est rarement sélectionnée en raison de l'absence d'une station de base de contrôle connaissant les conditions du canal entre toutes les stations de base et tous les utilisateurs afin d'effectuer le regroupement des utilisateurs partageant le même canal.

Pour mesurer l'équité entre les utilisateurs en termes de débit de données, nous utilisons l'indice de Jain. Cet indice mesure la proportion du débit moyen attribué à chaque utilisateur par rapport au débit total (Jain *et al.*, 1984). Par conséquent, l'indice de Jain est calculé comme suit :

$$J = \frac{R^2}{U \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} r_{ubc}^2 x_{ubc}}. \quad (3.32)$$

Il est à noter qu'un indice de Jain de 1 indique une répartition parfaitement équitable, tandis qu'un indice plus proche de 0 indique une répartition inéquitable.

La figure 3.9 montre l'indice de Jain pour 10 utilisateurs. Il est clair que lorsque les deux techniques d'accès OMA et MSC-NOMA sont considérées, l'indice de Jain de GRA, GA et DRLA est supérieur à 0,4 quand D_u^{max}/D_u^{min} est inférieur à 30 et diminue quand D_u^{max}/D_u^{min} augmente. Comme prévu, les solutions proposées fournissent une plus grande équité pour de petites valeurs de D_u^{max}/D_u^{min} puisque plus d'utilisateurs sont satisfaits.

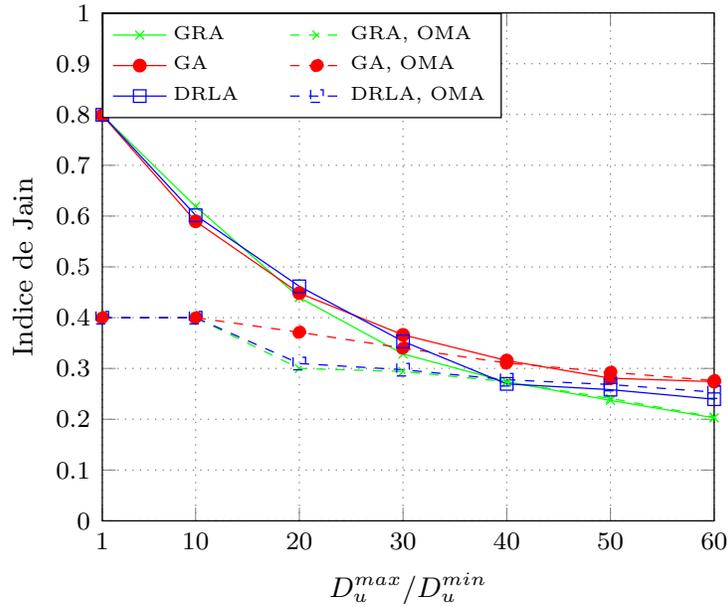


FIGURE 3.9 – Indice de Jain en fonction de D_u^{max}/D_u^{min} ($U = 10$, $B_s = B_m = 1$ et $C_s = C_m = 2$).

Dans ce qui suit, nous illustrons l'impact de la mobilité sur les performances des algorithmes proposés. Nous considérons 10 utilisateurs mobiles, $B_m = B_s = 1$ et $C_m = C_s = 2$. Nous supposons que chaque utilisateur peut se déplacer indépendamment des autres utilisateurs à une vitesse fixe. Afin de garder le même nombre d'utilisateurs dans la zone géographique, quand un utilisateur atteint la limite de la zone, il inverse la direction.

Le tableau 3.3 montre que DRLA surpasse GRA et GA en termes de débit de données. Aussi, toutes les solutions proposées (c'est-à-dire GRA, GA et DRLA) atteignent, comme prévu, un débit total descendant moins élevé dans les scénarios mobiles. Contrairement à GRA et GA où les solutions

d'allocation de ressources et d'association des utilisateurs aux stations de base ne sont pas mémorisées pour être exploitées dans les futures exécutions, DRLA stocke les expériences passées dans la mémoire de relecture. Les expériences accumulées sont utilisées pour s'adapter au mouvement des utilisateurs, c'est-à-dire aux changements de gain du canal des utilisateurs. En outre, lorsque la vitesse des utilisateurs est fixée à 6 km/h, l'écart de performance entre GA et GRA est plus que à 20% et entre GA et DRLA est plus que 28%. Cela prouve les performances de DRLA dans les scénarios mobiles.

TABLE 3.3 – Impact de la mobilité sur les performances des algorithmes proposés.

Vitesse de mobilité	0 km/h	3 km/h	6 km/h
Débit total de GRA	7.1284261e+08	5.1953863e+08	4.6206875e+08
Débit total de GA	8.6033306e+08	6.7224280e+08	5.8182660e+08
Débit total de DRLA	9.2482546e+08	8.3500551e+08	8.1163318e+08

3.8 Conclusions

Dans ce chapitre, nous avons étudié le problème d'allocation des ressources, notamment l'association des utilisateurs aux stations de base, l'allocation de puissance et l'attribution des sous-canaux et de la technique d'accès dans un réseau cellulaire à bandes multiples utilisant les deux bandes de fréquences sub-6 GHz et mmWave ainsi que les deux techniques d'accès OMA et MSC-NOMA. Nous avons modélisé le problème comme un programme non linéaire partiellement en nombres entiers et nous avons étudié sa NP-difficulté en utilisant la technique par réduction. Pour le résoudre, nous avons modélisé le problème sous forme d'un processus de décision de Markov et nous avons proposé un algorithme centralisé et un autre distribué multiagent basés sur l'approche d'apprentissage par renforcement profond. À des fins de comparaison, un algorithme glouton et un autre génétique ont été proposés. Enfin, nous avons démontré les performances des solutions d'apprentissage par renforcement profond proposées par rapport aux solutions de référence. Nous avons aussi prouvé les avantages de l'utilisation des techniques d'accès OMA et NOMA dans les systèmes à bandes multiples en termes de débit de données descendant, de nombre d'utilisateurs associés et d'équité entre les utilisateurs.

CHAPITRE 4

EFFICACITÉ ÉNERGÉTIQUE DANS UN RÉSEAU À BANDES MULTIPLES

Dans ce chapitre, nous étudions le problème d'efficacité énergétique dans un réseau cellulaire utilisant les trois bandes de fréquences sub-6 GHz, mmWave et THz.

Ce travail a été soumis dans le journal *IEEE Transactions on Wireless Communications* en septembre 2023. Des résultats préliminaires de ce travail ont été présentés à la conférence *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Toronto, Canada, 2023* (Chaieb *et al.*, 2023b).

4.1 Résumé

Comme mentionné dans les chapitres précédents, l'utilisation des communications à bandes multiples est une solution prometteuse pour faire face à la pénurie du spectre et aux exigences des diverses applications mobiles. La gestion des ressources radio demeure cruciale dans le but d'optimiser les performances des systèmes de communication sans fil, en particulier lorsqu'il s'agit de répondre aux diverses exigences de qualité de service. Dans ce chapitre, nous abordons le problème d'optimisation conjointe d'allocation des ressources dans un réseau cellulaire utilisant trois bandes de fréquences sub-6 GHz, mmWave et THz. L'objectif est de maximiser l'efficacité énergétique du système tout en respectant les contraintes de qualité de service et de puissance. Une nouvelle mesure d'efficacité énergétique qui quantifie le nombre d'utilisateurs satisfaits (ou associés) par unité de puissance est considérée. Le problème conjoint d'association des utilisateurs aux stations de base, d'allocation des canaux et de puissance est formulé comme un programme partiellement en nombres entiers, impliquant à la fois des variables de décision réelles et d'autres binaires. Nous montrons que ce problème est NP-difficile et nous proposons des solutions algorithmiques peu complexes. La première solution consiste à diviser le problème en deux sous-problèmes que nous résolvons d'une manière itérative jusqu'à atteindre la convergence des résultats. Nous commençons par résoudre le problème d'association des utilisateurs aux stations de base et d'allocation des canaux tout en respectant les contraintes du système. Ensuite, nous abordons le problème d'allocation de puissance qui a pour objectif de maximiser l'efficacité énergétique du réseau en se basant sur les résultats du premier sous-problème. La deuxième solution est une approche multiagent basée sur l'apprentissage

par renforcement profond QL. Chaque utilisateur est un agent qui apprend la meilleure association indépendamment des autres utilisateurs en maximisant son score. La récompense est calculée d'une manière distribuée et repose uniquement sur le score de l'utilisateur. Les résultats de simulation illustrent l'efficacité de la solution d'apprentissage par renforcement profond et de la solution itérative, en comparaison avec d'autres algorithmes basés sur l'approche gloutonne.

4.2 Introduction

Les systèmes de communication THz offrent une large bande passante, ce qui les rend parfaitement adaptés aux communications à bande ultra-large et à haut débit. Cependant, les communications THz sont confrontées à plusieurs problèmes de propagation affectant la qualité des signaux, notamment l'atténuation atmosphérique, la sensibilité aux obstacles et la dispersion des signaux. Une des approches possibles pour maintenir le rapport signal sur interférence et bruit (SINR) à un niveau donné est d'augmenter la puissance de transmission. Néanmoins, la puissance est une ressource limitée et cette augmentation peut entraîner des défis de gestion de l'énergie et de conception des composants radio. Par conséquent, le problème conjoint d'optimisation des ressources dans les systèmes de communication à large bande passante est donc un problème complexe et crucial.

4.2.1 Revue de littérature

Les travaux de recherche les plus importants liés aux réseaux cellulaires à bandes multiples et à l'optimisation de l'efficacité énergétique sont discutés ci-après. Les auteurs dans (Muhammed *et al.*, 2023) ont étudié le problème d'allocation de puissance et de partition de la bande passante dans les réseaux hétérogènes mmWave utilisant la technique d'accès non orthogonal. Les auteurs ont optimisé l'efficacité énergétique, définie comme étant le rapport entre le débit et la puissance de transmission totale, sous des contraintes de puissance et de débit minimal. Les auteurs ont transformé le problème initial, qui est convexe, en un autre problème équivalent quasi-convexe et ils ont proposé des algorithmes distribués d'allocation de puissance. Dans (Jin *et al.*, 2022), les auteurs ont utilisé la technologie de multiconnectivité dans un réseau cellulaire mmWave permettant à un utilisateur de s'associer simultanément à la station de base macro-cellule et aux petites cellules. Ils ont formulé le problème d'association des utilisateurs aux stations de base et d'allocation de puissance sous des contraintes de qualité de service. L'objectif était d'équilibrer la charge entre les stations de base et de maximiser l'efficacité énergétique. Les auteurs ont proposé un schéma hybride

de précodage et des algorithmes distribués d'allocation de puissance. Les travaux (Muhammed *et al.*, 2023; Jin *et al.*, 2022) ont proposé des solutions algorithmiques d'allocation des ressources dans les réseaux mmWave sans utiliser l'apprentissage machine.

Dans (Wang *et al.*, 2021), les auteurs ont étudié le problème conjoint d'association des utilisateurs aux stations de base et d'allocation de puissance dans un réseau cellulaire utilisant les bandes de fréquences mmWave et THz. Ils ont formulé le problème comme un programme d'optimisation non linéaire et non convexe avec l'objectif de maximiser une fonction d'utilité du réseau sous des contraintes de puissance, de qualité de service et de distances maximales entre les stations de base et les utilisateurs. Les auteurs ont décomposé le problème en deux sous-problèmes : association des utilisateurs aux stations de base et allocation de puissance. Les deux sous-problèmes sont résolus d'une manière itérative en se basant sur la dualité lagrangienne et la méthode de descente du gradient. Les auteurs ont montré que la coexistence de communications mmWave et THz peut grandement améliorer le débit de données du système. Dans (Wang et Gursoy, 2021), les auteurs ont considéré un réseau hétérogène composé d'une station de base macro-cellule utilisant la bande de fréquences sub-6 GHz, des petites cellules utilisant la bande de fréquences mmWave et des points d'accès opérant sur la bande de fréquences THz. En se basant sur la géométrie stochastique, les auteurs ont montré que le déploiement dense de points d'accès THz augmente la puissance du signal reçu. Cependant, cela diminue la probabilité de couverture du rapport signal sur interférence et bruit.

Dans l'article (Rasti *et al.*, 2022), les auteurs ont étudié le problème conjoint de sélection de réseau et d'allocation des sous-canaux dans un réseau cellulaire 6G à bandes multiples. Le réseau étudié offre une connectivité aérienne, terrestre, satellitaire et sous-marine afin de fournir aux utilisateurs terrestres et aériens des services de communication mobile à large bande améliorés (en anglais : *further enhanced mobile broadband*) ainsi que des communications ultra-fiables à faible latence améliorées (en anglais : *enhanced ultra-reliable low-latency communications*). Les auteurs ont proposé des solutions algorithmiques basées sur l'apprentissage profond, notamment QL, double QL et dueling QL. Les résultats de simulation ont illustré l'efficacité des réseaux à bandes multiples en termes de débit de données. En outre de ne pas étudier la complexité computationnelle des problèmes formulés dans (Rasti *et al.*, 2022; Wang *et al.*, 2021), les auteurs ont considéré des fonctions objectif et des modèles du système différents des nôtres.

Dans (Liao *et al.*, 2019), les auteurs ont étudié trois fonctions objectif visant à maximiser l'efficacité spectrale, l'efficacité énergétique et l'équité. Ils ont considéré un réseau ultra-dense avec des informations limitées sur le canal de transmission et ils ont proposé une solution multiagent basée sur l'apprentissage par renforcement profond. Dans (Iqbal *et al.*, 2021), les auteurs ont étudié le problème d'allocation de ressources avec activation et désactivation des stations. Ils ont proposé des solutions d'apprentissage par renforcement dans le but de maximiser l'efficacité énergétique du réseau. Dans (Adam *et al.*, 2022), les auteurs ont présenté une solution itérative et une autre basée sur les réseaux de neurones convolutifs profonds pour résoudre le problème d'allocation des ressources dans les réseaux cellulaires utilisant la technique d'accès non orthogonal. D'autre part, les auteurs dans (Sharma et Kumar, 2022; Mughees *et al.*, 2020) ont optimisé l'efficacité énergétique dans les réseaux cellulaires traditionnels, où seulement la communication sur la bande de fréquences sub-6 GHz est considérée. Ils ont proposé des algorithmes d'allocation de ressources basés sur la technique d'apprentissage par renforcement profond. Dans tous les travaux mentionnés, les auteurs n'ont pas considéré la même définition de l'efficacité énergétique que la notre et dans (Adam *et al.*, 2022) les auteurs n'ont pas utilisé l'apprentissage par renforcement. De plus, les auteurs n'ont pas étudié la complexité des problèmes formulés dans (Adam *et al.*, 2022; Sharma et Kumar, 2022; Iqbal *et al.*, 2021; Mughees *et al.*, 2020; Liao *et al.*, 2019).

