



HAL
open science

Resource allocation in a cell-less context for 5G wireless networks

Christopher Merlhe

► **To cite this version:**

Christopher Merlhe. Resource allocation in a cell-less context for 5G wireless networks. Networking and Internet Architecture [cs.NI]. Université de Rennes, 2024. English. NNT : 2024URENS042 . tel-04851338

HAL Id: tel-04851338

<https://theses.hal.science/tel-04851338v1>

Submitted on 20 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES

ÉCOLE DOCTORALE N° 601

*Mathématiques, Télécommunications, Informatique, Signal, Systèmes,
Électronique*

Spécialité : *Informatique*

Par

Christopher MERLHE

Resource Allocation in a Cell-less Context for 5G Wireless Networks

Thèse présentée et soutenue à « IMT Atlantique, le 16 Septembre 2024

Unité de recherche : Institut de Recherche en Informatique et Systèmes Aléatoires

Rapporteurs avant soutenance :

Marceau COUPECHOUX Professeur des Universités, Télécom Paris

Thi-Mai-Trang NGUYEN Professeur des Universités, Université Sorbonne Paris-Nord

Composition du Jury :

Président :

Xavier LAGRANGE

Professeur des Universités, IMT Atlantique

Examineurs :

Lila BOUKHATEM

Maître de Conférences, Université de Paris-Sud

Marceau COUPECHOUX

Professeur des Universités, Télécom Paris

Thi-Mai-Trang NGUYEN

Professeur des Universités, Université Sorbonne Paris-Nord

Dir. de thèse :

Bernard COUSIN

Professeur des Universités, Université de Rennes

Co-encadrant de thèse :

Cédric Guéguen

Maître de Conférences, Université de Rennes

À Alain et Françoise,

ACKNOWLEDGEMENT

Cette section est l'occasion de remercier toutes les personnes qui ont contribué à cette thèse, qui m'ont apporté leur aide ou leur soutien.

Je tiens d'abord à remercier mon directeur et mon encadrant de thèse, Bernard Cousin et Cédric Guéguen, sans qui cette thèse n'aurait pas eu lieu. J'ai eu la chance de pouvoir bénéficier de leurs précieux conseils, de leur expérience ainsi que de leur savoir. Cela m'a permis de développer et d'enrichir des compétences qui me seront précieuses dans la suite de ma carrière. Je suis également très reconnaissant de la liberté qu'ils m'ont octroyée, me permettant d'explorer des sujets variés et de travailler avec des personnes provenant d'horizons divers.

Je remercie également les rapporteurs Marceau Coupechoux et Thi-Mai-Trang Nguyen pour leurs attentives relectures du manuscrit et leurs remarques pertinentes. Je remercie Xavier Lagrange et Lila Boukhatem d'avoir accepté de faire partie de mon jury. L'ensemble des échanges avec les membres du jury fut une expérience enrichissante.

Je tiens aussi à exprimer mes remerciements à l'ensemble de l'équipe de recherche ADOPNET pour leur accueil et leurs conseils. En particulier, Xavier Lagrange auprès de qui j'ai beaucoup appris et qui s'est toujours rendu disponible pour répondre à mes questions. Que ce soit en stage ou pendant ma thèse, travailler avec lui a été très enrichissant sur de nombreux aspects. Je suis également reconnaissant envers Loutfi Nuyami qui m'a permis de découvrir et de creuser de nouvelles thématiques de recherche avec lui. D'un point de vue professionnel comme personnel, je suis heureux que nos chemins se soient croisés. Je remercie également Cesar Vargas qui m'a permis de pouvoir bénéficier de son expérience et de revenir sur des choses plus concrètes tout en améliorant ma connaissance des normes 4G et 5G.

Sur un plan plus personnel, je tiens en premier lieu, à remercier ma famille ainsi que ma belle famille pour leur soutien indéfectible. Je tiens aussi à adresser mes plus profonds remerciements à Marine qui, grâce à sa présence au quotidien et ses encouragements, m'a aidé à surmonter les épreuves auxquelles je faisais face durant ces quinze dernières

années.

Je tiens aussi à remercier tous mes amis. En particulier, Malo qui a traversé les mêmes difficultés que moi durant sa thèse et qui sans nul doute, comprenait parfaitement ma situation. Son soutien et ses conseils m'ont beaucoup aidé. Je remercie également toutes les personnes avec qui j'ai noué des liens grâce à cette thèse et qui ont été une réelle bouffée d'oxygène suite à la pandémie. En particulier, Tania, Clément, Matthieu et Philippe.

Je tiens également à souligner que cette thèse est l'aboutissement de longues années d'étude. Cela n'aurait jamais été rendu possible sans l'existence d'un enseignement de qualité, publique et gratuit. Cette école de la République a concrétisé les rêves de l'enfant issu d'un milieu social défavorisé que j'étais, en me permettant d'accéder à un niveau d'étude et à des métiers qui me paraissaient alors inaccessibles. En découle nécessairement un sentiment de redevabilité qui, aujourd'hui et je l'espère dans la suite de ma carrière, me poussera à continuer de contribuer à l'intérêt général en travaillant pour des organismes publics.

Enfin, je dédie cette thèse à mes grands-parents Alain et Françoise, avec qui j'aurai aimé avoir l'occasion de célébrer la fin de ces travaux.

RÉSUMÉ FRANÇAIS

L'augmentation significative du nombre d'utilisateurs ayant des besoins toujours croissants, l'émergence de nouveaux services et de nouvelles applications ont amené les réseaux mobiles à évoluer. En outre, face à des problématiques telles que le réchauffement climatique ou l'inflation du coût de l'énergie, les réseaux mobiles doivent non seulement être en mesure de fournir un bon débit et une bonne qualité de service aux utilisateurs mais également veiller à leur empreinte énergétique.

C'est dans ce contexte que la 5G apporte des évolutions notables comparée à la 4G, qui visent notamment à améliorer l'efficacité spectrale, la zone de couverture, la capacité du système et la latence tout en veillant à l'efficacité énergétique du réseau.

Cette thèse s'inscrit pleinement dans ce cadre et a pour objet l'étude de l'allocation des ressources dans un contexte « Cell-less » dans les réseaux sans fil 5G. Cela sous-entend que le périmètre de cette thèse est restreint à l'étude du réseau d'accès radio et qu'elle s'intéresse tout particulièrement à la nouvelle interface radio appelée New Radio en 5G.

Usuellement, lorsqu'on parle d'allocation des ressources, cela fait directement référence à l'allocation des ressources radio qui est réalisée par l'ordonnanceur au sein d'une station de base. Son rôle est d'attribuer les ressources radio aux utilisateurs afin qu'ils puissent transmettre ou recevoir des données tout en sachant qu'ils ont des qualités de canal radio et des besoins différents. Comme les ressources radio sont en nombre limité à un instant donné, la façon dont l'ordonnanceur alloue les temps de parole aux utilisateurs est donc primordiale pour les performances du réseau et a un impact direct sur la qualité de services des utilisateurs.

Cependant et on le verra dans la suite, cette thèse ne se cantonne pas à l'étude des ordonnanceurs. En effet, l'allocation des ressources comprend également d'autres thématiques. C'est le cas, par exemple, du routage où l'on alloue des flux à des chemins mais également de la gestion des interférences qui peut être vue comme un ordonnancement multi-cellulaire. Un des premiers enjeux de cette thèse est donc d'identifier quel est le type d'allocation de ressources qui sied le mieux au contexte « Cell-less ». Cette approche innovante consiste à envisager le réseau non plus comme une somme de cellules indépendantes mais comme une hyper cellule où l'on peut avoir des prises de décisions de

façon logiquement centralisées. Cela permet notamment de réenvisager l'allocation des ressources classique telle qu'elle est réalisée actuellement, en y apportant des améliorations significatives. Pour atteindre ce premier objectif, il est donc nécessaire d'acquérir une profonde connaissance des différents types d'allocation de ressources et de comprendre les mécanismes qui influent sur les performances de ces solutions.

C'est dans ce contexte, que la première contribution de cette thèse fournit une analyse en profondeur de la diversité multi-utilisateur. La diversité multi-utilisateur peut être définie comme le nombre d'utilisateurs qui peuvent être ordonnancés à un instant donné (c'est-à-dire qui peuvent transmettre ou recevoir des données). La nouveauté de ce travail consiste non plus à voir la diversité multi-utilisateur comme une conséquence de l'allocation des ressources, mais comme un paramètre sur lequel il est possible d'influer et donc de traiter cela comme une problématique à part entière. En effet, la diminution volontaire de l'utilisation de la diversité multi-utilisateur permet de compresser le temps de transmission d'un utilisateur et donc de lui faire économiser de l'énergie. À l'inverse, l'augmentation de l'utilisation de la diversité multi-utilisateur permet d'accroître l'efficacité spectrale et donc de consommer moins de bande passante, ce qui augmente la capacité du système. Cette contribution fournit également des éléments sur la façon de concevoir des solutions d'allocation de ressources et d'évaluer ces solutions via des simulations numériques. Par exemple, pour simuler des profils de trafics d'utilisateurs, le 3GPP recommande principalement deux types de sources : le full buffer ou le non full buffer. Derrière ces aspects, se cache l'utilisation de la diversité multi-utilisateur qui en sera faite par les solutions d'allocation des ressources. Suivant le profil utilisé, les performances de ces solutions en sortie de simulation peuvent fortement varier voire être antagonistes, amenant à des résultats trompeurs aux regards des vraies performances de ces solutions.

La deuxième contribution de cette thèse est une nouvelle solution d'ordonnancement. Cette solution est multi-objectif. Quand de nombreuses ressources radios sont disponibles et que le réseau arrive facilement à assurer une bonne qualité de service aux utilisateurs, la solution proposée réduit volontairement son utilisation de la diversité multi-utilisateurs, ce qui permet d'économiser l'énergie des terminaux mobiles. Lorsque le nombre de ressources radios devient plus restreint, par exemple, suite à une augmentation de la charge de trafic au sein du réseau, la solution va détecter que les utilisateurs commencent à être insatisfaits grâce à une métrique de qualité de service. Elle va alors dynamiquement changer de priorité en se concentrant sur l'optimisation de l'efficacité spectrale, pour réduire le nombre de ressources radios consommées. L'évaluation de performance montre

que cette solution s'approche à la fois des performances des ordonnanceurs spécialisés dans l'économie d'énergie mais également de ceux spécialisés dans la qualité de service et l'équité, permettant de tirer profit du meilleur de ces deux mondes.

La troisième contribution de cette thèse est une extension de la précédente. Il s'agit d'un méta-ordonnanceur, qui peut se greffer aux ordonnanceurs les plus connus. Il va agir juste avant le processus d'allocation des ressources radios et va restreindre ou augmenter l'utilisation de la diversité multi-utilisateur de l'ordonnanceur. Cette augmentation ou diminution de la diversité multi-utilisateur est réalisée en fonction du nombre de ressources radios disponible et de la charge de trafic de chaque utilisateur. L'évaluation de performance montre que cette solution permet aux ordonnanceurs les plus connus d'économiser de l'énergie à basse charge, tout en conservant leurs propriétés intrinsèques (par ex. optimisation de l'efficacité spectrale, latence, équité).

La quatrième contribution de cette thèse s'intéresse au routage. La première difficulté sur cette thématique est de définir quel est le meilleur chemin. Le calcul du meilleur chemin repose donc sur une métrique qui doit être définie et adaptée à notre contexte. Dans notre cas, nous croyons que le meilleur chemin est celui sur lequel le délai est le plus faible. En s'inspirant de la loi de Little, nous proposons une nouvelle solution de routage qui tient compte de l'état des nœuds (taux d'occupation des buffers) et de l'état des liens (qualité du canal radio), afin de réduire le délai de nos chemins. L'évaluation de performance montre que la solution proposée fournit des performances supérieures à celles de l'état de l'art en termes de débit et de latence avec une signalisation comparable voire plus faible.

Les dernières contributions de cette thèse s'intéressent à la gestion des interférences inter-cellulaires. La plupart des solutions actuelles de gestion des interférences n'ont pas connaissance de l'allocation des ressources qui est réalisée dans les cellules voisines. Cela amène souvent à des solutions sous-optimales où l'on traite un problème multi-cellulaire (c'est-à-dire la gestion des interférences) dans le domaine intra-cellulaire. Par exemple, certaines solutions de l'état de l'art vont diviser leur bande passante afin de ne pas transmettre de données sur la même bande de fréquence que les cellules voisines. Cela permet d'éviter les interférences inter-cellulaires au détriment d'un fort gâchis de bande passante. Fort de ces analyses et des contributions précédentes, le contexte Cell-less semble particulièrement adapté à la gestion des interférences inter-cellulaires.

La cinquième contribution de cette thèse propose une nouvelle solution de gestion des interférences dans un contexte Cell-less. Sur une ressource radio donnée et en fonction de

l'état du canal radio des utilisateurs, la solution proposée va soit éviter les interférences de se produire (en empêchant les cellules voisines d'allouer la même ressource radio à un autre utilisateur), soit laisser les interférences se produire. L'évaluation de performance montre que la solution proposée amène à une meilleure efficacité spectrale et une meilleure capacité du système tout en fournissant un meilleur délai aux utilisateurs, particulièrement ceux en bordures de cellule.

La dernière contribution de cette thèse est une extension de la précédente. La solution proposée se base sur l'utilisation du Joint-Transmission Coordinated MultiPoint. Cela permet à plusieurs stations de base de pouvoir communiquer avec un même utilisateur sur la même ressource radio. Ainsi, les interférences sont évitées. Cependant l'un des principaux inconvénients du Joint Transmission est qu'il est très consommateur de ressources radios. Il est donc nécessaire de l'utiliser avec parcimonie en protégeant les utilisateurs des interférences uniquement lorsque c'est nécessaire. La solution proposée se base sur l'état du canal radio des utilisateurs pour savoir quand il est nécessaire de les protéger des interférences via l'utilisation du Joint-Transmission. L'évaluation de performance montre que la solution proposée fournit une meilleure efficacité spectrale et protège efficacement les utilisateurs en bordure de cellules.

TABLE OF CONTENTS

1	Introduction	19
1.1	Mobile networks and 5G overview	19
1.1.1	Mobile and cellular networks	19
1.1.2	Signal attenuation	22
1.1.3	From 4G to 5G	23
1.2	Presentation and challenges of the radio interface	24
1.2.1	Transmission principle	24
1.2.2	Frequency and time organization	26
1.2.3	Resource allocation	27
1.3	Contributions and thesis outlines	29
1.3.1	Motivations	29
1.3.2	Thesis outline	30
1.3.3	Main contributions	31
1.3.4	List of publications	34
2	Influence of the multi-user diversity on resource allocation solutions	35
2.1	Introduction	35
2.2	System description	39
2.3	Contribution: analysis of the multi-user diversity	42
2.3.1	On system performance	42
2.3.2	On the design of resource allocation solutions	45
2.3.3	On the impact of simulation parameters	47
2.4	Discussion	49
2.5	Conclusion	51
3	Scheduling	53
3.1	Introduction	53
3.2	Related work	56
3.2.1	Round Robin (RR)	56

TABLE OF CONTENTS

3.2.2	Maximum Signal-to-Noise Ratio (MaxSNR)	57
3.2.3	Proportional Fair (PF)	57
3.2.4	Weighted Fair Opportunistic (WFO)	58
3.2.5	Opportunistic Energy Aware (OEA)	59
3.2.6	Dynamic Trade-off (DT)	60
3.3	Scheduler to reduce energy consumption and increase system capacity	60
3.3.1	System description	61
3.3.2	Contribution: Fair, Energy efficient and high system Capacity scheduler (FEC)	66
3.3.3	Performance evaluation	68
3.3.4	Conclusion	81
3.4	Meta scheduler to reduce UE's energy consumption	81
3.4.1	System Description	83
3.4.2	Contribution : Adaptative Multi-User Diversity Meta Scheduler (AMUDS)	84
3.4.3	Performance evaluation	91
3.4.4	Conclusion	97
3.5	Conclusion	97
4	Routing	99
4.1	Introduction	99
4.2	Related work	103
4.2.1	Optimized Link State Routing (OLSR)	104
4.2.2	Adhoc On-Demand Distance Vector (AODV)	105
4.2.3	WCETT-LBA	106
4.2.4	Link State Opportunistic Routing (LSOR)	106
4.2.5	Discussion	106
4.3	Node and link state routing	107
4.3.1	Contribution: Opportunistic Buffer Occupancy Routing (OBOR)	108
4.3.2	Operation modes of OLSR, AODV, LSOR and OBOR solutions	111
4.3.3	Performance evaluation	114
4.3.4	Discussion about the overhead	121
4.4	Conclusion	124

5	Interference management in a cell-less context	127
5.1	Introduction	128
5.2	Related work	131
5.2.1	Reuse 1 (R1)	131
5.2.2	Sliced Bandwidth (SB)	132
5.2.3	Fractional Frequency Reuse	133
5.2.4	Soft Frequency Reuse	133
5.2.5	Hybrid static (HS)	135
5.2.6	Power Level Difference CoMP (PLDCoMP)	135
5.3	System description	136
5.4	Interference management with logically centralized decision making	138
5.4.1	Contribution: Dynamic cell-less Radio Access Network Meta Scheduler (DC-RAN-MS)	138
5.4.2	Performance evaluation	142
5.4.3	Conclusion	150
5.5	Interference management with Coordinated Multi-Point	150
5.5.1	Contribution: Hybrid Joint-Transmission CoMP (H-JT-CoMP)	151
5.5.2	Operation mode	151
5.5.3	Performance evaluation	154
5.5.4	Conclusion	160
5.6	Conclusion	160
6	Conclusion	163
6.1	Context	163
6.2	Summary of the main results	163
6.2.1	Multi-user diversity	164
6.2.2	Scheduling	164
6.2.3	Routing	165
6.2.4	Interference management	165
6.3	Perspectives	166
	Bibliography	169