4.2.2 Contributions

Dans ce chapitre, nous étudions le problème d'allocation des ressources dans un réseau cellulaire composé des stations de base fonctionnant sur l'une des trois bandes de fréquences : sub-6 GHz, mmWave et THz. Nos contributions sont résumées comme suit :

- Nous proposons une nouvelle métrique d'efficacité énergétique qui diffère de la définition considérée dans la plupart des recherches actuelles. Plutôt que de se concentrer sur le débit de transmission par unité de puissance, notre définition se focalise sur le nombre d'utilisateurs satisfaits par unité de puissance.
- Nous modélisons conjointement l'association des utilisateurs aux stations de base, l'allocation des canaux et de puissance dont le but de maximiser l'efficacité énergétique du système. Deux exigences de qualité de service par utilisateur sont définies par des seuils de débit de transmission et du rapport signal sur interférence et bruit. Nous formulons le problème sous forme d'un programme non linéaire partiellement en nombres entiers et nous démontrons qu'il

est dans la classe NP-difficile par restriction. Par conséquent, il est difficile de le résoudre d'une manière optimale dans un temps raisonnable. Pour relever ce défi, nous proposons des solutions algorithmiques efficaces.

- Nous décomposons le problème en deux sous-problèmes, que nous résolvons d'une manière itérative. Tout d'abord, nous considérons une répartition égale de la puissance disponible entre les canaux. Ensuite, nous abordons le problème d'association et d'allocation des canaux tout en tenant compte les contraintes de qualité de service des utilisateurs. La solution du premier sous-problème constitue une solution initiale et réalisable. Enfin, nous résolvons le problème d'allocation de puissance afin de maximiser l'efficacité énergétique du réseau. Nous répétons d'une manière itérative la résolution des sous-problèmes jusqu'à ce que l'efficacité énergétique converge.
- Nous proposons une solution d'apprentissage multiagent basée sur l'apprentissage par renforcement profond QL. Chaque utilisateur est considéré comme un agent indépendant, capable d'apprendre et de converger vers une solution d'allocation de ressources qui optimise l'efficacité énergétique. La fonction de récompense considérée repose sur le score de chaque utilisateur, ce qui rend notre solution entièrement distribuée sans échange d'informations.
- Nous proposons une solution basée sur l'approche gloutonne centralisée et une autre distribuée efficaces pour des fins de comparaison. Enfin, nous illustrons les performances des algorithmes proposés par simulation et nous montrons que la coexistence des communications à bandes multiples améliore l'efficacité énergétique du réseau.

4.2.3 Organisation

Le reste de ce chapitre est organisé comme suit. Dans la section 4.3, nous introduisons le modèle du système ainsi que la nouvelle métrique d'efficacité énergétique. La section 4.4 formule mathématiquement le problème et la section 4.5 étudie sa complexité. La solution itérative et les algorithmes gloutons sont présentés respectivement dans les sous-sections 4.6.1 et 4.6.2. La sous-section 4.6.3 détaille la solution d'apprentissage par renforcement profond. Enfin, la section 4.7 montre les performances des solutions proposées utilisant les simulations Monte-Carlo et la section 4.8 conclut ce chapitre.

4.3 Modèle du système

Comme illustré dans la figure 4.1, nous considérons la coexistence de B_s stations de base sub-6 GHz (sBSs), B_m stations de base mmWave (mBSs) et B_z stations de base THz (zBSs) dans un réseau cellulaire. Les U utilisateurs sont uniformément répartis dans la zone géographique. Soit \mathcal{U} l'ensemble d'utilisateurs et $\mathcal{B} = \mathcal{B}_s \cup \mathcal{B}_m \cup \mathcal{B}_z$ l'ensemble de stations de base, où \mathcal{B}_s , \mathcal{B}_m et \mathcal{B}_z représentent respectivement les ensembles de sBSs, mBSs et zBSs. De plus, $\mathcal{C} = \mathcal{C}_s \cup \mathcal{C}_m \cup \mathcal{C}_z$ désigne l'ensemble de canaux, où \mathcal{C}_s , \mathcal{C}_m et \mathcal{C}_z représentent respectivement les ensembles des canaux sub-6 GHz, mmWave et THz. La transmission sur le lien descendant est considérée. Chaque utilisateur peut être associé à une seule station de base (sBS, mBS ou zBS) en utilisant un seul canal, d'une manière à ce que toutes ses exigences de qualité de service soient satisfaites. Chaque station de base peut simultanément associer plusieurs utilisateurs sur différents canaux. Les utilisateurs ont différentes exigences de qualité de service. Ces exigences sont définies par un débit élevé pour les services à large bande passante (p. ex., réalité augmentée, réalité virtuelle, etc.) et une haute fiabilité (p. ex., conduite autonome, usines et industries intelligentes, etc.) (Saad *et al.*, 2019; Liu *et al.*, 2023).

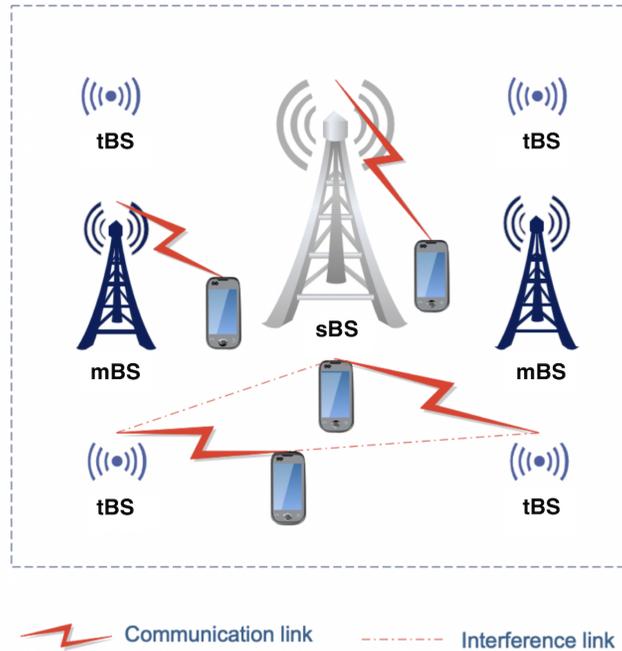


FIGURE 4.1 – Un exemple du modèle du système ($B_s = 1$, $B_m = 2$ et $B_z = 4$).

Pour formuler le problème mathématiquement, nous définissons la matrice binaire d'association $\mathbf{X} = [x_{u,b}^c]$, où $x_{u,b}^c = 1$ si l'utilisateur $u \in \mathcal{U}$ est associé à la station de base $b \in \mathcal{B}$ en utilisant le

canal $c \in \mathcal{C}$ et $x_{u,b}^c = 0$ sinon. Le rapport signal sur interférence et bruit entre l'utilisateur u et la station de base b est exprimé comme suit, pour tout $u \in \mathcal{U}$, $b \in \mathcal{B}$, $c \in \mathcal{C}$:

$$\gamma_{u,b}^c = \frac{p_{u,b}^c h_{u,b}^c}{w_b^c \sigma^2 + i_{u,b}^c}, \quad (4.1)$$

où $p_{u,b}^c$ représente le niveau de puissance alloué par la station de base b , $h_{u,b}^c$ est le coefficient du canal entre l'utilisateur u et la station de base b et σ^2 est la densité de puissance du bruit. L'interférence intercellules $i_{u,b}^c$ est exprimée comme suit, pour tout $u \in \mathcal{U}$, $b \in \mathcal{B}$, $c \in \mathcal{C}$:

$$i_{u,b}^c = \sum_{u' \in \mathcal{U}'} \sum_{b' \in \mathcal{B}'} p_{u',b'}^c h_{u',b'}^c x_{u',b'}^c, \quad (4.2)$$

où $\mathcal{U}' = \mathcal{U} \setminus \{u\}$ et $\mathcal{B}' = \mathcal{B} \setminus \{b\}$. Le débit de transmission possible en bits/s entre l'utilisateur u la station de base b est approximé utilisant la capacité du canal de Shannon. Ce dernier est donné par :

$$r_{u,b}^c = w_b^c \log_2(1 + \gamma_{u,b}^c), \quad (4.3)$$

où w_b^c représente la bande passante du canal c alloué par la station de base b .

On suppose que toutes les zBSs sont dotées d'un réseau d'antennes sectorielles. La zone à couvrir est découpée en secteurs distincts, chacun étant couvert par des antennes directionnelles qui concentrent principalement leur puissance dans une direction spécifique. Par conséquent, le gain d'antenne g_b^c de la zBS b est modélisé comme dans (Humadi *et al.*, 2021; Sayehvand et Tabassum, 2020), pour tout $b \in \mathcal{B}_z$, $c \in \mathcal{C}_z$:

$$g_b^c = \begin{cases} g^m, & \text{si } x_{u,b}^c = 1 \\ g^m, & \text{si } x_{u,b}^c = 0 \text{ et } |\theta| \leq \theta^t / 2\pi \\ g^s, & \text{sinon} \end{cases}, \quad (4.4)$$

où g^m et g^s sont le gain du lobe principal et celui de tous les lobes secondaires, respectivement. θ désigne l'angle de la direction visée, qui est modélisé comme une variable aléatoire uniformément répartie dans l'intervalle $[-\pi, \pi]$ et θ^t est la largeur du lobe principal. Par conséquent, le gain d'antenne entre un utilisateur non associé à la zBS b est modélisé en tant que variable aléatoire discrète prenant la forme de $g_b^c = g^m$ (respectivement $g_b^c = g^s$) avec une probabilité $\theta^t / 2\pi$ (respectivement

$1 - \theta^t/2\pi$). Nous supposons que toutes les zBSs sont capables d'ajuster l'orientation des antennes pour atteindre le gain de directivité maximal (Bai et Heath, 2015).

Le canal entre l'utilisateur u et la zBS b à une distance $d_{u,b}$ est modelisé en ligne de visée de propagation LoS comme dans (Sayehvand et Tabassum, 2020; Kokkonen *et al.*, 2017), pour tout $u \in \mathcal{U}$, $b \in \mathcal{B}_z$ et $c \in \mathcal{C}_z$:

$$h_{u,b}^c = g_b^c \left(\frac{\rho}{4\pi f} \right)^2 \frac{\exp(-\xi(f)d_{u,b})}{d_{u,b}^{\alpha_t}}, \quad (4.5)$$

où f est la fréquence de fonctionnement, ρ est la vitesse de la lumière, α_t est l'exposant d'affaiblissement des communications THz et $\xi(f)$ est le coefficient d'absorption moléculaire qui est une fonction de la fréquence de fonctionnement. Il est à noter que dans ce chapitre, les gains du canal sub-6 GHz et mmWave sont similaires aux gains définis dans le deuxième chapitre.

4.4 Formulation du problème

Nous introduisons une nouvelle métrique de performance du système, nommée efficacité énergétique. Cette métrique est dénotée par η et elle est exprimée comme la somme sur tous les utilisateurs, les stations de base et les canaux du rapport de la variable binaire d'association sur le niveau de puissance alloué à l'utilisateur associé. Mathématiquement, l'efficacité énergétique η est exprimée en watt^{-1} et elle est donnée par :

$$\eta = \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} \frac{x_{u,b}^c}{p_{u,b}^c}. \quad (4.6)$$

En utilisant cette métrique, nous formulons un problème d'optimisation qui maximise l'efficacité énergétique du système sous des contraintes de qualité de service définies par chaque utilisateur. Maximiser l'efficacité énergétique revient à optimiser la puissance allouée aux utilisateurs et à éviter d'allouer des niveaux de puissance disparates.

À notre connaissance, ce travail présenté dans ce chapitre est le premier dans la littérature à introduire et à optimiser la métrique η dans un réseau cellulaire à bandes multiples et à proposer un algorithme d'apprentissage par renforcement profond qui résout le problème d'allocation de ressources sous différentes contraintes de qualité de service.

En utilisant les notations précédentes, le problème d'optimisation d'association des utilisateurs aux

stations de base et d'allocation de ressources est formulé mathématiquement comme suit :

$$\underset{\mathbf{X}, \mathbf{P}}{\text{maximiser}} \quad \eta \quad (\text{P5a})$$

$$\text{sous les contraintes: } x_{u,b}^c \in \{0, 1\}, \forall u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}, \quad (\text{P5b})$$

$$r_{u,b}^c x_{u,b}^c \geq r_u^{\min}, \forall u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}, \quad (\text{P5c})$$

$$\gamma_{u,b}^c x_{u,b}^c \geq \gamma_u^{\min}, \forall u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}, \quad (\text{P5d})$$

$$p_{u,b}^c > 0, \forall u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}, \quad (\text{P5e})$$

$$\sum_{u \in \mathcal{U}} \sum_{c \in \mathcal{C}} p_{u,b}^c x_{u,b}^c \leq p_b, \forall b \in \mathcal{B}, \quad (\text{P5f})$$

$$\sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} x_{u,b}^c \leq 1, \forall u \in \mathcal{U}, \quad (\text{P5g})$$

$$\sum_{u \in \mathcal{U}} x_{u,b}^c \leq 1, \forall b \in \mathcal{B}, c \in \mathcal{C}. \quad (\text{P5h})$$

Les contraintes (P5b) garantissent que les variables d'association soient binaires. Les contraintes (P5c) et (P5d) garantissent respectivement le débit r_u^{\min} et le rapport signal sur interférence et bruit γ_u^{\min} de l'utilisateur u . Les contraintes (P5e) et (P5f) assurent que le niveau de puissance alloué à chaque utilisateur soit positif et que la somme des niveaux de puissance alloués par chaque station de base ne dépasse pas sa limite de puissance disponible. Enfin, les contraintes (P5g) et (P5h) garantissent que chaque utilisateur soit associé à une seule station de base utilisant un seul canal.

Le problème (P5) est formulé comme un programme mathématique non linéaire partiellement en nombres entiers et non convexe en raison de la fonction logarithmique et des variables d'association binaires. Nous notons que, les contraintes (P5g) garantissent toujours une solution réalisable pour (P5). Dans ce qui suit, la démonstration de la NP-difficulté est détaillée.

4.5 NP-difficulté

Théorème 4.5.1. *Le problème (P5) est NP-difficile.*

Démonstration : Afin d'étudier la complexité computationnelle du problème (P5), nous utilisons

la technique par réduction à partir du problème d'affectation généralisé (en anglais : *Generalized assignment problem*, GAP) qui est un problème NP-difficile (Fisher *et al.*, 1986).

Étant donné un ensemble d'objets avec des profits et des poids différents, ainsi qu'un ensemble de sacs avec des budgets différents, le problème GAP consiste à assigner chaque objet à exactement un sac sans dépasser son budget. La fonction objectif est de maximiser la somme des profits (Chekuri et Khanna, 2005).