ACRONYMS

- 1G** 1st Generation. 19
- 2G** 2nd Generation. 19
- 3G** 3rd Generation. 20
- 3GPP** 3rd Generation Partnership. 35, 46, 47, 52, 54, 121, 130
- 4G** 4th Generation. 20, 23–27
- 5G** 5th Generation. 19, 20, 23, 24, 26–30, 39, 45, 54, 55, 99, 130, 163
- 5QI** 5G QoS Identifier. 46
- AMUDS** Adaptative Multi User Diversity meta Scheduler. 32, 53
- AODV** Ad-hoc On-demand Distance Vector. 33
- BBU** Baseband Unit. 129, 137
- C-RAN** Cloud Radio Access Network. 128, 129, 137, 138, 142, 150, 160
- CoMP** Coordinated MultiPoint. 33, 127, 130, 136, 151, 153, 157, 158, 160, 161, 166
- CP** Cyclic Prefix. 26
- CQI** Channel Quality Indicator. 49–51
- CSI** Channel State Information. 33, 35, 40, 49–52, 87, 88, 127, 129, 130, 137, 139, 151, 160
- CSI-RS** Channel State Information-Reference Signal. 49
- DC-RAN-MS** Dynamic Cell-less Radio Access Network Meta- Scheduler. 33, 130, 131, 139, 160, 161
- DCI** Downlink Control Information. 27, 28, 39, 49, 55
- DRX** Discontinuous Reception. 43, 55
- DT** Dynamic Trade-off. 60, 67, 71, 73, 74, 76, 78, 79, 81
- EE** Energy Efficiency. 30, 54

- eMBB** Enhanced Mobile BroadBand. 23, 30, 45
- FEC** Fair Energy efficient scheduler for high system Capacity. 31, 32, 53
- gNB** next Generation Node B. 20–23, 25, 27–30, 33, 35, 36, 38–40, 46–49, 53, 57, 58, 61, 68, 73, 74, 83–85, 87, 91, 92, 94, 95, 97, 127–129, 131, 135–137, 160, 161, 166
- H-JT-CoMP** Hybrid Joint-Transmission Coordinated MultiPoint. 33, 131, 150, 151, 153, 157, 158, 160, 161, 166
- ICI** Inter-Cell-Interference. 21, 22, 29–31, 33, 46, 127–132, 136, 138, 139, 141, 142, 144–146, 149–151, 153–155, 157, 158, 160, 165, 166
- ICT** Information and Communications Technology. 30, 54
- IoT** Internet of Things. 23
- JT-CoMP** Joint Transmission Coordinated MultiPoint. 33, 130, 131, 135, 150, 151, 160, 161, 166
- KPI** Key Performance Indicator. 24, 27, 29, 30, 42, 46, 54, 70, 74, 81, 92, 114, 117, 143, 155, 158
- LSOR** Link State Opportunistic Routing. 33
- LTE** Long Term Evolution Advance. 24–27, 41, 130
- MaxSNR** Maximum Signal-to-Noise Ratio. 36–38, 46–50, 54, 56, 57, 63, 66, 68, 71, 73, 74, 76, 78, 79
- MCS** Modulation and Coding Scheme. 23, 28, 39–41, 84, 138
- mIoT** Massive Internet of Things. 23
- mMTC** Massive Machine Type Communications. 23, 46
- MUD** Multi-User Diversity. 31, 32, 35, 37–39, 42–53, 67, 81, 97
- NR** New Radio. 24, 26–29
- O-RAN** Open Radio Access Network. 128
- OEA** Opportunistic Energy Aware. 56, 59, 64–68, 71, 73, 74, 76, 78, 79, 81
- OFDM** Orthogonal Frequency Division Multiplexing. 24, 26, 39, 83, 136, 137

- OLSR** Optimized Link State Routing. 33
- PDCCH** Physical Downlink Control CHannel. 27, 28, 35, 39, 43, 55
- PDSCH** Physical Downlink Shared CHannel. 27, 28
- PF** Proportional Fair. 36–38, 42, 48, 49, 51, 54, 63
- PMI** Precoding Matrix Indicator. 49
- PRB** Physical Resource Block. 25–29, 33, 35, 36, 39, 40, 50, 53, 55, 61–67, 70, 71, 73, 74, 76, 78, 83, 85–87, 89, 90, 127, 131–133, 135, 137–139, 141–147, 149–151, 153, 155, 157, 161, 164
- QCI** QoS Class Identifier. 46
- QoS** Quality of Service. 22–24, 27, 29–32, 34, 36, 42, 44–47, 51, 53, 60, 62, 66–70, 73, 74, 76, 78, 81, 82, 84, 85, 89, 93–95, 97–99, 106, 108, 128, 132, 142, 143, 147, 149–151, 154, 158, 160, 163, 164
- RA** Random Access. 36
- RAN** Radio Access Network. 30, 54
- RE** Resource Element. 25
- RNTI** Radio Network Temporary Identifier. 28
- RR** Round Robin. 36, 46, 48, 49, 54, 57, 70, 71, 73, 74, 76, 79
- RRH** Remote Radio Head. 129, 130, 137–139, 141, 143–145, 150, 151, 153, 158
- SBS** Sub-Carrier Spacing. 26
- SINR** Signal-to-Interference-and-Noise Ratio. 128, 130, 137, 153, 157, 160
- SNR** Signal-to-Noise Ratio. 36, 38, 40, 57, 67, 68, 74, 83, 84, 137
- SSB** Synchronization Signal Block. 49, 50
- TBS** Transport Block Size. 25
- TS** Time Slot. 86, 139
- UE** User Equipment. 20–25, 27–33, 35–40, 42–49, 51, 53, 55, 62, 63, 65, 66, 70, 83–89, 91–95, 97–99, 128, 130–133, 135–139, 141–147, 149–151, 153–155, 157, 158, 160, 161, 164–167
- uRLLC** Ultra-Reliable Low Latency Communications. 23, 26, 46
- WFO** Weighted Fair Opportunistic. 63, 67, 71, 73, 74, 76, 78, 79, 81

INTRODUCTION

The main purpose of this chapter is to provide the reader with an overview of background, issues and contributions related to this thesis. This chapter begins with an introduction of foundations and key concepts of mobile networks allowing the context of this work to be clearly identified. Then, knowing that the radio resource allocation is the main concern of this thesis, challenges and issues of the radio interface are presented. This chapter ends with the motivations, outlines and the contributions of this thesis.

1.1 Mobile networks and 5G overview

This section firstly describes the history and evolution of cellular and mobile networks. Then, this section provides the challenges and limitations of such concept as well as a description of signal attenuation occurring in wireless environment. This section ends with an introduction on 5G networks.

1.1.1 Mobile and cellular networks

The concept of cellular networks has been firstly defined by K. Bullington¹ and H.J. Shuttle² in the early 50s and 60s, respectively. The main idea was to provide transmitter-receiver installations at fixed locations in order to replace wired phones by wireless cell-phones allowing mobile connectivity.

Mobile operators took up this concept up nearly three decades later with the emergence of analog phones (1G) and digital phones (2G). It was the beginning and foundations of cellular networks as we currently know them. For many years, the main service of these networks was voice calls. In the early 2000s, the emergence of wireless internet

1. Bullington, K., « Frequency Economy in Mobile Radio Bands », *in: Transactions of the IRE Professional Group on Vehicular Communications PGVC-3* (1953), pp. 4–27.

2. Schulte, H. J. and Cornell, W. A., « Multi-area Mobile Telephone System », *in: Transactions of the IRE Professional Group on Vehicular Communications VC-9.1* (1960), pp. 49–53.

(3G) and the appearance, a decade later, of smartphones and mobile broadband (4G) have revolutionized phone usage. Nowadays, the main service is data packets³, requiring cellular networks to evolve.

In cellular networks, base stations are at fixed locations and cover the territory in order to provide any devices with wireless access a variety of data services. The territory in the radio range of a base station constitutes a *cell*. This range depends mainly on the radio environment, on the transmission power and the frequency used. Into a cellular network, a device is connected to only one base station, the one that provides the strongest signal.

In the rest of this manuscript, a wireless device is called a User Equipment (UE) and a base station is called a next Generation Node B (gNB) to fit 5G standard. Figure 1.1 illustrates a cellular network with seven gNBs and UEs, where hexagons represents the range of each gNB (i.e. cells).

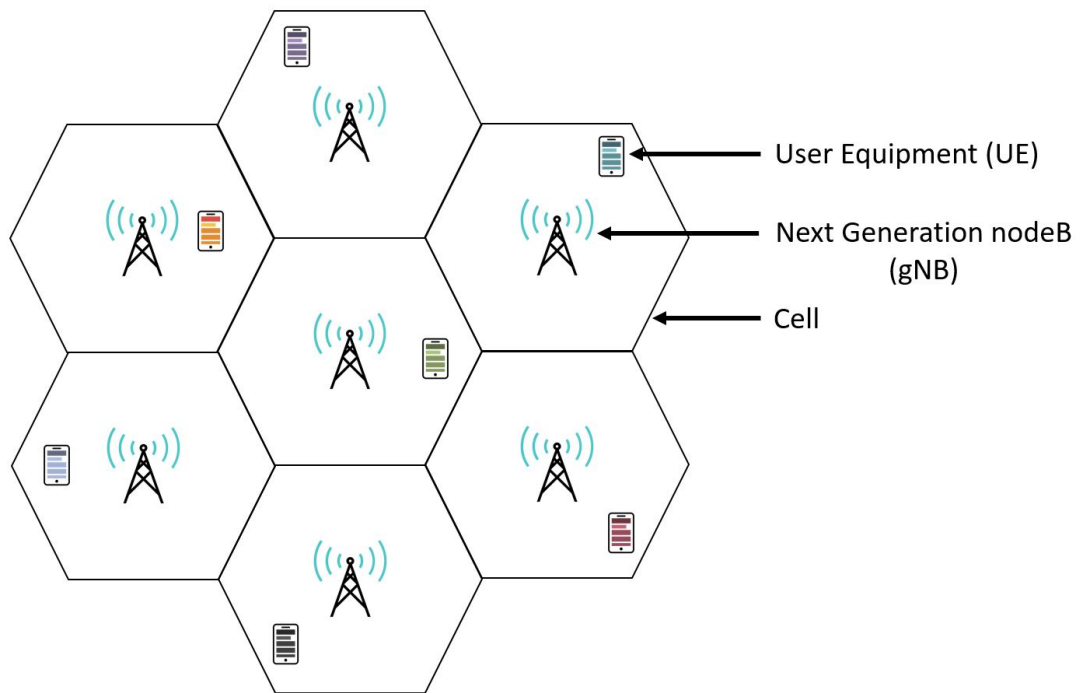


Figure 1.1 – A cellular network.