Une instance de (P5) peut être construite par analogie comme suit. Les objets correspondent aux utilisateurs. Les sacs correspondent aux stations de base. Le budget de chaque sac est la puissance de transmission disponible. Le profit de chaque objet est l'inverse du niveau de puissance alloué à chaque utilisateur associé ($1/p_{u,b}^c$, pour tout $u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}$). Enfin, le poids de chaque objet correspond au niveau de puissance alloué. Les niveaux de puissance alloués sont calculés en fonction des exigences minimales de qualité de service des utilisateurs. Sous cette restriction, (P5) consiste à maximiser la somme des profits des utilisateurs associés. Puisque (P5) et GAP sont équivalents et étant donné que GAP est NP-difficile, nous pouvons en conclure que (P5) est également NP-difficile. Cela prouve le théorème 4.5.1. ■

Trouver la solution optimale du problème (P5) dans un temps raisonnable en utilisant la recherche exhaustive est difficile. Par conséquent, des algorithmes efficaces et peu complexes sont proposés dans la suite.

4.6 Solutions algorithmiques

4.6.1 Solution itérative (IRAA)

Le problème (P5) est non linéaire et non convexe en raison des contraintes (P5c) et (P5d). Pour le résoudre efficacement, nous divisons (P5) en deux sous-problèmes : (i) association des utilisateurs aux stations de base et allocation des canaux, et (ii) attribution de puissance. Dans la suite, nous détaillons les deux sous-problèmes, puis nous proposons un algorithme itératif qui résout (P5).

Association des utilisateurs aux stations de base et allocation des canaux

Étant donné une solution d'allocation de puissance réalisable qui satisfait les contraintes de puissance, (P5) peut être simplifié en un problème conjoint d'association des utilisateurs aux stations de base et d'allocation des canaux comme suit :

$$\underset{\mathbf{X}}{\text{maximiser}} \quad \eta, \text{ sous les contraintes: (P5b), (P5c), (P5d), (P5g), (P5h).} \quad (\text{P6a})$$

Le problème (P6) est non convexe et non linéaire en raison des contraintes (P5b), (P5c) et (P5d). Pour résoudre (P6), nous utilisons la dualité lagrangienne et la méthode de descente du gradient. Pour avoir un saut de dualité (la différence entre les valeurs prises par le problème primal et le problème dual aux points solutions) nul, nous relâchons $x_{u,b}^c$ en tant que valeur réelle continue dans l'intervalle $[0, 1]$. De plus, nous considérons le pire scénario d'interférence comme dans (Bethanabhotla *et al.*, 2014) en supposant que toutes les stations de base interfèrent entre-elles.

La fonction lagrangienne associée à (P6) est définie comme suit :

$$\begin{aligned} \mathcal{L}(\mathbf{X}, \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = & \eta + \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} \left(\alpha_{u,b}^c (r_{u,b}^c x_{u,b}^c - r_u^{\min}) \right. \\ & \left. + \mu_{u,b}^c (\gamma_{u,b}^c x_{u,b}^c - \gamma_u^{\min}) \right) + \sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} \lambda_b^c \left(1 - \sum_{u \in \mathcal{U}} x_{u,b}^c \right), \end{aligned} \quad (4.7)$$

où $\boldsymbol{\alpha} = [\alpha_{u,b}^c]$, $\boldsymbol{\mu} = [\mu_{u,b}^c]$ et $\boldsymbol{\lambda} = [\lambda_b^c]$ sont les multiplicateurs de Lagrange des contraintes de débit de données et de rapport signal sur interférence et bruit. Le problème dual lagrangien est formulé comme suit :

$$\underset{\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\lambda}}{\text{minimiser}} \quad g(\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\lambda}), \text{ sous les contraintes: } \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\lambda} \geq \mathbf{0}, \quad (\text{P7a})$$

où la fonction duale $g(\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\lambda})$ est donnée par :

$$g(\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = \underset{\mathbf{X}}{\text{maximiser}} \quad \mathcal{L}(\mathbf{X}, \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\lambda}) \quad (\text{P8a})$$

$$\text{sous les contraintes: (P5g),} \quad (\text{P8b})$$

$$0 \leq x_{u,b}^c \leq 1, \forall u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}. \quad (\text{P8c})$$

Le problème lagrangien peut être résolu en le décomposant en deux problèmes, nommés intérieur

et extérieur. Le problème intérieur est défini dans (P8) et le problème extérieur est défini dans (P7). Pour le problème intérieur, la fonction lagrangienne, notée \mathcal{L}^i , peut être réécrite comme suit :

$$\mathcal{L}^i(\mathbf{X}, \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} x_{u,b}^c \left(\frac{1}{p_{u,b}^c} + \alpha_{u,b}^c r_{u,b}^c + \mu_{u,b}^c \gamma_{u,b}^c - \lambda_b^c \right), \quad (4.8)$$

où la solution optimale est, pour tout $u \in \mathcal{U}$:

$$x_{u,b^*}^{c^*} = \begin{cases} 1, & (b^*, c^*) = \arg \max_{(b,c)} \left(\frac{1}{p_{u,b}^c} + \alpha_{u,b}^c r_{u,b}^c + \mu_{u,b}^c \gamma_{u,b}^c - \lambda_b^c \right) \\ 0, & \text{sinon} \end{cases}. \quad (4.9)$$

Finalement, pour résoudre le problème extérieur, nous utilisons la méthode du gradient, où les variables du gradient sont calculées comme suit :

$$\alpha_{u,b}^c{}^{(t+1)} = \left[\alpha_{u,b}^c{}^{(t)} - \delta_1^{(t)} (r_{u,b}^c x_{u,b}^c - r_u^{min}) \right]^+, \quad (4.10)$$

$$\mu_{u,b}^c{}^{(t+1)} = \left[\mu_{u,b}^c{}^{(t)} - \delta_2^{(t)} (r_{u,b}^c x_{u,b}^c - \gamma_u^{min}) \right]^+, \quad (4.11)$$

$$\lambda_b^c{}^{(t+1)} = \left[\lambda_b^c{}^{(t)} - \delta_3^{(t)} (1 - x_{u,b}^c) \right]^+, \quad (4.12)$$

où $\delta_1^{(t)}$, $\delta_2^{(t)}$ et $\delta_3^{(t)}$ sont les pas d'itération pour t ($t \in 1, 2, \dots, t_{max}$), où t_{max} est le nombre maximal d'itérations. Notez que les δ doivent satisfaire les conditions suivantes : $\sum_{t=1}^{t_{max}} \delta_j^{(t)} = \infty$ and $\lim_{t \rightarrow \infty} \delta_j^{(t)} = 0, \forall j \in \{1, 2, 3\}$ (Zhang *et al.*, 2015).

Allocation de puissance

Le problème d'optimisation de puissance peut être formulé comme suit :

$$\underset{\mathbf{P}}{\text{maximiser}} \eta, \text{ sous les contraintes: (P5c), (P5d), (P5e), (P5f)}. \quad (\text{P9a})$$

Le problème (P9) est un problème non convexe en raison de la fonction logarithmique présente dans les contraintes (P5c). La solution optimale de (P9), notée $\mathbf{P}^* = [p_{u,b}^{c,*}]$, est calculée comme

suit, pour tout $u \in \mathcal{U}, b \in \mathcal{B}, c \in \mathcal{C}$:

$$p_{u,b}^{c*} = \mathbf{max}(p_{u,b}^{c(P5c)}, p_{u,b}^{c(P5d)}) + p_{min}, \quad (4.13)$$

où p_{min} est une constante déterminée par la consommation de l'amplificateur de puissance et du circuit (Li *et al.*, 2014). $p_{u,b}^{c(P5c)}$ et $p_{u,b}^{c(P5d)}$ sont respectivement les niveaux de puissance qui respectent les contraintes (P5c) et (P5d), et qui sont donnés par :

$$p_{u,b}^{c(P5c)} = \left(2 \frac{\gamma_u^{min}}{w_b^c} - 1\right) \frac{w_b^c \sigma^2 + i_{u,b}^c}{h_{u,b}^c}, \quad (4.14)$$

$$p_{u,b}^{c(P5d)} = \frac{\gamma_u^{min}(w_b^c \sigma^2 + i_{u,b}^c)}{h_{u,b}^c}. \quad (4.15)$$

Algorithme itératif (IRAA)

Le pseudo-code de l'algorithme itératif proposé, dénoté par IRAA, est détaillé dans l'algorithme 8. IRAA résout d'une manière itérative le problème intérieur et le problème extérieur jusqu'à atteindre la convergence de l'efficacité énergétique. IRAA commence par trouver une solution réalisable de (P5) et initialiser la matrice d'association et la matrice de puissance. Cette solution est obtenue en utilisant l'algorithme glouton proposé dans la sous-section 4.6.2. Ensuite, la boucle répéter itère jusqu'à ce que la convergence de η soit atteinte ou que le nombre maximal d'itérations y_{max} soit atteint. IRAA résout le problème d'association des utilisateurs (P8) selon l'équation (4.9) et met à jour les multiplicateurs de Lagrange du problème (P7). Enfin, le niveau de puissance alloué est calculé selon l'équation (4.13). La complexité computationnelle de IRAA dans le pire des cas est $\mathcal{O}(y_{max} U^2 B_z^2 C_z)$ en supposant que $B_z > B_m > B_s$ et $C_s = C_m = C_z$.

Algorithme 8 Pseudo-code de l'algorithme IRAA

- 1: Initialiser $\mathbf{X}, \mathbf{P}, \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\lambda}$ et y .
 - 2: **répéter**
 - 3: Résoudre (P8) selon (4.9).
 - 4: Résoudre (P7) et mettre à jour les multiplicateurs de Lagrange selon (4.10), (4.11) et (4.12).
 - 5: Résoudre (P9) selon (4.13).
 - 6: $y \leftarrow y + 1$
 - 7: **jusqu'à** la converge de η ou $y = y_{max}$
 - 8: **Retourner** η
-

4.6.2 Algorithmes gloutons (CGRA et DGRA)

Dans ce qui suit, nous détaillons les algorithmes gloutons proposés pour résoudre (P5). L'idée principale de l'algorithme centralisé, appelé CGRA, est d'itérer sur l'ensemble de stations de base et de canaux. À chaque itération, CGRA commence par associer les utilisateurs ayant les meilleurs gains du canal afin d'optimiser les niveaux de puissance alloués. Si toutes les contraintes du (P5) sont respectées, CGRA met à jour la matrice d'association et la matrice de puissance. CGRA s'arrête lorsque le nombre maximal d'itérations z_{max} est atteint ou qu'il n'y a plus de canaux disponibles. Le pseudo-code de CGRA est donné dans l'algorithme 9. Sa complexité dans le pire des cas est $\mathcal{O}(z_{max}B_z^2C_zU^3 \log U)$ avec $B_z > B_m > B_s$ et $C_s = C_m = C_z$.

Algorithme 9 Pseudo-code de l'algorithme CGRA

```

1:  $\mathbf{X} \leftarrow \mathbf{0}$ ,  $\eta \leftarrow 0$ ,  $z \leftarrow 0$ .
2: Tant que  $z < z_{max}$  et  $|\mathbf{X}| < B_s C_s + B_m C_m + B_z C_z$  faire
3:   Pour chaque canal  $c \in \mathcal{C}$  disponible de la station  $b \in \mathcal{B}$  faire
4:     Trier les utilisateurs non associés en fonction de leurs gains du canal.
5:     Calculer  $p_{u,b}^c$  utilisant l'équation (4.13).
6:     Calculer  $\eta$  selon l'équation (4.6).
7:     Si les contraintes (P5c), (P5e) et (P5f) sont respectées alors
8:       Mettre à jour  $\mathbf{X}$ ,  $\mathbf{P}$  et  $\eta$ .
9:        $z \leftarrow z + 1$ .
10:    Fin Si
11:  Fin Pour
12: Fin Tant que
13: Retourner  $\eta$ 

```

L'algorithme glouton distribué, appelé DGRA, utilise l'idée de CGRA d'une manière distribuée. Nous supposons que chaque station de base connaît uniquement les informations sur l'état du canal entre elle-même et tous les utilisateurs. La procédure d'association est décrite comme suit. Tout d'abord, chaque utilisateur diffuse ses exigences de qualité de service aux stations de base. Ensuite, afin de maximiser l'efficacité énergétique du système, l'utilisateur choisit la station de base avec laquelle il a le meilleur gain du canal et envoie une demande d'association. Chaque station de base prend une décision locale. Elle peut accepter ou refuser la demande envoyée par l'utilisateur, en fonction de la puissance et des canaux disponibles. Il est à noter que DGRA est une solution distribuée évolutive, ce qui la rend efficace pour un réseau à grande échelle (un grand nombre d'utilisateurs et de stations de base).

4.6.3 Solution d'apprentissage par renforcement profond (MADRA)

Dans cette section, nous modélisons le problème d'allocation de ressources comme un processus décisionnel de Markov (en anglais : *Markov decision process*), puis nous proposons un algorithme multiagent basé sur l'apprentissage par renforcement profond.

Processus décisionnel de Markov

Un processus décisionnel de Markov est un modèle stochastique permettant de décrire un problème de décision. Chaque agent peut apprendre une décision, c'est-à-dire choisir une action à exécuter dans un état donné, tout en recevant une récompense en retour. Nous modélisons (P5) comme un processus de Markov multiagent défini par $(U, \mathcal{S}, \mathcal{A}, R, \text{Pr}_{ss'})$, où U est le nombre d'agents, \mathcal{S} et \mathcal{A} désignent les ensembles d'états et d'actions du système, R désigne la fonction de récompense et $\text{Pr}_{ss'}$ est la probabilité de transition de l'état s à l'état s' . Il est à noter que cette probabilité est inconnue pour les agents en raison de l'instabilité de l'environnement (Zhao *et al.*, 2019). L'ensemble d'états $\mathcal{S}_u(t)$ de l'agent u à l'itération t est défini comme les valeurs de gain du canal entre l'agent et toutes les stations de base. L'action $a_u(t) \in \mathcal{A}$ de l'utilisateur u consiste à choisir la station de base avec laquelle il souhaite s'associer et à sélectionner le canal de transmission. En conséquence, chaque utilisateur dispose d'un total de $B_s C_s + B_m C_m + B_z C_z$ actions possibles. Après chaque action, l'utilisateur reçoit une récompense. Des collisions se produisent lorsque plusieurs utilisateurs choisissent la même station de base et le même canal. Plusieurs utilisateurs peuvent être en collision s'ils sélectionnent la même station de base et le même canal. Comme dans (Sana *et al.*, 2019), pour éviter les collisions, une récompense négative est attribuée à un utilisateur s'il est impliqué dans une collision et n'est pas sélectionné par la station de base. La récompense de l'utilisateur u , appelée récompense individuelle, dépend uniquement de son score et elle est calculée comme suit :

$$R_u(t) = \begin{cases} \sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} \frac{x_{u,b}^c(t)}{p_{u,b}^c(t)}, & \text{si aucune collision n'a lieu} \\ -1, & \text{sinon} \end{cases} . \quad (4.16)$$

Par la suite, en raison de larges espaces d'états et d'actions, nous utilisons l'apprentissage par renforcement profond, notamment le QL profond, pour estimer les valeurs Q associées à chaque combinaison état-action.

MADRA

Le pseudo-code de l'algorithme d'apprentissage par renforcement profond, nommé MADRA, est décrit dans l'algorithme 10. MADRA est un algorithme distribué multiagent. Les agents sont les utilisateurs. Chaque utilisateur apprend indépendamment la meilleure association. Tout d'abord, les agents initialisent les poids du réseau de neurones avec des valeurs aléatoires. Les stations de base initialisent les niveaux de puissance en utilisant un schéma d'allocation de puissance équitable, c'est-à-dire $p_{u,b}^c = p_b/C_b$. Chaque utilisateur u choisit une action aléatoire $a_u(t) \in \mathcal{A}$ (avec une probabilité d'exploration ϵ), ou l'action ayant la valeur Q la plus élevée en sortie du réseau de neurones. Ensuite, chaque utilisateur envoie une demande d'association à la station de base choisie et sélectionne le canal de transmission. Si deux utilisateurs ou plus choisissent le même canal et la même station de base, la station de base sélectionnera l'utilisateur ayant le gain du canal le plus élevé. Les autres utilisateurs seront considérés comme étant en collision. Chaque station de base associe les utilisateurs d'une manière à ce que la puissance totale allouée ne dépasse pas sa capacité (c'est-à-dire $\sum_{u \in \mathcal{U}} \sum_{c \in \mathcal{C}} x_{u,b}^c p_{u,b}^c \leq p_b$, pour la station de base b). Les stations de base calculent la puissance allouée aux utilisateurs associés selon l'équation (4.13). Chaque utilisateur associé calcule sa récompense. Pour permettre aux agents d'apprendre d'une manière autonome, chaque agent dispose de sa propre mémoire de lecture (en anglais : *replay memory*). À chaque itération, $s_u(t), a_u(t), R_u(t)$ et $s_u(t+1)$ sont enregistrés dans la mémoire de lecture et le modèle d'apprentissage est entraîné en échantillonnant un mini-lot aléatoire à partir de la mémoire. Les poids du réseau de neurones, $\tau_u(t)$ et $\tau_u^-(t)$, sont mis à jour en minimisant la fonction de perte par la descente du gradient. La fonction de perte est donnée par :

$$L(\tau_u) = (R_u(t) + \Lambda Q_u(s_u(t+1), a_u(t+1); \tau_u(t))) - Q_u(s_u(t), a_u(t); \tau_u^-(t))^2, \quad (4.17)$$

où Λ représente le taux d'actualisation.

4.7 Résultats de simulation

Dans cette section, nous présentons les résultats de simulation Monte-Carlo pour évaluer les performances des algorithmes proposés. Nous considérons une zone géographique sous forme d'un carré de 300 mètres de côté. Différents scénarios sont pris en compte dans les simulations, où les utilisateurs sont générés d'une manière uniforme et aléatoire. Dans le premier scénario, les nombres de stations

Algorithme 10 Pseudo-code de l'algorithme MADRA

```
1: Pour chaque utilisateur  $u \in \mathcal{U}$  faire
2:   Initialiser les poids du réseau de neurones.
3:   Initialiser  $p_{u,b}^c$ .
4: Fin Pour
5: Pour chaque épisode  $e$  faire
6:   Observer l'état initial  $s(0)$ .
7:   Pour chaque itération  $t$  faire
8:     Pour chaque utilisateur  $u \in \mathcal{U}$  faire
9:       Choisir une action  $a_u(t)$ .
10:    Fin Pour
11:    Pour chaque station de base  $b \in \mathcal{B}$  faire
12:      Choisir les utilisateurs à associer.
13:      Mettre à jour  $p_{u,b}^c$  selon (4.13).
14:    Fin Pour
15:    Pour chaque utilisateur  $u \in \mathcal{U}$  faire
16:      Calculer la récompense selon (4.18) et observer l'état suivant  $s_u(t+1)$ .
17:      Enregistrer la transition  $(s_u(t), a_u(t), R_u(t), s_u(t+1))$  dans la mémoire de relecture.
18:      Échantillonner un mini-lot aléatoire de transitions.
19:      Effectuer une descente de gradient pour minimiser la perte et mettre à jour les poids du
      réseau utilisant (4.17).
20:    Fin Pour
21:    Observer l'état  $s(t+1)$ .
22:  Fin Pour
23:  Retourner  $\eta$ 
24: Fin Pour
```

de base sont $B_s = 1$, $B_m = 2$, $B_z = 4$ et dans le deuxième scénario sont $B_s = 1$, $B_m = 4$, $B_z = 9$. Les positions des stations de base sont présentées dans la figure 4.2. Dans le troisième et le quatrième scénarios, les stations de base sont modélisées par un processus ponctuel homogène de Poisson (en anglais : *homogeneous Poisson point process*). Les nombres de stations sont indépendants et les positions sont générées d'une manière aléatoire dans la zone. Les densités des sBSs, mBSs et zBSs en station/m² sont $\lambda_s = 1 \cdot 10^{-5}$, $\lambda_m = 5 \cdot 10^{-5}$, $\lambda_z = 9 \cdot 10^{-5}$ dans le troisième scénario et $\lambda_s = 1 \cdot 10^{-5}$, $\lambda_m = 10 \cdot 10^{-5}$, $\lambda_z = 20 \cdot 10^{-5}$ dans le quatrième scénario. Enfin, le tableau 4.1 liste les autres paramètres de simulation.

Les figures 4.3, 4.4, 4.5, 4.6, 4.7 et 4.8 montrent l'impact des paramètres du système sur les performances de CGRA en termes d'efficacité énergétique, du nombre d'utilisateurs associés et de puissance utilisée. Les figures 4.3, 4.5 et 4.7 comparent les différents scénarios. Nous notons qu'augmenter le nombre de stations de base déployées dans la même zone (de $B_s = 1$, $B_m = 2$, $B_z = 4$ à $B_s = 1$, $B_m = 4$, $B_z = 9$) augmente l'efficacité énergétique et le nombre d'utilisateurs associés puisque plus

TABLE 4.1 – Notations et paramètres de simulation.

Notations	Paramètres	Valeurs
$p_b, \forall b \in \mathcal{B}_s$	Puissance maximale de la sBS	20 dBm
$p_b, \forall b \in \mathcal{B}_m$	Puissance maximale de la mBS	25 dBm
$p_b, \forall b \in \mathcal{B}_z$	Puissance maximale de la zBS	30 dBm
$w_b^c, \forall b \in \mathcal{B}_s, c \in \mathcal{C}_s$	Bande passante sub-6 GHz	20 MHz
$w_b^c, \forall b \in \mathcal{B}_m, c \in \mathcal{C}_m$	Bande passante mmWave	500 MHz
$w_b^c, \forall b \in \mathcal{B}_z, c \in \mathcal{C}_z$	Bande passante THz	1 THz
α_z	Exposant d'affaiblissement	2
σ^2	Densité spectrale de puissance du bruit	-174 dBm/Hz
θ^z, g^m, g^s	Paramètres d'antenne THz	10°, 25 dB, -10 dB (Kokkonieni <i>et al.</i> , 2017)
f, ξ	Coefficient d'absorption moléculaire	0.3 THz, 0.06 (Kokkonieni <i>et al.</i> , 2017)
ζ	Taux d'apprentissage	0.05
$C_s = C_m = C_z$	Nombre de canaux disponibles	2
Λ	Taux d'actualisation	0.9
T	Nombre d'itérations	5000
ϵ	Probabilité d'exploration	0.001
p_{min}	Puissance consommée par le circuit	13 dBm (Li <i>et al.</i> , 2014)

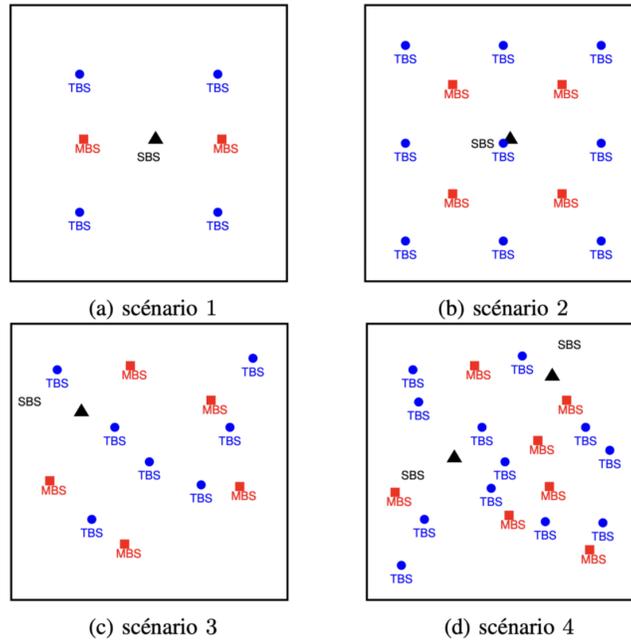


FIGURE 4.2 – Positionnement des stations de base.

de canaux sont disponibles. En outre, l'utilisation du processus ponctuel homogène de Poisson dans les scénarios 3 et 4 n'affecte pas le comportement des courbes de performance. Nous observons un comportement similaire à celui obtenu avec un nombre fixe de stations de base. Les résultats d'efficacité énergétique dans les scénarios 2 et 3 sont remarquablement proches, de même que le nombre

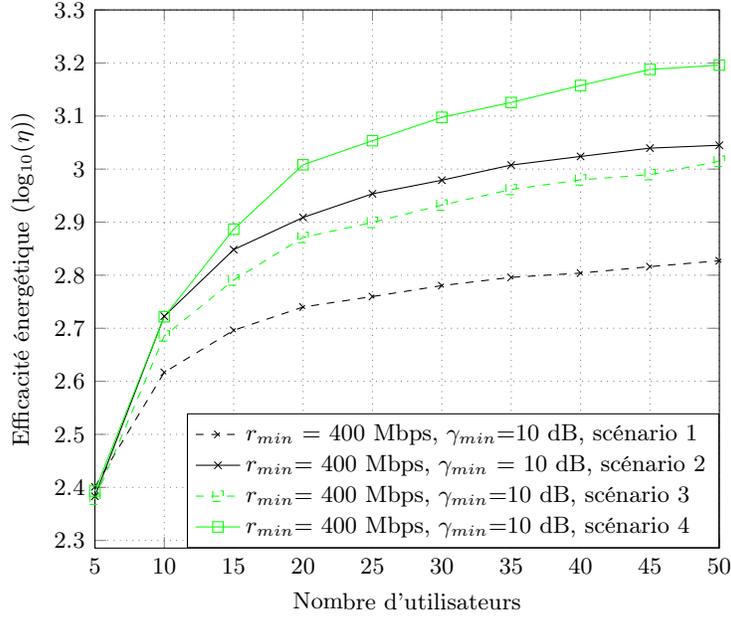


FIGURE 4.3 – Efficacité énergétique ($\log_{10}(\eta)$) de CGRA pour différents scénarios.

d'utilisateurs satisfaits et la puissance consommée. Cela est dû au fait que les nombres moyens des stations de base dans le scénario 3 ($B_s = 1$, $B_m = 5$ et $B_z = 8$, arrondis à l'unité près) sont très proches de ceux du scénario 2. Nous pouvons observer qu'augmenter le nombre d'utilisateurs augmente l'efficacité énergétique et le nombre d'utilisateurs associés. En revanche, comme illustré dans les figures 4.7 et 4.8, la puissance totale allouée augmente lorsque le nombre d'utilisateurs dans le réseau est faible, et diminue lorsque ce nombre est grand. Cette observation s'explique par le fait que, avec un plus grand nombre d'utilisateurs parmi lesquels choisir, CGRA privilégie ceux qui consomment moins de puissance. Ces figures montrent également qu'augmenter le débit minimal (de 400 Mbps à 800 Mbps) diminue l'efficacité énergétique et le nombre d'utilisateurs associés. De plus, augmenter γ_{min} (de 5 dB à 10 dB) diminue l'efficacité énergétique car cela réduit le nombre d'utilisateurs associés tout en maintenant une consommation de puissance du système comparable.

La figure 4.9 illustre la convergence de MADRA. Nous considérons le scénario 2 avec 20 utilisateurs, $r_{min} = 400$ Mbps et $\gamma_{min} = 10$ dB. Nous pouvons observer que MADRA en utilisant la fonction de récompense individuelle (4.16) converge approximativement en environ 4000 itérations dans les trois épisodes qui ont été arbitrairement sélectionnés.

Nous introduisons une autre fonction de récompense, appelée récompense système, qui calcule d'une

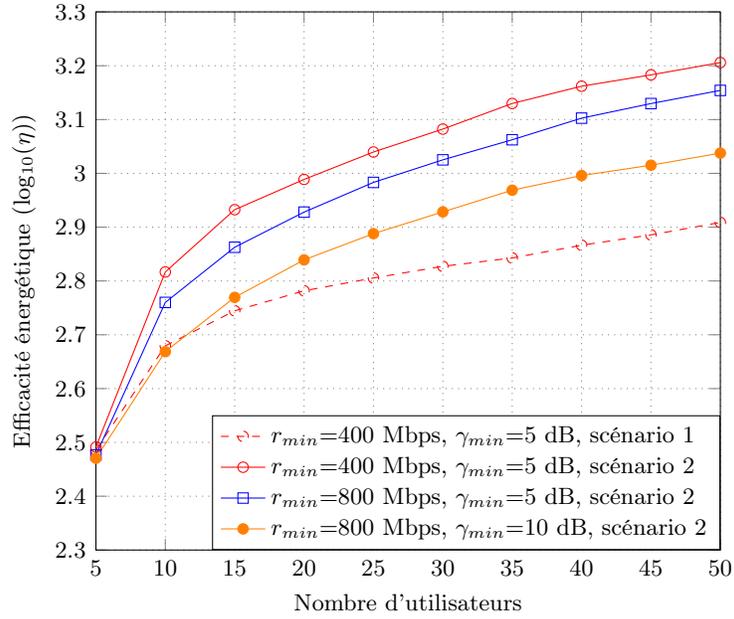


FIGURE 4.4 – Efficacité énergétique ($\log_{10}(\eta)$) de CGRA pour différentes exigences de qualité de service.