In order to cope with the increasing number of devices and the variety of services, wireless standards have set up different types of cells in terms of power transmission and frequency used. These parameters vary according to the environment and the number

3. Ericsson, in: *Ericsson Mobility report* (2022).

of devices (e.g. rural/sub-urban/urban/dense-urban) but also on the needs (e.g. outdoor/indoor). This leads cells to have different sizes varying from a range of a few dozen meters to few kilometers. These different cells have various designation to be able to distinguish them (e.g. femto, pico, micro and macro cells) and allow operators to also have a better energy management of their networks⁴.

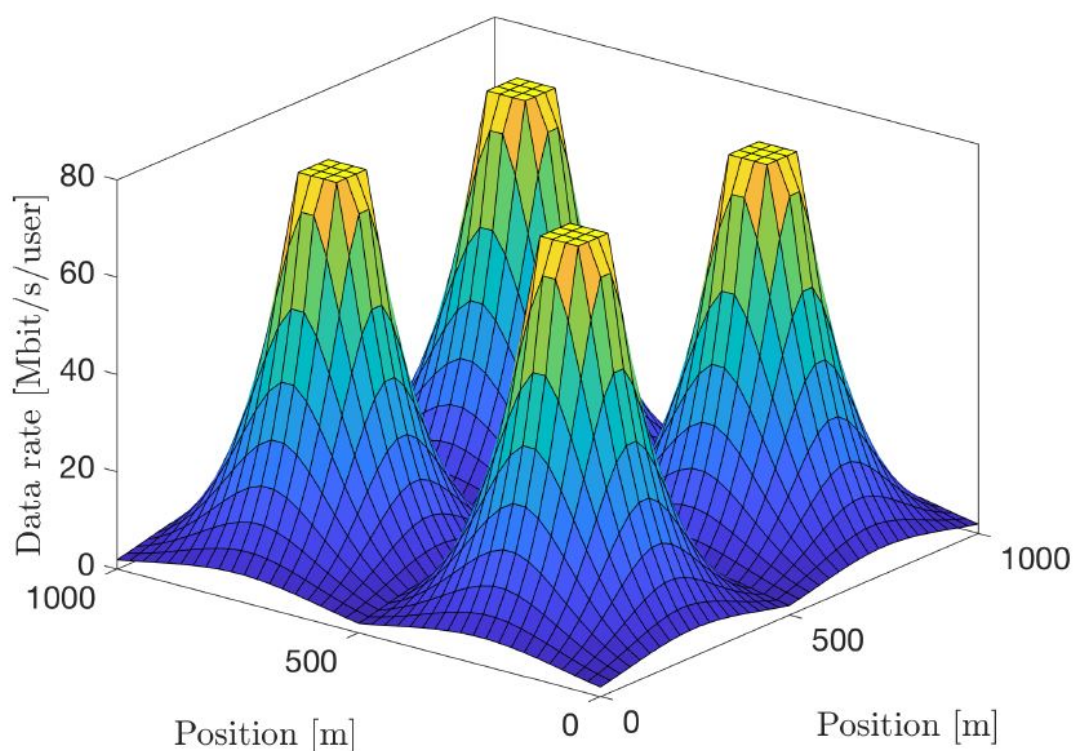


Figure 1.2 – 4 gNBs with a 9dBi fixed-gain antenna (from Emil Björnson’s book⁵).

Inter-Cell-Interference (ICI) occurs when a cell transmits or receives signal at the same time and frequency (i.e. on the same radio resource) than its neighboring cells. The more UEs are close to interfering cells, the more their transmission quality is reduced. Usually, ICI is managed by frequency plans so adjacent cells use different frequency radio resources. This allows to avoid ICI at the expense of a high bandwidth waste. In addition, extending the number of cells of variable sizes leads to overlapping areas or cells included in each

4. 3GPP, « Energy Saving Management (ESM) », in: *3rd Generation Partnership Project (3GPP), TS 32.551 v11.0.0* (2012).

5. Demir, Ö., Björnson, E., and Sanguinetti, L., *Foundations of User-Centric Cell-Free Massive MIMO*, vol. 14, 3-4, Now Publishers, 2021, pp. 164–472.

others. This increases the probability that ICI occurs.

Figure 1.2⁶ illustrates the maximum reachable throughput for any UEs with 4 gNBs. Each gNB has a range of 500 m and transmits with full power. This figure emphasizes that the maximum data rate can be achieved at the center of cells and highly decays with the distance (cf. Section 1.1.2) and interference induced by neighboring cells. Consequently, operators cannot rely on cell densification to ensure broadband services and face the exponential increase of the number of devices in coming years.

Investigating a new approach for mobile network, which differs from the cellular view is thus, a major concern for the future of wireless mobile networks. Based on this analysis, this thesis investigates the benefits of an innovative approach known as *cell-less*. In this context, gNBs cooperate with each other and UEs can dynamically be connected to different gNBs at the same time (unlike in usual cellular networks). This allows to have a dynamical network management by leveraging gNB cooperation to design more efficient resource allocation solutions (particularly in ICI field). Moreover, this thesis aims to combine the benefits of cell-less with an opportunistic approach of resource allocation. Within a cell, various signal attenuation may occur that highly impact the transmission efficiency. Opportunistic solutions take into account the radio channel quality between each couple of UEs and gNB. This allows to reduce these negative impacts to provide a better energy and spectral efficiency, increase the system capacity and the Quality of Service (QoS).

1.1.2 Signal attenuation

A significant signal attenuation observed in wireless networks is multipath fading. Since propagation is not guided, the radio waves emitted by an antenna could propagate in many directions, could encounter different obstacles and some of them will eventually recombine at the receiving antenna. As those waves travel on paths with different lengths, they reach the receiving antenna with a different phase. Depending on how much in phase (or out-of-phase) these waves arrive, constructive or destructive interference occurs. If these waves are in phase, constructive interference produces a strong received signal which may be harnessed using a high order modulation to obtain a high short term throughput. If these waves are close to phase opposition, destructive interference yields poor received signal power and results in low short term throughput. As a consequence of this multipath

6. *Ibid*

fading phenomenon, channel state varies quickly across time. On a longer time scale, the channel state also varies due to path loss and shadowing if UEs are mobile.

In current mobile networks, the Modulation and Coding Scheme (MCS) is adapted to the channel quality. This allows UEs to properly code/decode the data transmitted/received to avoid errors during the transmission and limit retransmission. Consequently, signal attenuation highly impact the MCS and thus, the achievable throughput.

To guarantee a good channel quality is thus, not always possible and the network has to take into account these variations to preserve the users' QoS. In particular, gNBs being masters of the radio access network, the whole issue is to authorize UEs to transmit or receive data at the appropriate moment, depending on their channel quality. This is one of the challenge of the opportunistic resource allocation. This allows to make a good use of the potential gNB's bandwidth and potentially, be able to accept more users in the network.

1.1.3 From 4G to 5G

The ever-increasing number of equipment and needs in terms of throughput and delay constraints, the emergence of Internet of Things (IoT) and the increase of applications (such as gaming, streaming with higher definition video formats or even virtual reality) lead mobile networks to take up a major place in the everyday-life. All these elements have an interdependent relationship: the increase of throughput leads to develop new applications that create needs both for users and new advanced applications, and *vice versa*. This required mobile networks to evolve.

In particular, the transition from 4G to 5G networks has been made in this context. Faced with this multitude of applications with various needs, 5G standard defines three services. The first is Ultra-Reliable Low Latency Communications (uRLLC) that is intended for highly critical applications with the lowest possible latency and high network reliability such as autonomous vehicle, industrial automation or remote medicine. The second one is Massive Internet of Things (mIoT) also known as Massive Machine Type Communications (mMTC). It aims to offer a massive connectivity particularly for smart energy networks, industry or agriculture usage. However, uRLLC and mIoT are mainly for highly specialized, niche services and not for general use cases. For most 5G human users, these services are simply not relevant. This thesis is thus, focused on the third service that is Enhanced Mobile BroadBand (eMBB). This service can be seen as an extension of 4G networks which aims to provide higher data rates, improved latency and

coverage area.

In this context and following Key Performance Indicator (KPI) of 5G, this thesis aims to increase the spectral efficiency to provide a better throughput to UEs, increase the system capacity to accept more UEs in the network while ensuring a good QoS to UEs. Moreover, due to the global warming and the inflation related to energy cost, this thesis is also focused on the energy efficiency by wisely allocating radio resources in order to save energy from UE-side.