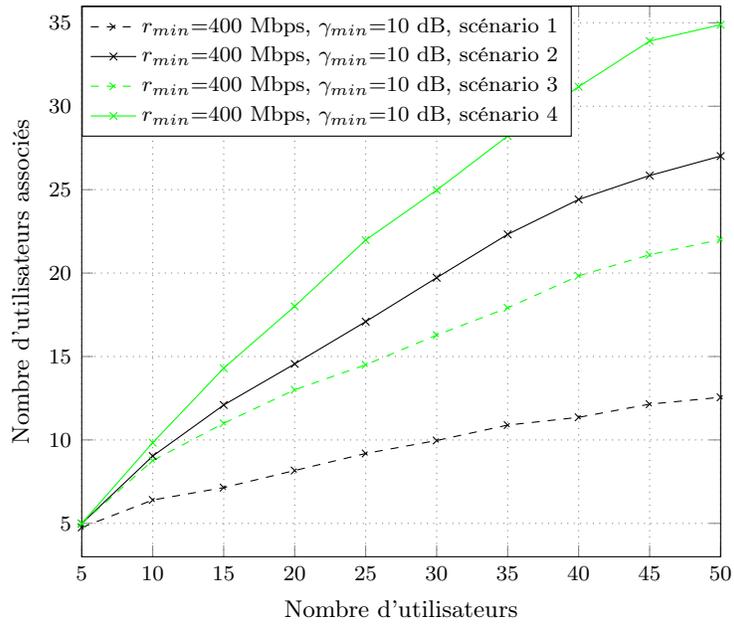


FIGURE 4.5 – Nombre d'utilisateurs associés de CGRA pour différents scénarios.

manière centralisée la récompense $R'(t)$. Cette récompense est la somme des scores de tous les utilisateurs, favorisant ainsi l'apprentissage coopératif. La fonction de récompense système $R'(t)$ est

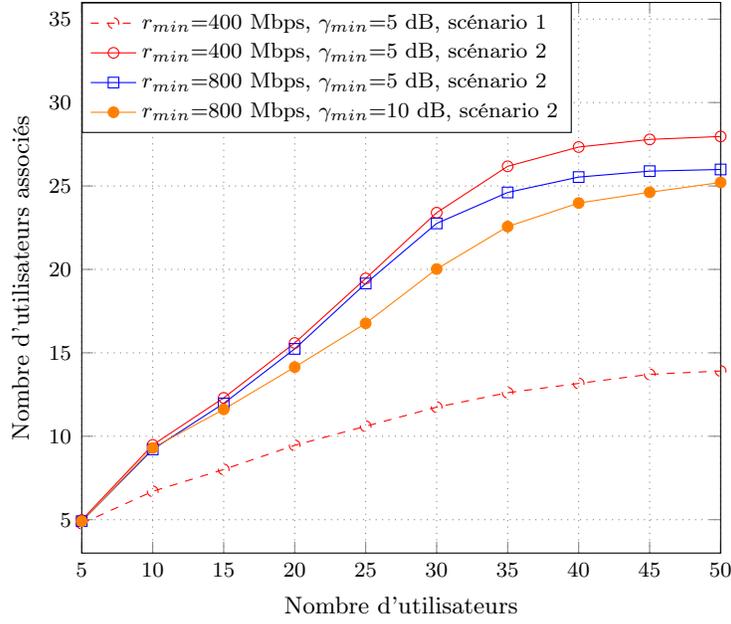


FIGURE 4.6 – Nombre d'utilisateurs associés de CGRA pour différentes exigences de qualité de service.

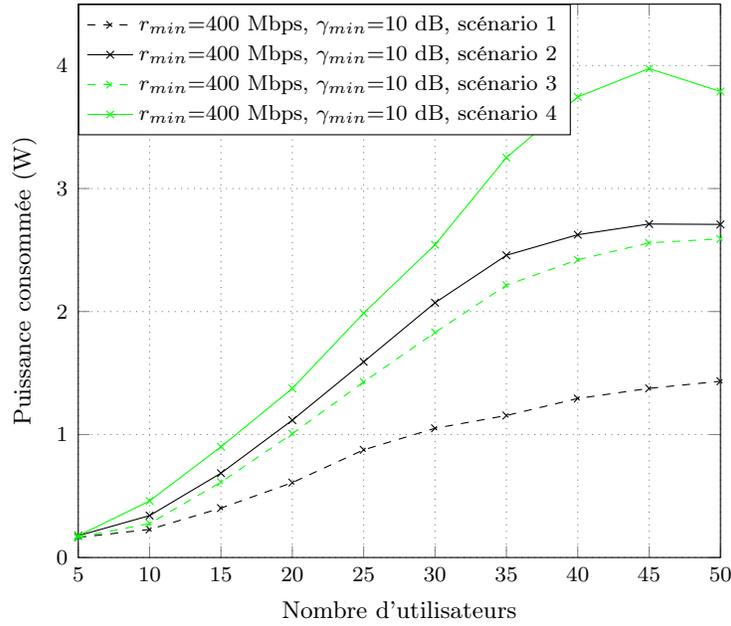


FIGURE 4.7 – Puissance consommée de CGRA pour différents scénarios.

la suivante :

$$R'(t) = \begin{cases} \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \sum_{c \in \mathcal{C}} \frac{x_{u,b}^c(t)}{p_{u,b}^c(t)}, & \text{si aucune collision n'a lieu} \\ -1, & \text{sinon} \end{cases} \quad (4.18)$$

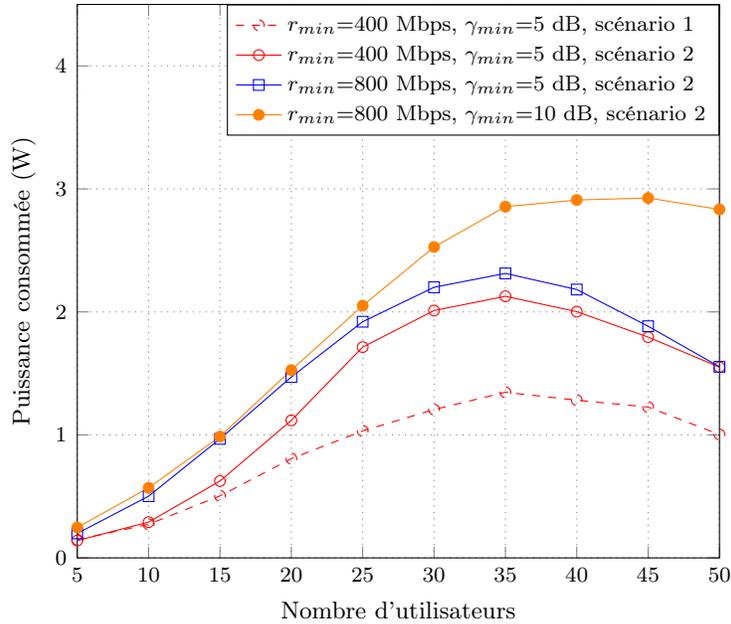


FIGURE 4.8 – Puissance utilisée de CGRA pour différentes exigences de qualité de service.

La figure 4.10 compare les performances de MADRA avec les deux fonctions de récompense individuelle et système. Nous considérons le scénario 2, $U = 20$ utilisateurs, $r_{min} = 400$ Mbps et $\gamma_{min} = 10$ dB. Pour le même nombre d'itérations 5000, l'efficacité énergétique obtenue avec la fonction de récompense individuelle surpasse celle obtenue avec la récompense système. Ceci s'explique par le fait que les résultats de MADRA avec des récompenses calculées d'une manière centralisée n'ont pas convergé après 5000 itérations. Augmenter le nombre d'itérations (de $T = 5000$ à $T = 9000$) améliore l'efficacité énergétique. Nous remarquons que les meilleurs résultats en termes d'efficacité énergétique sont obtenus avec 9000 itérations. Il est clair que la convergence vers une solution optimale ou quasi-optimale exige un plus grand nombre d'itérations lorsqu'une récompense système est considérée. Cette différenciation est due au fait que la fonction de récompense système prend en compte non seulement l'efficacité énergétique de l'agent, mais aussi celle de tous les utilisateurs associés. Il est aussi important de noter qu'un échange d'informations pour calculer la récompense d'une manière centralisée est nécessaire et la quantité de données échangées est proportionnelle au nombre de stations de base et au nombre d'utilisateurs. Étant donné les contraintes de temps d'exécution dans les systèmes réels, dans le reste de ce chapitre, nous fixons le nombre d'itérations à 5000 et nous considérons MADRA avec des récompenses individuelles.

Les figures 4.11, 4.12 et 4.13 illustrent les performances des algorithmes proposés en considérant le

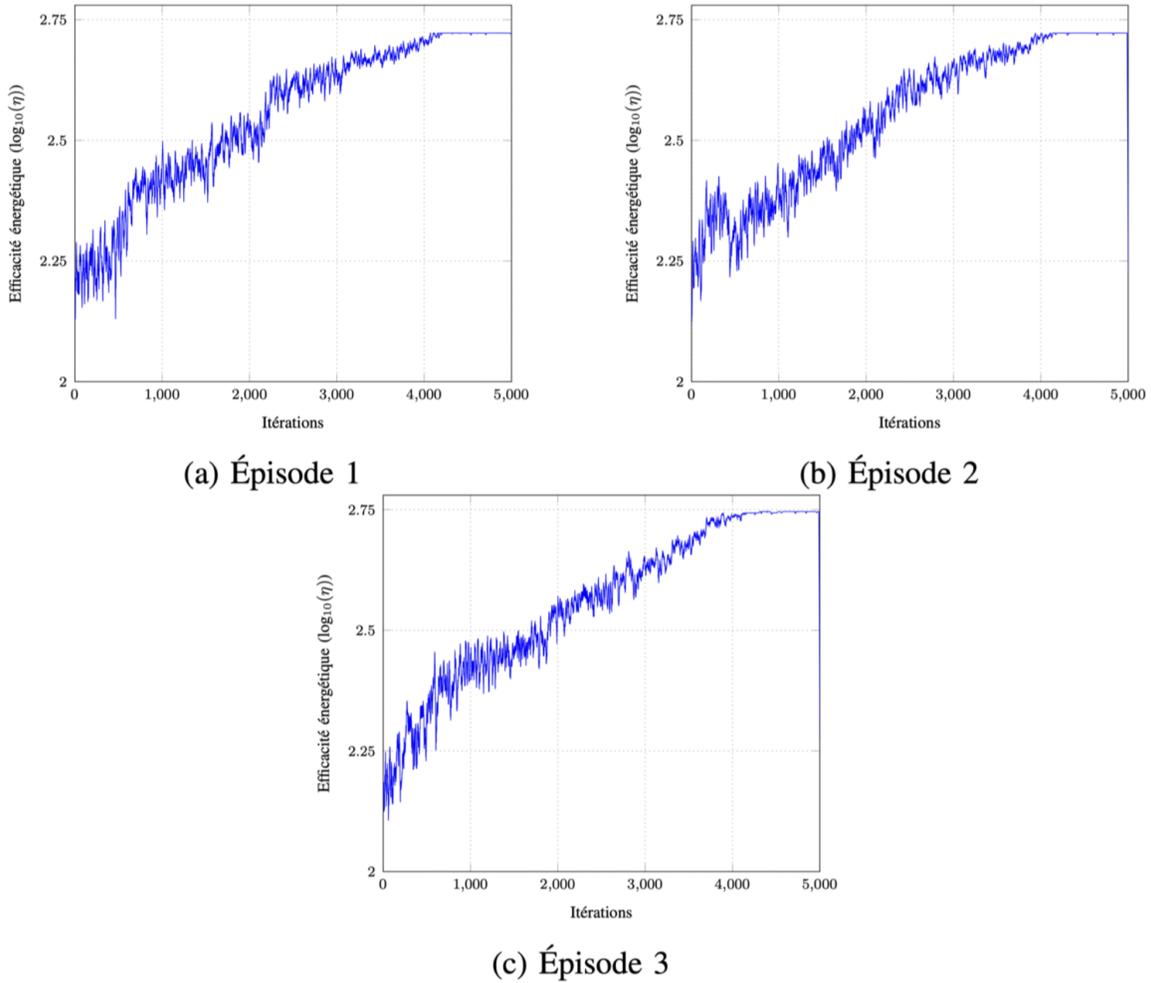


FIGURE 4.9 – Convergence de MADRA – Fonction de récompense individuelle.

scénario 2, $r_{min} = 400$ Mbps et $\gamma_{min} = 10$ dB. En comparant les résultats des algorithmes distribués, nous observons que MADRA surpasse DGRA avec une amélioration d'environ 6% en termes d'efficacité énergétique. Par exemple, avec 40 utilisateurs, MADRA associe un utilisateur de plus en comparaison avec DGRA tout en consommant 56% de moins de puissance. En comparant les résultats des algorithmes centralisés, IRAA surpasse légèrement CGRA en termes d'efficacité énergétique. Cependant, il est à noter que IRAA associe légèrement plus d'utilisateurs dans le cas d'un grand nombre d'utilisateurs et consomme également un peu plus de puissance. Il est clair que les solutions centralisées sont mieux que les solutions distribuées en termes de performances, mais elles sont généralement moins évolutives pour les réseaux à grande échelle. Pour des raisons de comparaison, nous adoptons IRAA afin d'optimiser l'efficacité énergétique de trois réseaux chacun opérant sur une seule bande de fréquences : sub-6 GHz, mmWave ou THz. L'algorithme est nommé S-IRAA

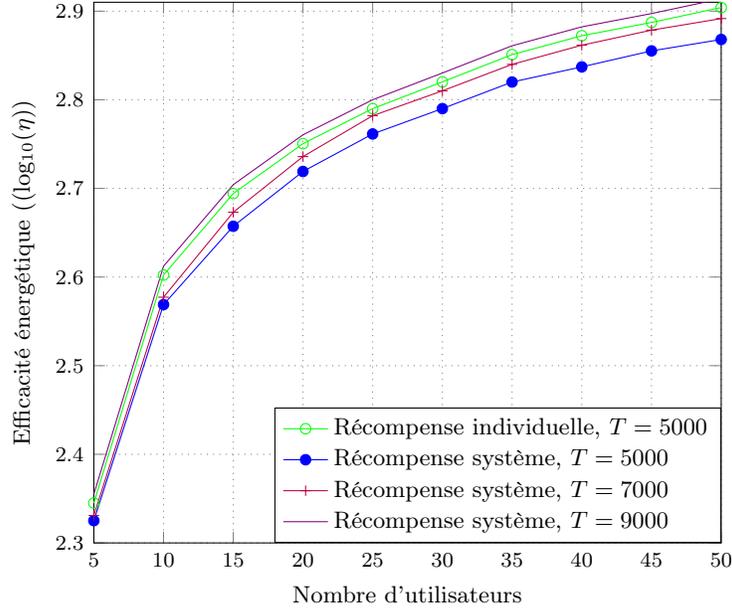


FIGURE 4.10 – Comparaison de l'efficacité énergétique pour MADRA.

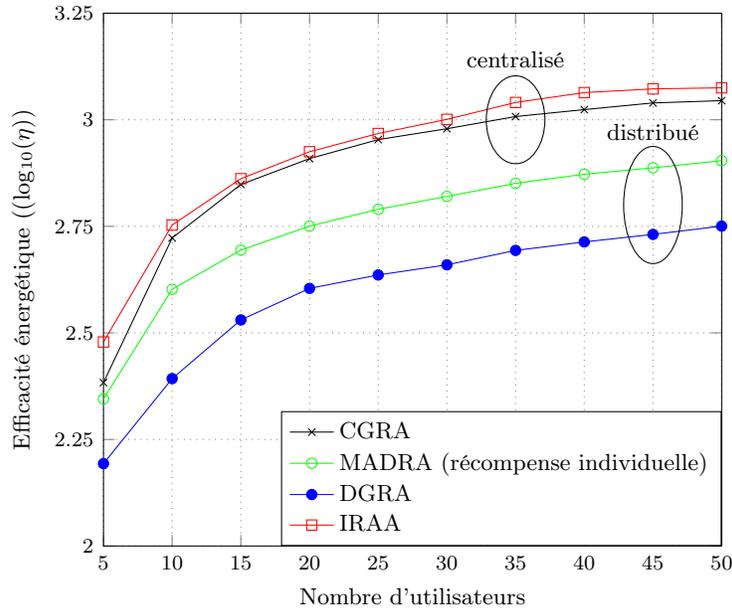


FIGURE 4.11 – Comparaison de l'efficacité énergétique des différents algorithmes.

et il est résumé comme suit. S-IRAA alloue les ressources séparément dans les réseaux sub-6 GHz, mmWave et THz tout en maintenant un nombre égal d'utilisateurs dans chaque réseau. Par exemple, si le nombre total d'utilisateurs est de 30, nous répartissons les utilisateurs en trois groupes de 10. La figure 4.14 illustre les avantages de communication à bandes multiples. Nous considérons le scénario 2, $r_{min} = 400$ Mbps et $\gamma_{min} = 10$ dB. Il est clair que la coexistence de communications à bandes

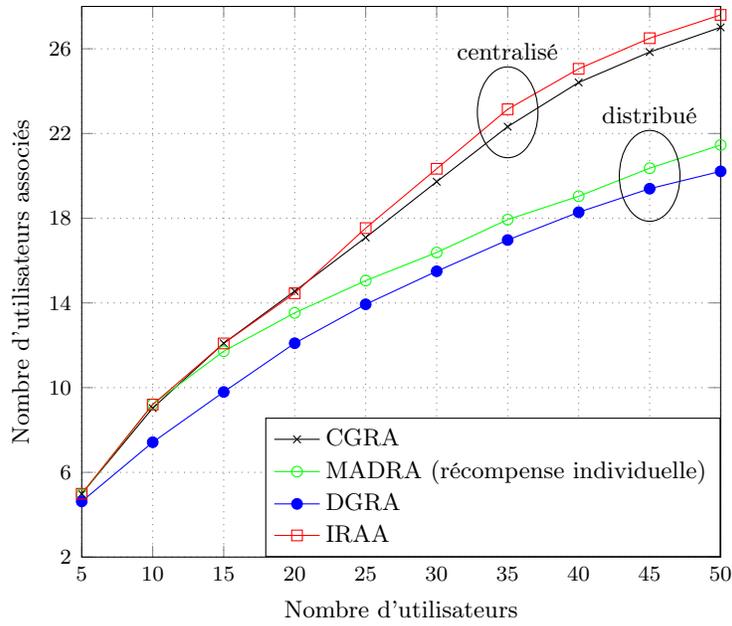


FIGURE 4.12 – Comparaison de nombre d'associations des différents algorithmes.

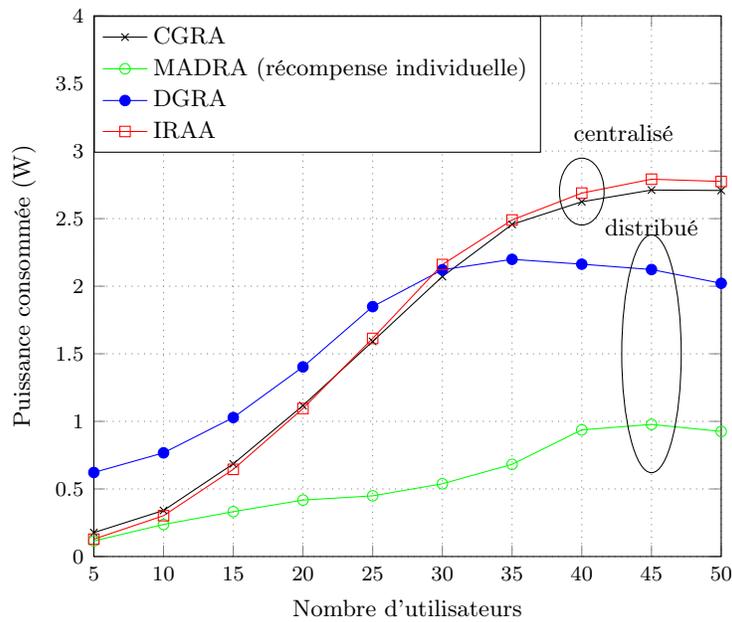


FIGURE 4.13 – Comparaison de la consommation de puissance des différents algorithmes.

multiple augmente l'efficacité énergétique pour différents nombres d'utilisateurs. Par exemple, pour $U = 30$ utilisateurs, la différence d'efficacité énergétique entre IRAA et S-IRAA dépasse 5%.

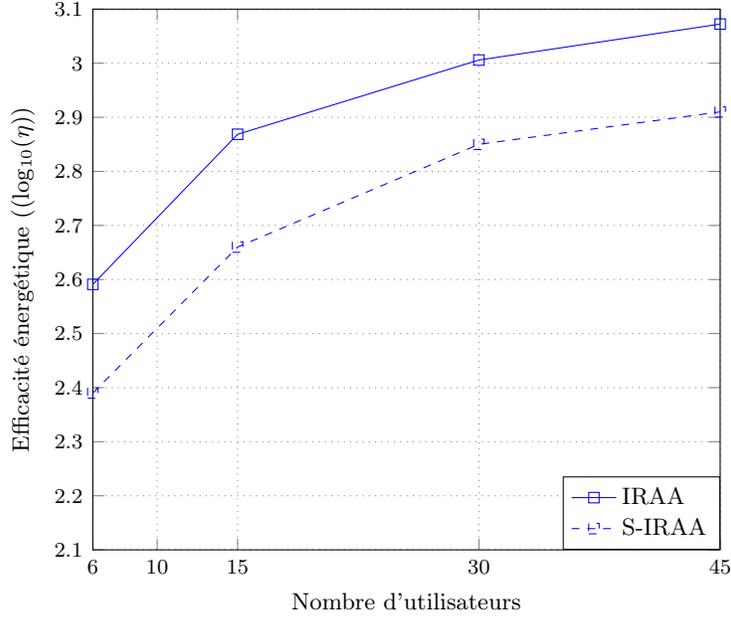


FIGURE 4.14 – Comparaison de l’efficacité énergétique de IRAA et S-IRAA

4.8 Conclusions

Dans le but d’optimiser l’efficacité énergétique du système, nous avons étudié dans ce chapitre le problème d’allocation des ressources dans un réseau cellulaire à bandes multiples. Trois bandes de fréquences ont été utilisées sub-6 GHz, mmWave et THz. Nous avons introduit une nouvelle mesure d’efficacité énergétique qui quantifie le nombre d’utilisateurs satisfaits par unité de puissance. Deux contraintes de qualité de service par utilisateur ont été considérées : un débit et un rapport signal sur interférence et bruit. Tout d’abord, nous avons formulé le problème d’association des utilisateurs aux stations de base, d’allocation des canaux et de puissance comme un programme mathématique non linéaire partiellement en nombres entiers. Puis, nous avons montré que ce problème est NP-difficile et nous avons proposé des algorithmes gloutons peu complexes et un algorithme itératif basé sur la dualité lagrangienne et la méthode de descente du gradient. Nous avons divisé le problème en deux sous-problèmes qui sont résolus itérativement jusqu’à ce que les résultats convergent. Nous avons aussi proposé un autre algorithme multiagent basé l’apprentissage par renforcement profond QL, où les agents sont des utilisateurs qui apprennent la solution d’association et le choix du canal d’une manière indépendante. Enfin, nous avons illustré les performances des algorithmes proposés utilisant des simulations Monte-Carlo et nous avons prouvé l’amélioration de l’efficacité énergétique dans les réseaux à bandes multiples.

CHAPITRE 5

CONCLUSIONS ET TRAVAUX FUTURS

5.1 Conclusions

Dans cette thèse, nous avons proposé des solutions algorithmiques d'allocation des ressources radio (puissance, canaux, etc.) et d'association des utilisateurs aux stations de base pour les futurs réseaux cellulaires à bandes multiples. Trois bandes de fréquences ont été utilisées sub-6 GHz, mmWave et THz. Afin de répondre au mieux aux exigences de futurs systèmes de communication, nous avons optimisé le nombre d'utilisateurs associés, le débit du système et l'efficacité énergétique sous des contraintes de qualité de service et du système. Nos objectifs consistent principalement à formuler mathématiquement les problèmes d'optimisation, à étudier leurs complexités et à proposer des solutions algorithmiques efficaces et peu complexes.

Premièrement, nous avons étudié le problème conjoint d'association des utilisateurs et d'allocation des ressources temporelles dans un réseau hétérogène à bandes multiples. Le modèle du système considéré est composé d'une station de base macro-cellule opérant sur des fréquences sub-6 GHz et des petites stations de base opérant sur des fréquences mmWave et d'autres sur des fréquences sub-6 GHz. Le problème d'interférence a été pris en considération puisque les deux types de stations de base partagent la bande de fréquence sub-6 GHz. Nous avons considéré que chaque utilisateur ne peut être associé qu'à une seule station de base et que plusieurs utilisateurs peuvent être associés à la même station de base. Nous avons formulé le problème mathématiquement sous forme d'un programme linéaire en nombres entiers et qui a pour objectif de maximiser le nombre d'utilisateurs associés et de minimiser le nombre de créneaux temporels alloués. Nous avons démontré que le problème formulé est NP-difficile en utilisant une preuve par réduction polynomiale et nous avons proposé des solutions algorithmiques peu complexes. Le premier est un algorithme glouton centralisé implémenté au niveau d'un point central et le deuxième est un algorithme distribué basé sur l'apprentissage automatique par renforcement. L'objectif de l'apprentissage est de trouver la meilleure association de chaque utilisateur en fonction de son expérience passée, automatiquement et indépendamment des autres. À des fins de comparaison, nous avons aussi proposé un algorithme glouton distribué, où l'association des utilisateurs est basée sur la distance qui les sépare des stations de base. Nous avons utilisé le solveur IBM ILOG CPLEX pour trouver les associations optimales. Les résultats de

simulation Monte-Carlo ont démontré que les performances des algorithmes proposés sont proches des solutions optimales avec une réduction importante de la complexité de calcul.

Deuxièmement, nous avons étudié le problème conjoint d'association des utilisateurs, d'allocation de puissance, d'affectation des sous-canaux et de la technique d'accès multiple dans un réseau cellulaire à bandes multiples. Le modèle du système considéré est composé des stations de base utilisant la bande de fréquences sub-6 GHz et d'autres utilisant la bande de fréquences mmWave. Ainsi, les deux techniques d'accès OMA et MSC-NOMA sont utilisées. Nous avons supposé que les stations de base peuvent choisir, pour chaque sous-canal alloué, entre les deux techniques d'accès. L'objectif du problème formulé, qui non linéaire et non convexe, est d'optimiser le débit de données tout en respectant les exigences d'un débit minimal par utilisateur et les contraintes de puissance. Nous avons montré son appartenance à la classe des problèmes NP-difficile par restriction et nous avons proposé des solutions algorithmiques efficaces. Tout d'abord, nous avons proposé un algorithme glouton de complexité polynomiale ainsi qu'une métaheuristique basée sur la méthode génétique. Puis, nous avons proposé deux autres algorithmes, un centralisé et un autre multiagent, basés sur la technique d'apprentissage automatique par renforcement QL. Nous avons montré, via des simulations Monte-Carlo, l'impact des paramètres du système sur les performances des solutions proposées et les avantages de l'utilisation de deux techniques d'accès dans les réseaux cellulaires à bandes multiples, en termes de débit de données et d'équité entre les utilisateurs.

Troisièmement, nous avons étudié le problème d'optimisation conjointe d'allocation des ressources dans un réseau cellulaire utilisant trois bandes de fréquences sub-6 GHz, mmWave et THz. Nous avons proposé une nouvelle mesure d'efficacité énergétique qui quantifie le nombre d'utilisateurs satisfaits par unité de puissance. Ensuite, nous avons maximisé l'efficacité énergétique du système tout en respectant les contraintes de qualité de service et de puissance. Le problème d'association des utilisateurs aux stations de base, d'allocation des canaux et de puissance a été formulé comme un programme non linéaire partiellement en nombres entiers. Nous avons étudié la complexité du problème et nous avons proposé des solutions algorithmiques efficaces et peu complexes. La première solution divise le problème en deux sous-problèmes que nous avons résolus itérativement jusqu'à atteindre la convergence de l'efficacité énergétique. Le sous-problème d'association des utilisateurs aux stations de base et d'allocation des canaux a été résolu tout en respectant les contraintes du système. Ensuite, nous avons étudié le sous-problème d'allocation de puissance en nous basant

sur les résultats du premier sous-problème pour maximiser l'efficacité énergétique du réseau. La deuxième solution est distribuée multiagent basée sur l'apprentissage par renforcement profond QL. Les agents sont les utilisateurs qui apprennent indépendamment la meilleure association en maximisant leurs scores. Nous avons présenté les résultats de deux fonctions de récompense. La première fonction dépend des scores de tous les utilisateurs, tandis que la deuxième fonction repose uniquement sur le score de l'utilisateur. Les résultats de simulation ont confirmé l'efficacité de la solution d'apprentissage et de la solution itérative, en comparaison avec d'autres solutions basées sur l'approche gloutonne.

Dans cette thèse, nous avons proposé des algorithmes efficaces d'allocation de ressources et d'association des utilisateurs aux stations de base. Nos résultats ont mis en évidence les avantages des réseaux cellulaires utilisant plusieurs bandes de fréquences, notamment en termes de débit de données, de nombre d'utilisateurs satisfaits et d'efficacité énergétique, ainsi que l'impact de l'utilisation de deux techniques d'accès multiples sur les performances des réseaux. Toutes nos solutions algorithmiques sont de complexité raisonnable, les rendant ainsi adaptées à une utilisation future dans les réseaux mobiles cellulaires.

5.2 Travaux futurs

Ci-dessous, nous énumérons quelques pistes de recherche en continuité avec nos travaux réalisés :

- Dans le chapitre 4, nous avons proposé une nouvelle définition de l'efficacité énergétique qui se focalise sur le nombre d'utilisateurs satisfaits par unité de puissance. Changer la fonction objectif à maximiser le nombre d'utilisateurs satisfaits est une piste de recherche importante. Cela nous permet d'évaluer l'impact de l'optimisation de l'efficacité énergétique sur le nombre d'associations.
- Prendre en considération la mobilité des utilisateurs dans le chapitre 4 constitue une autre piste de recherche importante. Nous pouvons nous inspirer des algorithmes d'allocation de ressources et d'association des utilisateurs proposés dans cette thèse.
- Le déploiement des surfaces intelligentes reconfigurables et l'utilisation des drones en tant que stations de base aériennes représentent une piste de recherche prometteuse visant à améliorer la couverture, à réduire les interférences et à minimiser la consommation d'énergie

en adaptant la puissance de transmission nécessaire. Étudier l'optimisation de l'efficacité énergétique dans ce nouveau modèle du système est un défi de taille pour les futurs réseaux cellulaires.