In order to make the transition from 4G to 5G easier, the standardization body has chosen to keep the same principles for the radio access network architecture. Nevertheless, the radio interface has evolved.

1.2 Presentation and challenges of the radio interface

The radio interface in 4G-Long Term Evolution Advance (LTE) enabled to reach throughput over 100 Mbit/s, which was impossible with previous technologies. However, constraints inherent to its design in terms of spectral and energy efficiencies have led to improve this interface. In 5G, the radio interface inherited from 4G-LTE is called New Radio (NR). In particular, 5G networks rely on this enhanced radio interface to achieve aforementioned KPIs. As this thesis is focused on the radio part of the network, particularly on resource allocation, this section introduces the 5G-NR radio interface. Short but fundamental descriptions of transmission principle and organization in time and frequency of the 5G-NR frame are provided. This allows to understand the challenge and issues involved in the resource allocation process, described at the end of this section.

1.2.1 Transmission principle

In the manner of 4G-LTE, the transmission on 5G-NR radio interface relies on the Orthogonal Frequency Division Multiplexing (OFDM). The total available bandwidth is divided in frequency bands called sub-carriers. By using orthogonal sub-carriers to each other, the spectral efficiency and thus the global throughput is increased as almost the whole available bandwidth could be used. OFDM therefore, allows to simultaneously perform multiple radio transmissions.

In 4G-LTE, the maximum number of sub-carriers is 1200, against 3300 in 5G-NR

(when carrier bandwidth is 200 MHz and subcarrier spacing is 60 kHz)⁷. This is one of the levers that allows an increase in throughput. The smallest element of resource, called Resource Element (RE), is composed of one symbol on one sub-carrier. The gNB adapts the modulation to the channel quality of each UE. So the number of bits per symbol varies (e.g. 2 bits in QAM, 4 bits in 16-QAM and so on.). However, it is not wise to allocate radio resources sub-carrier by sub-carrier. For this purpose, the concept of Physical Resource Block (PRB) is defined. This block corresponds to a block of 12 sub-carriers in the frequency domain. In the time domain, a PRB (unlike 4G-LTE) is defined on a variable number of symbols depending on the Transport Block Size (TBS). A PRB is therefore composed of at least, 12 RE (Fig. 1.3). Consequently, gNBs allow UEs to receive or transmit data by allocating them PRBs.

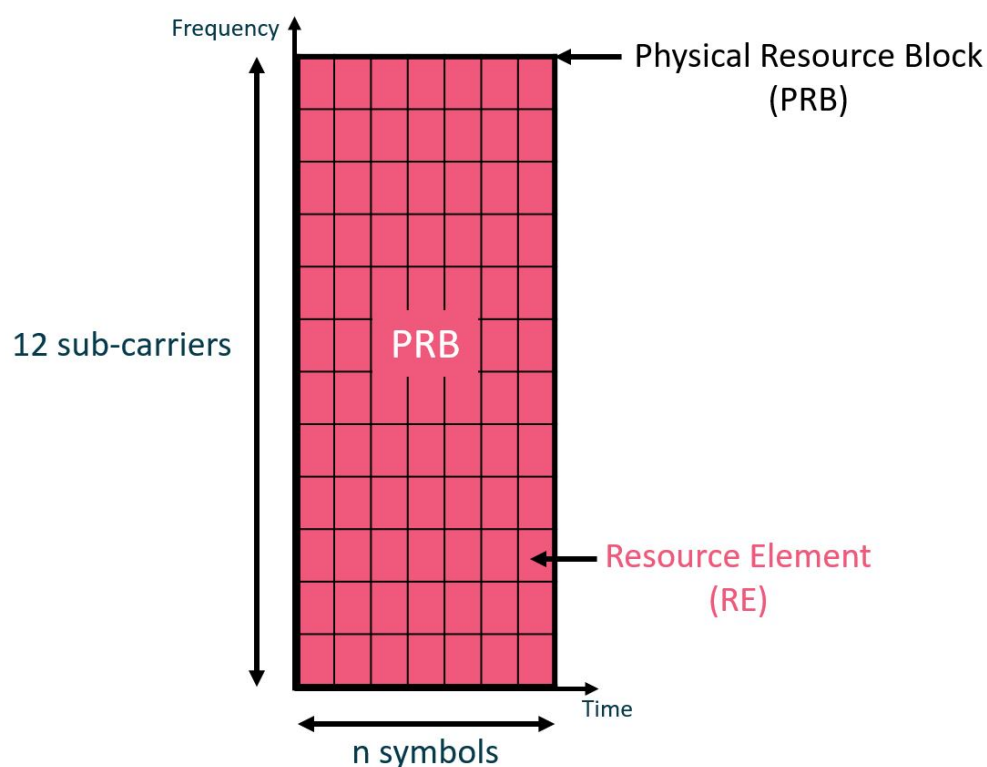


Figure 1.3 – Physical Resource Block.

⁷ 3GPP, « 5G NR, Physical channels and modulation », in: *3rd Generation Partnership Project (3GPP), TS 38.211 version 15.2.0 Release 15* (2018).

1.2.2 Frequency and time organization

The number of sub-carriers is correlated to the number of available PRBs. The potential throughput that can be provided to end users is conditioned by the size of the channel bandwidth but also by the Sub-Carrier Spacing (SCS). In 4G-LTE, the SCS, denoted Δf , is fixed to 15 kHz. It means that the maximum bandwidth is 1200×15 kHz, being 18 MHz (without the guard bands). In 5G-NR, the maximum bandwidth with the same SCS is around 50 MHz. As the maximum throughput of the system is directly correlated to the maximum bandwidth, apart from the 15 kHz, 5G-NR can use double, quadruple or even increase this value Δf eightfold (e.g. 30 kHz, 60 kHz, 120 kHz). Each value of Δf corresponds to a numerology. More precisely, the numerology n corresponds to $\Delta f = 2^n * 15$ kHz. Consequently, the numerology has a high influence on the maximum system bandwidth but also on the symbol duration.

The duration of one OFDM symbol is equal to $1/\Delta f$. The duration is all the more short as the numerology is high. This allows to reduce the transmission delay (e.g. uRLLC class of traffic). A Cyclic Prefix (CP) is included in the OFDM transmission in order to avoid inter-symbol interference. The standard has chosen to keep an identical ratio between the CP duration and the symbol duration. This induces that high numerology must be only used in small cells to take into account the propagation delay and avoid inter-symbol interference.

On the 5G-NR radio interface, 14 successive OFDM symbols constitutes one slot. This slot is equal to 1 ms in numerology 0 and $1/2^n$ ms in numerology n . The standard defines the duration of a sub-frame to 1 ms in order to have a self-sufficient reference time of the numerology. It means that a sub-frame contains 1 slot in numerology 0, 2 slots in numerology 1 and so on. The NR frame has a fixed duration of 10 ms, being 10 sub-frames (Fig. 1.4). The structure of this frame is used to periodically transmit signal (e.g. control message).

On the one hand, 5G deployment plans to use higher frequencies compared to LTE. On these high frequencies, the total available bandwidth is often larger as the number of reserved radio frequencies is lower, especially in France. This leads network operators to overbid in order to obtain these frequencies. On the other hand, the numerology is designed to reach very low latency. However, most of the current and available smartphones are only compatible until numerology 1. Although 5G standard planned to use different numerology, the OFDM principle with orthogonal carriers prevents from using multiple

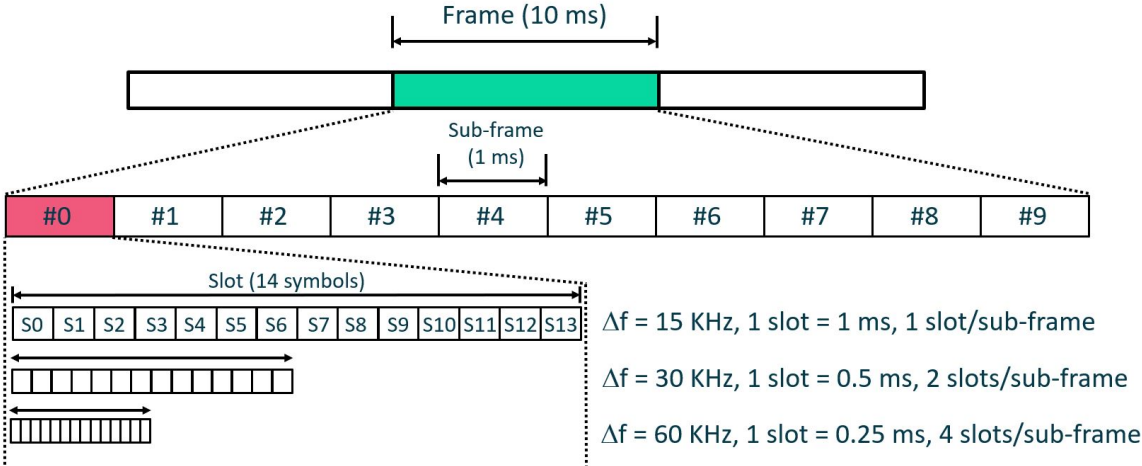


Figure 1.4 – Time organization on 5G-NR.

numerology in the same bandwidth part⁸ as the transmission duration is different depending on the numerology. Acquiring a large bandwidth on each frequency and make a good use of the numerology is therefore, a major concern for network operators but it is not the only lever to ensure aforementioned KPIs.