- Étudier les réseaux intégrant des systèmes satellitaires, des réseaux aériens et des communications terrestres en utilisant des surfaces intelligentes reconfigurables est une orientation de recherche. Les défis résident dans l'optimisation de la consommation d'énergie, l'association des utilisateurs et la configuration des surfaces intelligentes.

BIBLIOGRAPHIE

- 3GPP (Octobre 2022). 3GPP Release 18. Consulté le : juin 2024. Récupéré de <https://www.3gpp.org/specifications-technologies/releases/release-18>
- Adam, A. B. M., Wang, Z., Wan, X., Xu, Y. et Duo, B. (2022). Energy-Efficient Power Allocation in Downlink Multi-Cell Multi-Carrier NOMA : Special Deep Neural Network Framework. *IEEE Transactions on Cognitive Communications and Networking*, 8(4), 1770–1783.
- Adedoyin, M. A. et Falowo, O. E. (2020). Combination of Ultra-Dense Networks and Other 5G Enabling Technologies : A Survey. *IEEE Access*, 8, 22893–22932.
- Adeogun, R. et Berardinelli, G. (2023). Distributed Channel Allocation for Mobile 6G Subnetworks via Multi-Agent Deep Q-Learning. Dans *Dans IEEE Wireless Communications and Networking Conference (WCNC)*, 1–6.
- Agarwal, B., Togou, M. A., Ruffini, M. et Muntean, G.-M. (2022). A Comprehensive Survey on Radio Resource Management in 5G HetNets : Current Solutions, Future Trends and Open Issues. *IEEE Communications Surveys & Tutorials*, 24(4), 2495–2534.
- Akdeniz, M. R., Liu, Y., Samimi, M. K., Sun, S., Rangan, S., Rappaport, T. S. et Erkip, E. (2014). Millimeter Wave Channel Modeling and Cellular Capacity Evaluation. *IEEE Journal on Selected Areas in Communications*, 32(6), 1164–1179.
- Alam, S., Islam, S., Khandaker, M. R. A., Khan, R. T., Tariq, F. et Toding, A. (2023). Deep Q-Learning Based Resource Allocation in 6G Interference Systems With Outage Constraints. Dans *IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, 1–6.
- Ali, M. S., Hossain, E., Al-Dweik, A. et Kim, D. I. (2018). Downlink Power Allocation for CoMP-NOMA in Multi-Cell Networks. *IEEE Transactions on Communications*, 66(9), 3982–3998.
- Ali, M. S., Tabassum, H. et Hossain, E. (2016). Dynamic User Clustering and Power Allocation for Uplink and Downlink Non-Orthogonal Multiple Access (NOMA) Systems. *IEEE Access*, 4, 6325–6343.
- Alizadeh, A. et Vu, M. (2019). Load Balancing User Association in Millimeter Wave MIMO Networks. *IEEE Transactions on Wireless Communications*, 18(6), 2932–2945.
- Alizadeh, A. et Vu, M. (2022). Reinforcement Learning for User Association and Handover in mmWave-Enabled Networks. *IEEE Transactions on Wireless Communications*, 21(11), 9712–9728.
- Alsaedi, W. K., Ahmadi, H., Khan, Z. et Grace, D. (2023). Spectrum Options and Allocations for 6G : A Regulatory and Standardization Review. *IEEE Open Journal of the Communications Society*, 4, 1787–1812.
- AMPL (2020). Consulté le : juin 2024. Récupéré de <https://ampl.com/>

- Baghani, M., Parsaeefard, S., Derakhshani, M. et Saad, W. (2019). Dynamic Non-Orthogonal Multiple Access and Orthogonal Multiple Access in 5G Wireless Networks. *IEEE Transactions on Wireless Communications*, 67(9), 6360–6373.
- Bai, T. et Heath, R. W. (2015). Coverage and Rate Analysis for Millimeter Wave Cellular Networks. *IEEE Transactions on Wireless Communications*, 14(2), 1100–1114.
- Bartsiokas, I. A., Gkonis, P. K., Kaklamani, D. I. et Venieris, I. S. (2022). ML-Based Radio Resource Management in 5G and Beyond Networks : A Survey. *IEEE Access*, 10, 83507–83528.
- Bellman, R. (1957). A Markovian Decision Process. *Journal of mathematics and mechanics*, 679–684.
- Bethanabhotla, D., Bursalioglu, O. Y., Papadopoulos, H. C. et Caire, G. (2014). User association and load balancing for cellular massive MIMO. Dans *Information Theory and Applications Workshop (ITA)*, 1–10.
- Chaieb, C., Abdelkefi, F. et Ajib, W. (2020a). Joint User Association and Sub-channel Assignment in Wireless Networks with Heterogeneous Multiple Access and Heterogeneous Base Stations. Dans *IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 1–6.
- Chaieb, C., Abdelkefi, F. et Ajib, W. (2022). On the Sum-rate Maximization in Multi-access and Multi-band Wireless Networks. Dans *IEEE International Conference on Communications (ICC)*, 4902–4907.
- Chaieb, C., Abdelkefi, F. et Ajib, W. (2023a). Deep Reinforcement Learning for Resource Allocation in Multi-Band and Hybrid OMA-NOMA Wireless Networks. *IEEE Transactions on Communications*, 71(1), 187–198.
- Chaieb, C., Ajib, W. et Abdelkefi, F. (2023b). Maximizing the Energy Efficiency in Integrated Sub-6 GHz, mmWave and THz Wireless Networks. Dans *IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 1–6.
- Chaieb, C., Mlika, Z., Abdelkefi, F. et Ajib, W. (2017). On the Optimization of User Association and Resource Allocation in HetNets With mm-Wave Base Stations. Dans *IEEE 28th annual international symposium on personal, indoor, and mobile radio communications (PIMRC)*, 1–5.
- Chaieb, C., Mlika, Z., Abdelkefi, F. et Ajib, W. (2020b). On the Optimization of User Association and Resource Allocation in HetNets With mm-Wave Base Stations. *IEEE Systems Journal*, 14(3), 3957–3967.
- Chaieb, C., Mlika, Z., Abdelkefi, F. et Ajib, W. (2018). Mobility-Aware User Association in HetNets with Millimeter Wave Base Stations. Dans *IEEE 14th International Wireless Communications and Mobile Computing Conference (IWCMC)*, 153–157.
- Chekuri, C. et Khanna, S. (2005). A Polynomial Time Approximation Scheme for the Multiple

- Knapsack Problem. *SIAM Journal on Computing*, 35(3), 713–728.
- Cohen, E. et Megiddo, N. (1993). Maximizing concave functions in fixed dimension. In *Complexity in Numerical Optimization* 74–87. World Scientific.
- Cormen, T., Leiserson, C. et Rivest, R. (2010). *Introduction à l’algorithmique 3ème édition*. Paris, France : Dunod.
- Cui, J., Ding, Z., Fan, P. et Al-Dhahir, N. (2018). Unsupervised Machine Learning-Based User Clustering in Millimeter-Wave-NOMA Systems. *IEEE Transactions on Wireless Communications*, 17(11), 7425–7440.
- Dai, L., Wang, B., Ding, Z., Wang, Z., Chen, S. et Hanzo, L. (2018). A Survey of Non-Orthogonal Multiple Access for 5G. *IEEE Journal on Selected Areas in Communications*, 1–5.
- Ding, Z., Lei, X., Karagiannidis, G. K., Schober, R., Yuan, J. et Bhargava, V. (2017). A Survey on Non-Orthogonal Multiple Access for 5G Networks : Research Challenges and Future Trends. *IEEE Journal on Selected Areas in Communications* , 35(10), 2181–2195.
- Ding, Z., Xu, D., Schober, R. et Poor, H. V. (2022). Hybrid NOMA Offloading in Multi-User MEC Networks. *IEEE Transactions on Wireless Communications*, 21(7), 5377–5391.
- Elshaer, H., Kulkarni, M. N., Boccardi, F., Andrews, J. G. et Dohler, M. (2016). Downlink and Uplink Cell Association with Traditional Macrocells and Millimeter Wave Small Cells. *IEEE Transactions on Wireless Communications*, 15(9), 6244–6258.
- Ericsson (2022). Ericsson Mobility Report : la 5G atteindra 1 milliard d’abonnements en 2022 et 4,4 milliards en 2027. Consulté le : juin 2024. Récupéré de <https://www.ericsson.com/fr/press-releases/2022/6/ericsson-mobility-report-5g-to-top-one-billion-subscriptions-in-2022-and-4.4-billion-in-2027>
- Fan, Y., Zhang, Z. et Li, H. (2019). Message Passing Based Distributed Learning for Joint Resource Allocation in Millimeter Wave Heterogeneous Networks. *IEEE Transactions on Wireless Communications*, 18(5), 2872–2885.
- Fisher, M. L., Jaikumar, R. et Van Wassenhove, L. N. (1986). A multiplier adjustment method for the generalized assignment problem. volume 32 1095–1103. INFORMS.
- Gkonis, P. K. (2022). A Survey on Machine Learning Techniques for Massive MIMO Configurations : Application Areas, Performance Limitations and Future Challenges. *IEEE Access*, 11, 67–88.
- Graesser, L. et Keng, W. L. (2019). *Foundations of deep reinforcement learning : theory and practice in Python*. Addison-Wesley Professional.
- He, C., Hu, Y., Chen, Y. et Zeng, B. (2019). Joint Power Allocation and Channel Assignment for NOMA With Deep Reinforcement Learning. *IEEE Journal on Selected Areas in Communications* , 37(10), 2200–2210.

- He, H., Qin, C., Chen, S., Jiang, X., Yang, J. et Hanzo, L. (2023). Sparse Bandit Learning Based Location Management for Space-Ground Integrated Networks. *IEEE Transactions on Vehicular Technology*, 72(8), 10314–10329.
- Hu, Y., MacKenzie, R. et Hao, M. (2017). Expected Q-learning for Self-Organizing Resource Allocation in LTE-U with Downlink-Uplink Decoupling. Dans *IEEE 23th European Wireless Conference*, 1–6.
- Humadi, K., Trigui, I., Zhu, W.-P. et Ajib, W. (2021). Coverage Analysis of User-Centric Dense Terahertz Networks. *IEEE Communications Letters*, 25(9), 2864–2868.
- IBM (2017). Consulté le : juin 2024. Récupéré de https://www.ibm.com/support/knowledgecenter/SSSA5P_12.8.0/ilog.odms.studio.help/pdf/gscplex.pdf
- Iqbal, A., Tham, M.-L. et Chang, Y. C. (2021). Double Deep Q-Network-Based Energy-Efficient Resource Allocation in Cloud Radio Access Network. *IEEE Access*, 9, 20440–20449.
- Iqbal, M. U., Ansari, E. A., Akhtar, S. et Khan, A. N. (2022). Improving the QoS in 5G HetNets Through Cooperative Q-Learning. *IEEE Access*, 10, 19654–19676.
- Islam, S. R., Avazov, N., Dobre, O. A. et Kwak, K.-S. (2017). Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems : Potentials and Challenges. *IEEE Communications Surveys & Tutorials*, 19(2), 721–742.
- Islam, S. R., Zeng, M., Dobre, O. A. et Kwak, K.-S. (2018). Resource allocation for downlink NOMA systems : Key techniques and open issues. *IEEE Wireless Communications*, 25(2), 40–47.
- Jain, R. K., Chiu, D.-M. W. et Hawe, W. R. (1984). A quantitative measure of fairness and discrimination. *Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA*.
- Jin, K., Cai, X., Du, J., Park, H. et Tang, Z. (2022). Toward energy efficient and balanced user associations and power allocations in multi-connectivity enabled mmWave networks. *IEEE Transactions on Green Communications and Networking*, 6(4), 1917–1931.
- Karp, R. M. (2010). *Reducibility among combinatorial problems*. Springer.
- Khan, M. Q., Gaber, A., Schulz, P. et Fettweis, G. (2023). Machine Learning for Millimeter Wave and Terahertz Beam Management : A Survey and Open Challenges. *IEEE Access*, 11, 11880–11902.
- Kokkonen, J., Lehtomäki, J. et Juntti, M. (2017). Stochastic geometry analysis for mean interference power and outage probability in thz networks. *IEEE Transactions on Wireless Communications*, 16(5), 3017–3028.
- Lamba, A. K., Kumar, R. et Sharma, S. (2020). Joint User Pairing, Subchannel Assignment and Power Allocation in Cooperative Non-Orthogonal Multiple Access Networks. *IEEE Transactions on Vehicular Technology*, 69(10), 11790–11799.

- Li, Y., Sheng, M., Wang, X., Zhang, Y. et Wen, J. (2014). Max–Min Energy-Efficient Power Allocation in Interference-Limited Wireless Networks. *IEEE Transactions on Vehicular Technology*, 64(9), 4321–4326.
- Li, Z., Chen, M., Wang, K., Pan, C., Huang, N. et Hu, Y. (2020). Parallel Deep Reinforcement Learning Based Online User Association Optimization in Heterogeneous Networks. Dans *IEEE International Conference on Communications Workshops (ICC Workshops)*, 1–6.
- Liao, X., Shi, J., Li, Z., Zhang, L. et Xia, B. (2019). A Model-Driven Deep Reinforcement Learning Heuristic Algorithm for Resource Allocation in Ultra-Dense Cellular Networks. *IEEE Transactions on Vehicular Technology*, 69(1), 983–997.
- Lima, B. K. S., Dinis, R., da Costa, D. B., Oliveira, R. et Beko, M. (2022). User Pairing and Power Allocation for UAV-NOMA Systems Based on Multi-Armed Bandit Framework. *IEEE Transactions on Vehicular Technology*, 71(12), 13017–13029.
- Liu, B., Liu, C. et Peng, M. (2020a). Resource Allocation for Energy-Efficient MEC in NOMA-Enabled Massive IoT Networks. *IEEE Journal on Selected Areas in Communications*, 39(4), 1015–1027.
- Liu, B., Liu, C., Peng, M., Liu, Y. et Yan, S. (2020b). Resource Allocation for Non-Orthogonal Multiple Access-Enabled Fog Radio Access Networks. *IEEE Transactions on Wireless Communications*, 19(6), 3867–3878.
- Liu, C.-H. (2019). Coverage-Rate Tradeoff Analysis in mmWave Heterogeneous Cellular Networks. *IEEE Transactions on Communications*, 67(2), 1720–1736.
- Liu, R., Chen, Q., Yu, G. et Li, G. Y. (2019). Joint User Association and Resource Allocation for Multi-Band Millimeter-Wave Heterogeneous Networks. *IEEE Transactions on Communications*, 67(12), 8502–8516.
- Liu, Y., Deng, Y., Nallanathan, A. et Yuan, J. (2023). Machine Learning for 6G Enhanced Ultra-Reliable and Low-Latency Services. *IEEE Wireless Communications*, 30(2), 48–54.
- Ma, T., Xiao, Y., Lei, X., Zhang, L., Niu, Y. et Karagiannidis, G. K. (2023). Reconfigurable Intelligent Surface-Assisted Localization : Technologies, Challenges, and the Road Ahead. *IEEE Open Journal of the Communications Society*, 4, 1430–1451.
- Maier, M. (2023). *6G Vision and Next G Alliance Roadmap*, Dans *6G and Onward to Next G : The Road to the Multiverse*, (p. 77–99).
- Majumdar, S., Trivisonno, R. et Carle, G. (2021). Understanding Exploration and Exploitation of Q-Learning Agents in B5G Network Management. Dans *IEEE Globecom Workshops (GC Wkshps)*, 1–6.
- Malila, B., Falowo, O. et Ventura, N. (2018). Intelligent NLOS Backhaul for 5G Small Cells. *IEEE Communications Letters*, 22(1), 189–192.
- Maniezzo, V., Boschetti, M. A. et Stützle, T. (2021). The generalized assignment problem.