Based on this analysis, this thesis is not particularly focused on high frequency nor on studying numerology. Notwithstanding, important work remains to be done to improve the 5G-NR efficiency. As PRBs are in a limited number on a given time interval, wisely allocating PRBs is a crucial issue. One of the main objectives of this thesis is to emphasize that a wise resource allocation can drastically change the performances in terms of spectral and energy efficiencies, system capacity and QoS of users (delay, throughput, battery lifetime...).

1.2.3 Resource allocation

In 5G-NR the radio access is similar to the one in 4G-LTE. The time is split in slots and at each slots, one or many resources are allocated to one transmission. On the downlink, the symbols carrying control messages form the Physical Downlink Control CHannel (PDCCH), while the symbols carrying data messages form the Physical Downlink Shared CHannel (PDSCH).

The gNB notifies an UE of the resource allocation decision via a Downlink Control Information (DCI) on the PDCCH. At each sub-frame, the UEs have to decode the

8. *Ibid.*

PDCCH in order to ascertain if a DCI is intended to them. This procedure is known as the blind decoding. As the transmission of DCIs occurs regularly over a short period of time (each ms), the gNB needs to identify each UE with short addresses. This short address is called Radio Network Temporary Identifier (RNTI) and is coded on 16 bits. RNTI is obtained after the connection and synchronization with the gNB (using random access procedure). RNTI allows UEs to identify the DCI that are intended to them. As the transmission is adapted to the channel quality of the UEs, the DCI not only carries information about the allocated PRB but also the Modulation and Coding Scheme (MCS). This allows UEs to properly decode information received on the PDSCH.

Figure 1.5 illustrates the organization of the 5G-NR frame in downlink. The two main physical channels regarding resource allocation are PDCCH and PDSCH. In order to receive data, the first step for UEs is to regularly control if a DCI is intended to them (blind decoding) on the PDCCH. If their RNTI are linked with at least, one DCI on the PDCCH, it means that gNB allocated them some PRBs. Then, the second step for UEs is to decode the PDSCH with the adequate MCS (also given by DCIs) in order to receive data.

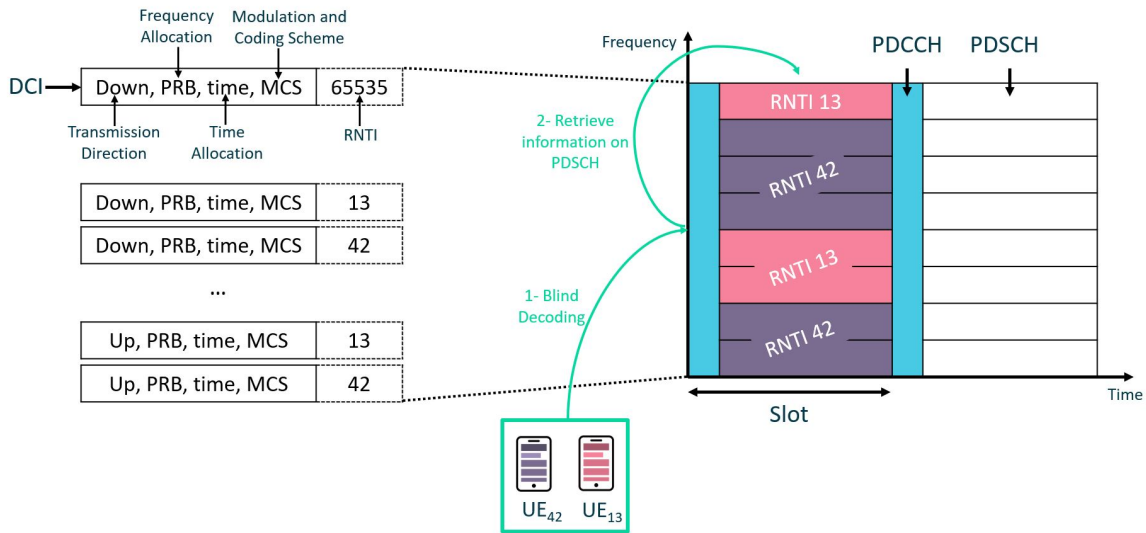


Figure 1.5 – Resource allocation for 2 UEs.

Knowing that the number of radio resources are limited over a period of time (cf. time and frequency organization of 5G-NR frame in the previous section), the main issue regarding the resource allocation for the gNB is to determine which PRB should be

allocated to which UE and find the appropriate time according to the channel quality and application requirements of UEs.

In the gNB, the scheduler is the part in charge of allocating PRB to UEs. The scheduler at once, has to take into account the variations of the channel quality of UEs (depending on signal attenuation) and the UEs' needs in terms of throughput and delay constraint. The global system capacity is directly related to these allocation decisions. In other words, wisely allocating radio resources allows to enhance the spectral efficiency and system capacity, thus to increase the number of UEs that an operator can accept in his network.

The resource allocation is not only restricted to the scheduling part. In the end, every parts of a network that directly or indirectly manage radio resources can be considered as an element of the resource allocation chain. For instance, ICI management directly influences PRB allocation to avoid interference between cells. To some extends, 5G routing can be also considered as a part of the resource allocation because information is forwarded from a node to another through a wireless link, so via the 5G-NR frame. Indeed the path chosen by the routing to transmit data from/to a UE could improve the KPI for this UE but also could optimized or degraded the KPI of others UEs using the same cells along the path.

Knowing that radio resources are in limited number, the main challenge of resource allocation lies in wisely managing the available bandwidth by allocating PRBs to UEs at the most suitable moment according to their needs (e.g. application requirements, delay constraints) and their radio channel quality. Resource allocation is thus, a crucial economic concern for network operators and allows to reach KPIs such as: spectral and energy efficiencies, system capacity and QoS of users.

1.3 Contributions and thesis outlines

1.3.1 Motivations

Following 5G KPIs, the main goals of this thesis are various. From gNB side (radio network operator), the objectives are to enhance the spectral efficiency and the system capacity to be able to accept more UEs in the network. From UE-side (users point of view), the main objective is to increase the QoS. It includes, increasing the effective throughput and reducing the latency. Moreover, in the wider context to reduce the

impact of Information and Communications Technology (ICT) equipment on the global environment and to reduce the network operator's expense, many studies have emerged on Energy Efficiency (EE) of mobile networks. Historically, the focus has been put on reducing energy consumption of Radio Access Network (RAN) as it is widely acknowledged that it constitutes the most energy consuming part of the mobile network. Work is now ongoing to provide energy efficient solutions for the whole 5G network. Another important objective of this thesis is thus, to help the design of sustainable communication medium by reducing the energy consumption on UE-side. This allows to also enhance their QoS as their battery lifetime is increased.

On the one hand, the most widespread and generic usage of 5G is the eMBB. On the other hand, the use of high numerology limits the context to small cells and numerology is not even compatible with current smartphones beyond numerology 1. Moreover, the use of different numerologies within the same sub-frame cannot currently be applied. This thesis is therefore, focused on eMBB traffic and numerology 0 as they concern most of human users and are the main challenges of current 5G networks and those to come.

In the first place, this thesis investigates the benefits of opportunistic resource allocation. This work emphasizes that most of 5G KPIs (e.g. spectral and energy efficiencies, system capacity) can be improved by the modification of scheduling and routing solutions. This work allowed us to identify the most suitable context for the use of cell-less. With cell-less, gNBs share information and cooperate with each other. This makes it easier to manage issues regarding the multi-cellular domain. This thesis therefore leverages cell-less context to improve the ICI management. Indeed, the interference issue is related to the multi-cellular domain. Classically, gNBs are prevented from transmitting on the same frequencies as their neighbors in order to avoid interference. This induces a high bandwidth waste and if the interference are badly managed, UEs located at cell-edges are strongly penalized (both by a high path-loss and a high magnitude of interference). Thanks to cell-less, gNBs are able to adopt a more dynamic approach on a shorter time scale to manage ICI. By combining the benefits of opportunistic resource allocation and cell-less, this thesis provides ICI management solutions that increase the spectral efficiency, system capacity and QoS of UEs.

1.3.2 Thesis outline

In order to provide a complete understanding of the work of this thesis to the reader, this manuscript is organized as follows. In Chapter 2, the main objective is to introduce

the radio resource allocation concern by a deep analysis of the multi-user diversity. This parameter has a strong influence on resource allocation performance and cannot be neglected or it could be subject to sub-optimal solutions. The two next chapters introduces the classical way to perform resource allocation (without cell-less approach). Chapter 3 introduces the mechanism related to energy efficiency of UEs. In this chapter, one new energy efficient QoS-oriented scheduler and one new meta scheduler providing energy efficiency for UEs on the most acknowledged schedulers are presented. In Chapter 4, a new solution of wireless routing that takes into account node and link states is presented. Finally, two new ICI management are introduced in Chapter 5 within the cell-less context.

1.3.3 Main contributions

The main contributions of this thesis are introduced below:

Multuser Diversity Regarding System Performance, Design and Evaluation of Radio Resource Allocation

According to Section 1.1.2, one of the main physical phenomenon that significantly impacts the wireless transmission quality is the multipath fading. The set of UEs that can transmit at a given time is called the Multi-User Diversity (MUD). Statistically, among a sufficient number of UEs, a subset of them will experience constructive multipath fading at any time. By allocating radio resources to this subset, MUD can be exploited to enhance the spectral efficiency and the system capacity. On the contrary, restraining MUD allows to compress the transmission time in order to reduce the energy consumption of UEs. In this study, the influence of multuser diversity on system performance is analyzed. We also show that overlooking MUD can lead to design sub-optimal solutions and to choose inaccurate simulation parameters leading to misleading performance evaluations of resource allocation solutions.