Matheuristics : Algorithms and Implementations, 3–33, Springer.

- Maraqa, O., Rajasekaran, A. S., Al-Ahmadi, S., Yanikomeroğlu, H. et Sait, S. M. (2020). A Survey of Rate-optimal Power Domain NOMA with Enabling Technologies of Future Wireless Networks. *IEEE Communications Surveys & Tutorials*, 22(4), 2192–2235.
- Matskani, E., Sidiropoulos, N. D., Luo, Z.-Q. et Tassiulas, L. (2008). Convex Approximation Techniques for Joint Multi-User Downlink Beamforming and Admission Control. *IEEE Transactions on Wireless Communications*, 7(7), 2682–2693.
- MINOS (2003). Consulté le : juin 2024. Récupéré de <http://web.stanford.edu/group/SOL/guides/minos551.pdf>
- Moltchanov, D., Sopin, E., Begishev, V., Samuylov, A., Koucheryavy, Y. et Samouylov, K. (2022). A Tutorial on Mathematical Modeling of 5G/6G Millimeter Wave and Terahertz Cellular Systems. *IEEE Communications Surveys & Tutorials*, 24(2), 1072–1116.
- Morgado, A. J., Huq, K. M. S., Rodriguez, J., Politis, C. et Gacanin, H. (2017). Hybrid resource allocation for millimeter-wave NOMA. *IEEE Wireless Communications*, 24(5), 23–29.
- Mughees, A., Tahir, M., Sheikh, M. A. et Ahad, A. (2020). Towards Energy Efficient 5G Networks Using Machine Learning : Taxonomy, Research Challenges, and Future Research Directions. *IEEE Access*, 8, 187498–187522.
- Muhammed, A. J., Chen, H., Seid, A. M., Han, Z. et Yu, Q. (2023). Energy-Efficient Resource Allocation for NOMA HetNets in Millimeter Wave Communications. *IEEE Transactions on Wireless Communications*, 22(6), 3790–3804.
- Muhammed, A. J., Ma, Z., Zhang, Z., Fan, P. et Larsson, E. G. (2020). Energy-Efficient Resource Allocation for NOMA Based Small Cell Networks With Wireless Backhails. *IEEE Transactions on Communications*, 68(6), 3766–3781.
- Parsaeefard, S. et Sharafat, A. R. (2012). Robust Worst-Case Interference Control in Underlay Cognitive Radio Networks. *IEEE Transactions on Vehicular Technology*, 61(8), 3731–3745.
- Pei, X., Chen, Y., Wen, M., Yu, H., Panayirci, E. et Poor, H. V. (2022). Next-Generation Multiple Access Based on NOMA With Power Level Modulation. *IEEE Journal on Selected Areas in Communications*, 40(4), 1072–1083.
- Pollakis, E., Cavalcante, R. L. et Stańczak, S. (2012). Base Station Selection for Energy Efficient Network Operation with the Majorization-Minimization Algorithm. Dans *IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 219–223.
- Rappaport, T. S., Murdock, J. N. et Gutierrez, F. (2011). State of the Art in 60-GHz Integrated Circuits and Systems for Wireless Communications. *Proceedings of the IEEE*, 99(8), 1390–1436.
- Rasilainen, K., Phan, T. D., Berg, M., Pärssinen, A. et Soh, P. J. (2023). Hardware Aspects of

- Sub-THz Antennas and Reconfigurable Intelligent Surfaces for 6G Communications. *IEEE Journal on Selected Areas in Communications*, 41(8), 2530–2546.
- Rasti, M., Taskou, S. K., Tabassum, H. et Hossain, E. (2022). Evolution Toward 6G Multi-band Wireless Networks : A Resource Management Perspective. *IEEE Wireless Communications*, 29(4), 118 – 125.
- Russell, S. et Norvig, P. (2010). *Artificial Intelligence : A Modern Approach*. Pearson Education France.
- Saad, W., Bennis, M. et Chen, M. (2019). A Vision of 6G Wireless Systems : Applications, Trends, Technologies, and Open Research Problems. *IEEE Network*, 34(3), 134–142.
- Sana, M., De Domenico, A. et Strinati, E. C. (2019). Multi-Agent Deep Reinforcement Learning Based User Association for Dense mmWave Networks. Dans *IEEE Global Communications Conference (GLOBECOM)*, 1–6.
- Sayehvand, J. et Tabassum, H. (2020). Interference and Coverage Analysis in Coexisting RF and Dense TeraHertz Wireless Networks. *IEEE Wireless Communications Letters*, 9(10), 1738–1742.
- Semiari, O., Saad, W., Bennis, M. et Debbah, M. (2017). Performance Analysis of Integrated Sub-6 GHz-Millimeter Wave Wireless Local Area Networks. Dans *IEEE Global Communications Conference (GLOBECOM)*, 1–7.
- Semiari, O., Saad, W., Bennis, M. et Debbah, M. (2019). Integrated Millimeter Wave and Sub-6 GHz Wireless Networks : A Roadmap for Joint Mobile Broadband and Ultra-Reliable Low-Latency Communications. *IEEE Wireless Communications*, 26(2), 109–115.
- Sharma, N. et Kumar, K. (2022). Energy Efficient Clustering and Resource Allocation Strategy for Ultra-Dense Networks : A Machine Learning Framework. *IEEE Transactions on Network and Service Management*, 20(2), 1884–1897.
- Shi, Y., Lian, L., Shi, Y., Wang, Z., Zhou, Y., Fu, L., Bai, L., Zhang, J. et Zhang, W. (2023). Machine Learning for Large-Scale Optimization in 6G Wireless Networks. *IEEE Communications Surveys & Tutorials, Early access in IEEE Xplore*, 1–1.
- Shimodaira, H., Tran, G. K., Araki, K., Nanba, S., Hayashi, T., Sakaguchi, K. et Konishi, S. (2014). Cell Association Method for Multiband Heterogeneous Networks. Dans *IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*.
- Simon, D. (2013). *Evolutionary optimization algorithms*. John Wiley & Sons.
- Sokun, H. U., Gohary, R. H. et Yanikomeroglu, H. (2015). QoS-Guaranteed User Association in HetNets via Semidefinite Relaxation.
- Sritharan, S., Weligampola, H. et Gacanin, H. (2020). A Study on Deep Learning for Latency Constraint Applications in Beyond 5G Wireless Systems. *IEEE Access*, 8, 218037–218061.

- Sutton, R. S. et Barto, A. G. (2018). *Reinforcement learning : An introduction*. MIT press.
- Tang, F., Zhou, Y. et Kato, N. (2020). Deep Reinforcement Learning for Dynamic Uplink/Downlink Resource Allocation in High Mobility 5G HetNet. *IEEE Journal on Selected Areas in Communications* , 38(12), 2773–2782.
- Tang, L., Tan, Q., Shi, Y., Wang, C. et Chen, Q. (2018). Adaptive Virtual Resource Allocation in 5G Network Slicing Using Constrained Markov Decision Process. *IEEE Access*, 6, 61184–61195.
- Tezergil, B. et Onur, E. (2022). Wireless Backhaul in 5G and Beyond : Issues, Challenges and Opportunities. *IEEE Communications Surveys & Tutorials*, 24(4), 2579–2632.
- Theis, G., Song, Z., Federico, G., Caratelli, D. et Smolders, A. B. (2021). A Design Framework for Beamforming Integrated Circuits Operating at mm-Wave Frequencies. *IEEE Access*, 9, 62232–62240.
- Uusitalo, M. A., Ericson, M., Richerzhagen, B., Soykan, E. U., Rugeland, P., Fettweis, G., Sabella, D., Wikström, G., Boldi, M., Hamon, M.-H., Schotten, H. D., Ziegler, V., Strinati, E. C., Latva-aho, M., Serrano, P., Zou, Y., Carrozzo, G., Martrat, J., Stea, G., Demestichas, P., Pärssinen, A. et Svensson, T. (2021). Hexa-X The European 6G flagship project. Dans *2021 Joint European Conference on Networks and Communications 6G Summit (EuCNC/6G Summit)*, 580–585.
- Uwaechia, A. N. et Mahyuddin, N. M. (2020). A comprehensive survey on millimeter wave communications for fifth-generation wireless networks : Feasibility and challenges. *IEEE Access*, 8, 62367–62414.
- Vaezi, M., Azari, A., Khosravirad, S. R., Shirvanimoghaddam, M., Azari, M. M., Chasaki, D. et Popovski, P. (2022). Cellular, Wide-Area, and Non-Terrestrial IoT : A Survey on 5G Advances and the Road Toward 6G. *IEEE Communications Surveys & Tutorials*, 24(2), 1117–1174.
- Van Hasselt, H., Guez, A. et Silver, D. (2016). Deep Reinforcement Learning with Double Q-learning. Dans *AAAI conference on artificial intelligence*, volume 30.
- Venugopal, K., Valenti, M. C. et Heath, R. W. (2015). Interference in Finite-Sized Highly Dense Millimeter Wave Networks. Dans *Information Theory and Applications Workshop (ITA)*, 175–180.
- Wang, C.-X., You, X., Gao, X., Zhu, X., Li, Z., Zhang, C., Wang, H., Huang, Y., Chen, Y., Haas, H., Thompson, J. S., Larsson, E. G., Renzo, M. D., Tong, W., Zhu, P., Shen, X., Poor, H. V. et Hanzo, L. (2023). On the Road to 6G : Visions, Requirements, Key Technologies and Testbeds. *IEEE Communications Surveys & Tutorials*, 25(2), 905–974.
- Wang, K., Liu, Y., Ding, Z., Nallanathan, A. et Peng, M. (2019). User Association and Power Allocation for Multi-Cell Non-Orthogonal Multiple Access Networks. *IEEE Transactions on Wireless Communications*, 18(11), 5284–5298.

- Wang, W., Wu, J., Li, N. et Song, R. (2021). Joint User Association and Resource Allocation for MmWave and THz Coexistence Networks.
- Wang, X. et Gursoy, M. C. (2021). Coverage in Networks with Hybrid Terahertz, Millimeter Wave, and Microwave Transmissions. *Dans IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)*, 1–4.
- Watkins, C. J. et Dayan, P. (1992). Q-learning. *Technical Note : Q-Learning*, 8(3-4), 279–292.
- Wu, C., Mu, X., Liu, Y., Gu, X. et Wang, X. (2022). Resource Allocation in STAR-RIS-Aided Networks : OMA and NOMA. *IEEE Transactions on Wireless Communications*, 21(9), 7653–7667.
- Xiao, M., Mumtaz, S., Huang, Y., Dai, L., Li, Y., Matthaiou, M., Karagiannidis, G. K., Björnson, E., Yang, K., I, C.-L. et Ghosh, A. (2017). Millimeter Wave Communications for Future Mobile Networks. *IEEE Journal on Selected Areas in Communications* , 35(9), 1909–1935.
- Xing, Y. et Rappaport, T. S. (2021). Millimeter Wave and Terahertz Urban Microcell Propagation Measurements and Models. *IEEE Communications Letters*, 25(12), 3755–3759.
- Xu, B., Chen, Y., El Kashlan, M., Zhang, T. et Wong, K.-K. (2016). User Association in Massive MIMO and mmWave Enabled HetNets Powered by Renewable Energy. *Dans IEEE Wireless Communications and Networking Conference (WCNC)*, 1–6.
- Xu, Y., Gui, G., Gacanin, H. et Adachi, F. (2021). A Survey on Resource Allocation for 5G Heterogeneous Networks : Current Research, Future Trends, and Challenges. *IEEE Communications Surveys & Tutorials*, 23(2), 668–695.
- Xu, Y. et Mao, S. (2017). User association in massive MIMO HetNets. *IEEE Systems Journal* , 11(1), 7–19.
- You, H., Pan, Z., Liu, N. et You, X. (2020). User Clustering Scheme for Downlink Hybrid NOMA Systems Based on Genetic Algorithm. *IEEE Access*, 8, 129461–129468.
- Zalghout, M., Helard, J.-F., Crussiere, M., Abdul-Nabi, S. et Khalil, A. (2017). A Greedy Heuristic Algorithm for Context-aware User Association and Resource Allocation in Heterogeneous Wireless Networks. *Dans IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 1–7.
- Zhang, H., Huang, S., Jiang, C., Long, K., Leung, V. C. et Poor, H. V. (2017). Energy efficient user association and power allocation in millimeter-wave-based ultra dense networks with energy harvesting base stations. *IEEE Journal on Selected Areas in Communications* , 35(9), 1936–1947.
- Zhang, H., Jiang, C., Beaulieu, N. C., Chu, X., Wang, X. et Quek, T. Q. (2015). Resource Allocation for Cognitive Small Cell Networks : A Cooperative Bargaining Game Theoretic Approach. *IEEE Transactions on Wireless Communications*, 14(6), 3481–3493.
- Zhang, Y., Xiong, L. et Yu, J. (2020). Deep Learning Based User Association in Heterogeneous Wireless Networks. *IEEE Access*, 8, 197439–197447.

- Zhao, N., Liang, Y.-C., Niyato, D., Pei, Y., Wu, M. et Jiang, Y. (2019). Deep Reinforcement Learning for User Association and Resource Allocation in Heterogeneous Cellular Networks. *IEEE Transactions on Wireless Communications*, 18(11), 5141–5152.
- Zheng, J., Tang, X., Wei, X., Shen, H. et Zhao, L. (2021). Channel Assignment for Hybrid NOMA Systems with Deep Reinforcement Learning. *IEEE Transactions on Wireless Communications*, 10(7), 1370–1374.
- Zhou, B., Liu, A. et Lau, V. (2019). Successive Localization and Beamforming in 5G mmWave MIMO Communication Systems. *IEEE Transactions on Signal Processing*, 67(6), 1620–1635.
- Zhu, J., Wang, J., Huang, Y., Fang, F., Navaie, K. et Ding, Z. (2020). Resource Allocation for Hybrid NOMA MEC Offloading. *IEEE Transactions on Wireless Communications*, 19(7), 4964–4977.