Fair Energy Efficient Scheduler Providing High System Capacity for Wireless Networks

Following results obtained in the previous contribution (that is further detailed and explained in Chapter 2), preventing resource allocation solutions from using the whole MUD allows to reduce energy consumption on UE-side at the expense of the spectral efficiency. Based on this analysis, a new scheduling solution called, Fair Energy efficient

scheduler for high system Capacity (FEC), is proposed. In underloaded network, the focus should be put on low energy consumption while, on the contrary in overloaded network, priority should be put on high system capacity. Even though efficient, this solution is unable to provide service differentiation and consequently full fairness and adequate QoS in many context particularly when UEs have different delay constraints or throughput requirements. FEC heightens the benefits of opportunistic scheduling (cf. Sections 1.1.2 and 1.2.3) by extending cross layer technique from physical layer to higher layers. Resource allocation is performed thanks to a system of weights that dynamically accounts for all three: the radio conditions, system energy profitability to select a specific user and its experienced QoS. Performance evaluation shows that this solution offers efficient global system energy consumption without downgrading the system capacity while enabling good QoS differentiation and fairness.

Adaptative Multi User Diversity Meta Scheduler to Reduce Energy Consumption

This contribution extends the concept of FEC scheduler (i.e. varying the MUD usage depending on the traffic load) to any schedulers. In order to address this challenge, we propose a meta-scheduler acting before the scheduling process. The Adaptative Multi User Diversity meta Scheduler (AMUDS) is able to manage the MUD usage made by schedulers depending on the traffic load of the cell and the spectral efficiency. This allows to adjust the adequate number of UEs in active mode during a time slot in order to minimize the transmission time and thus, to reduce their energy consumption. Performance evaluation emphasizes that this solution is compatible with the most acknowledged scheduling solutions while keeping their intrinsic properties.

Buffer Occupancy and Link State Opportunistic Routing for Wireless Mesh Networks

As aforementioned in Section 1.2.3, resource allocation is not limited to the scheduling and can be also extended to routing. This contribution is a new multihop wireless routing protocol that opportunistically takes profit from variations of radio conditions in terms of path loss, shadowing and multipath fading to maximize the system capacity. However, guaranteeing high system capacity should not evade the packet delay minimization objective. Consequently, the best path should not only be considered as the path with best throughput but a combination of a good link throughput and, in addition, low router

buffer occupancy load. Taking into account the available router buffer occupancy in its path selection, our proposal uses queuing theory information in order to also provide an efficient load balancing solution that adequately distributes the traffic load in the whole network. Exploiting this information, our solution dynamically adapts the selected path across time avoiding overexploited efficient links as well as low throughput link usage. This adaptation is performed considering each link state and the amount of channel information available. This improves the throughput and delay with only small marginal overhead cost. Our proposal applies to all wireless multihop networks, with increased benefit for extending cell coverage. We demonstrate through our simulation study that our solution raises the system capacity by more than 50% in several scenarios as well as reduces packet delays compared to state-of-the-art protocols such as Ad-hoc On-demand Distance Vector (AODV), Optimized Link State Routing (OLSR) and Link State Opportunistic Routing (LSOR).

Dynamic Cell-Less Radio Access Network Meta-Scheduler for High System Capacity Increase

Classical ICI management solutions have to restrict the usage of the bandwidth in order to avoid interference. By leveraging the new approach cell-less (cf. Section 1.1), gNBs cooperate between each other and ICI can be smoothly managed on each PRB. This contribution called Dynamic Cell-less Radio Access Network Meta-Scheduler (DC-RAN-MS) dynamically handles for each gNB the management of radio resources depending on the interference potentially experienced by users. Performance evaluation shows that the DC-RAN-MS offers an increased system capacity by optimizing the usage of bandwidth while reducing the magnitude of interference received.

Hybrid Joint-Transmission Multi-Point Coordination for Inter-Cell Interference Management

This contribution is an extension of DC-RAN-MS but this contribution investigates the benefits of the Joint Transmission Coordinated MultiPoint (JT-CoMP). With JT-CoMP, gNBs share between each other, UE data, scheduling information and UEs' channel quality. In addition JT-CoMP allows at least 2 gNBs to simultaneously transmit information to a given UE on the same PRB without ICI. The Hybrid Joint-Transmission Coordinated MultiPoint (H-JT-CoMP) dynamically performs its ICI management according to Channel State Information (CSI). This allows to make a wise CoMP usage according to

the magnitude of interference received. Performance evaluation highlights an increased QoS and system capacity with a better fairness between inner and edges of the cell.

1.3.4 List of publications

The publications of this thesis are listed below:

- "Multiuser Diversity Regarding System Performance, Design and Evaluation of Radio Resource Allocation" submitted in Springer Personal Wireless Communication journal, 2023 (Chapter 2)
- "Fair Energy Efficient Scheduler Providing High System Capacity for Wireless Networks" published in Springer Applied Sciences journal, 2020 (Chapter 3)
- "Adaptative Multi User Diversity Meta Scheduler to Reduce Energy Consumption" submitted in IEEE Transactions on Green Communications and Networking , 2023 (Chapter 3)
- "Buffer Occupancy and Link State Opportunistic Routing for Wireless Mesh Networks" published in Springer Wireless Networks journal, 2021 (Chapter 4)
- "Dynamic Cell-Less Radio Access Network Meta-Scheduler for High System Capacity Increase" published in IEEE WoWMoM conference, 2020 (Chapter 5)
- "Hybrid Joint-Transmission Multi-Point Coordination for Inter-Cell Interference Management", published in IEEE Vehicular Technology conference, 2021 (Chapter 5)

INFLUENCE OF THE MULTI-USER DIVERSITY ON RESOURCE ALLOCATION SOLUTIONS

This chapter introduces the first contribution of this thesis by providing an in-depth analysis of MUD. Although MUD is often mentioned in the literature and well-known to researchers working on resource allocation, the novelty of this work lies in considering MUD as a fully-fledged issue. Firstly, this work highlights the influence of MUD on resource allocation solutions' performance and by extension on network performance. It shows the advantages and disadvantages that can be obtained by increasing or decreasing MUD. In particular, regarding spectral and energy efficiencies. Secondly, this chapter reviews all the elements that can impact MUD (and thus network's performance): from the design of resource allocation solutions to their performance evaluation (related to 3rd Generation Partnership (3GPP) standards). Thirdly, this chapter ends with a discussion on theoretical performance versus practical performance that highly depends on the manner to obtain the required CSI for each UE. This chapter therefore, can be seen as an introduction of resource allocation issues.

2.1 Introduction

Radio resource allocation consists in wisely allocating PRBs according to the available gNB bandwidth and UEs needs. Within a gNB, the scheduler is in charge of directly allocating PRBs to UEs and it notifies them of the allocation decisions via the PDCCH (for more details, cf. Section 1.2.3). This work is particularly focused on the scheduling but the conclusions are also concordant with other resource allocations processes. For instance, inter-cell interference management can be seen as a multi-cellular scheduling. The analysis provided in this chapter is consequently also true for interference management solutions.

Usually, schedulers are split into two classes. The first class gathers non-opportunistic schedulers. These schedulers do not take into account the channel quality between the gNB and each UE. For instance, this class contains legacy schedulers such as Random Access (RA)¹ and Round Robin (RR)². These solutions have been inspired by operating systems. They were used to schedule processes. While having some major downsides, they are the simplest methods to uniformly distribute resources across a system. RA allocates PRBs randomly while RR allocates PRBs to one UE after another. However, due to their non-opportunistic behavior, they overlook the specificities of wireless networks (e.g. path-loss, shadowing, multipath fading, cf. section 1.1.2) leading to poor performance regarding spectral efficiency and QoS. Although these solutions are somewhat outdated they continue to be studied for their ease of implementation³ or to be used in scheduler performance studies as a standard of comparison⁴.

The second class gathers opportunistic schedulers. These solutions take into account the channel quality of UEs when allocating PRBs. This kind of schedulers has been design in order to take into consideration the swift and significant signal attenuation (cf. Section 1.1.2) that impacts the transmission quality. The Maximum Signal-to-Noise Ratio (MaxSNR)⁵ is one of the most acknowledged opportunistic scheduler. MaxSNR allocates a PRB to the UE with the best Signal-to-Noise Ratio (SNR) on this resource in order to increase the system capacity. Usually, the closest UE to the gNB has, on average, a better SNR due to a low path-loss. An important side effect of this strategy is the lack of fairness regarding the UE's distance from the gNB. To address this issue, the Proportional Fair (PF) introduces a correction factor⁶. This solution considers the best ratio between short (instantaneous SNR) and long term throughputs (average SNR), leading UEs to have the same probability to access resources regardless of their locations

1. Gokhan, M. and Lang, T., « Random Scheduling Medium access for Wireless Ad Hoc Networks », *in: Proc. IEEE Int. Conf. on MILCOM*, vol. 2, October 2002, pp. 868–872.

2. Kuurne, A. and Miettinen, A.P., « Weighted Round Robin Scheduling Strategies in (E)GPRS Radio Interface », *in: Proc. IEEE Int. Vehicular Technology Conference (VTC)*, vol. 5, Sept. 2004, pp. 3155–3159.

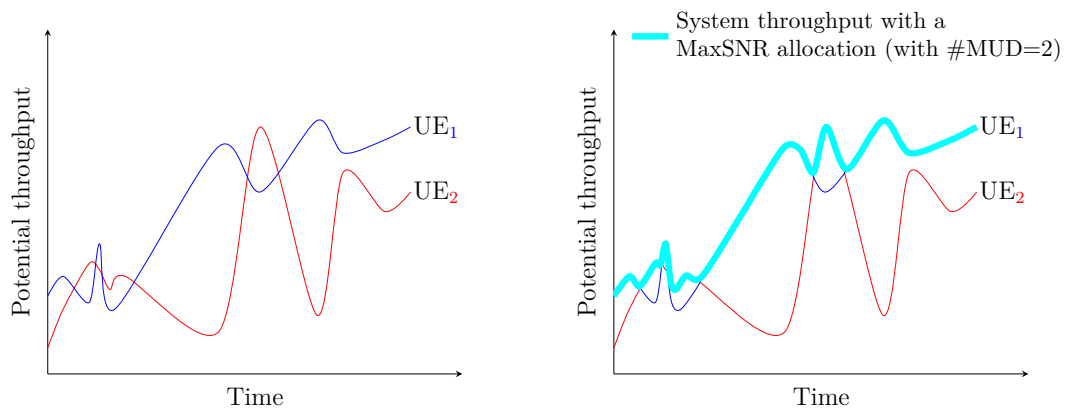
3. Prakash, P. and Chaitali, B., « Fair Resource Allocation to MIMO Wireless System Using Opportunistic Round Robin Scheduling Algorithm », *in: International Conference on Pervasive Computing (ICPC)*, 2015, pp. 1–3.

4. Minelli, M. et al., « Scheduling Impact on the Performance of Relay-Enhanced LTE-A Networks », *in: IEEE Transactions on Vehicular Technology* 65.4 (2016), pp. 2496–2508.

5. Tabatabaee, V. and Tassiulas, L., « Max-min Fair Self-randomized Scheduler for Input-buffered Switches », *in: Workshop on High Performance Switching and Routing (HPSR)*, 2004, pp. 299–303.

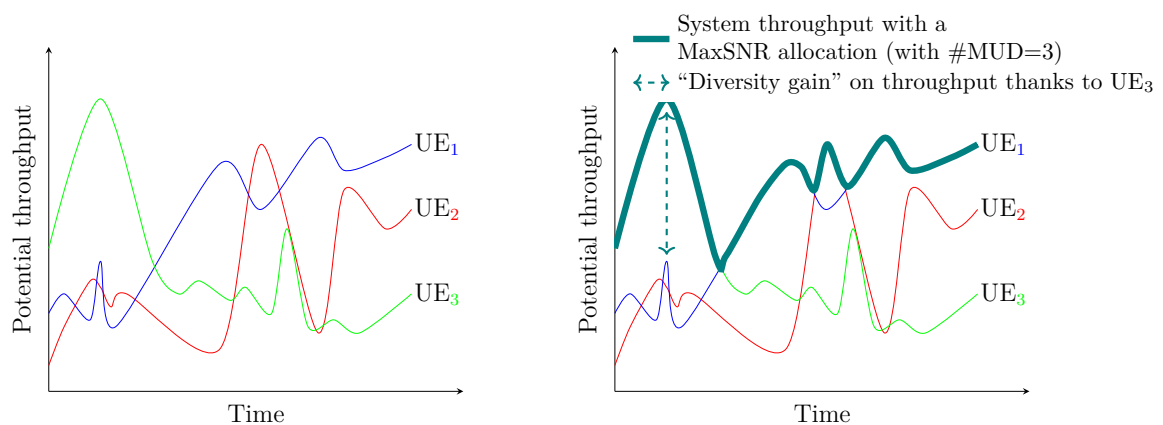
6. Yao, M., « Proportional Fair Scheduling for Downlink OFDMA », *in: Proc. IEEE Int. Conf. on Communications (ICC)*, June 2007, pp. 4843–4848.

in the cell. While MaxSNR focuses on short-term optimization, PF improves spectral efficiency over a long time scale. PF and MaxSNR introduce key concepts of opportunistic scheduling and various PF-based schedulers have been proposed since to further improve their performance in specific areas^{7 8}.



(a) Potential throughput of 2 UEs over the time. (b) System throughput with a MaxSNR allocation over the time.

Figure 2.1 – Potential throughput with a MUD of 2 UEs.



(a) Potential throughput of 3 UEs over the time. (b) System throughput with a MaxSNR allocation over the time.

Figure 2.2 – Potential throughput with a MUD of 3 UEs.

7. Hamouda H. Kabaou, M.O. and Bouhlel, M.S., « A Cross-Layer Downlink Scheduling Scheme for Balancing QoS in IEEE 802.16 Broadband Wireless Access Systems », in: *IEEE 86th Vehicular Technology Conference (VTC-Fall)*, Sept. 2017, pp. 1–5.

8. Ge, X., Jin, H., and Leung, V.C.M., « CDF-Based Scheduling Algorithm for Proportional Throughput Fairness », in: *IEEE Communications Letters* 20.5 (2016), pp. 1034–1037.

MUD can be defined as the pool of UEs that can be scheduled (i.e. available to transmit or receive data) at a given time. Figures 3.8(c) and 2.2 illustrate the potential throughput obtained with MaxSNR when the set of MUD contains two and three UEs, respectively. When MUD is composed of two UEs, the potential throughput of UEs highly varies over the time due to signal attenuation (Fig. 2.1(a)). According to its mode of operation, MaxSNR allocates resources to UEs which have the best SNR in order to increase the spectral efficiency (Fig. 2.1(b)). On Figure 2.2(a), MUD is now formed of three UEs. As UE_3 has the best radio conditions at the beginning, MaxSNR allocates resources to this specific UE (Fig. 2.2(b)). This provides an increase in the spectral efficiency in comparison with the MaxSNR allocation when 2 UEs were in the system (Fig. 2.1(b)). This difference is called the multi-user diversity gain. It means that the more UEs can be scheduled at a given time, the more resource allocation solutions can take benefits from the MUD to provide better performance (until a threshold related to the number of sub-carriers of the system).

Consequently, the more the MUD (i.e. the number of UEs that can be scheduled) is large, the more it is possible to leverage the multi-user diversity gain. Indeed, statistically among a sufficient number of users, a subset of them experience constructive multipath fading at any time. This particular mechanism can be defined as the driving force of any opportunistic resource allocation solutions. In fact, the differences in performance provided by aforementioned schedulers can be mostly attributed to the manner they use MUD because it highly impacts the network's performance. For instance, as MaxSNR highly favors UEs close to gNB, it reduces its MUD by serving first the closest UEs then the farthest ones. This explains why PF could provide a better spectral efficiency than MaxSNR as it does not segregate UEs according to their position and keeps a use of the whole MUD.

In the literature, MUD is mainly mentioned as a lever to increase performance of schedulers⁹ and networks^{10 11} but also to improve the physical layer security^{12 13}. Some

9. Taejoon, K. and Jong-Tae, L., « Queuing Analysis in a Multiuser Diversity System With Adaptive Modulation and Coding Scheme », *in: IEEE Transactions on Vehicular Technology* (2011), pp. 338–342.

10. Heidarpour, A.R., Ardakani, M., and Tellambura, C., « Multiuser Diversity in Network-Coded Cooperation: Outage and Diversity Analysis », *in: IEEE Communications Letters* 23.3 (2019), pp. 550–553.

11. Liwei, Y., Bo, B., and Wei, C., « On Energy Efficiency Maximization in Downlink MIMO Systems Exploiting Multiuser Diversity », *in: IEEE Communications Letters* 18.12 (2014), pp. 2161–2164.

12. Yujia, H. and Xiaofeng, T., « Secrecy Outage Analysis of Multiuser Diversity With Unequal Average SNR in Transmit Antenna Selection Systems », *in: IEEE Communications Letters* 19.3 (2015), pp. 411–414.

13. Inkyu, B., Su Min, K., and Dan Keun, S., « Effects of Multiple Antennas and Imperfect Channel

other works also attempted to find an optimal number of UEs for maximizing the MUD gain^{14 15}. However, these results are quite questionable as they highly depend on simulation parameters (further explained in Section 2.3.3). Consequently, MUD is an important parameter to design efficient resource allocation solutions but it is rarely studied as a fully-fledged issue.

The rest of this work is organized as follows. Section II introduces the system description, Section III provides an analysis of the MUD while Section IV discusses of the influence of the MUD in realistic environments. Finally, the main conclusions and implications are drawn in Section VI.

2.2 System description

In the radio access interface under study, the physical layer is considered to operate using a Time Division Duplex mode (TDD) which allows a good compatibility with the OFDM based transmission mode¹⁶. The numerology is assumed equal to zero. According to 5G specification with this numerology, the frame is divided in sub-frequency called sub-carriers in the frequency domain. In the time domain, the frame is divided in slots (1 ms duration)¹⁷. As the sub-carriers spacing is equal to 15 kHz (numerology zero) it means that a PRB is defined as 180 kHz (12 sub-carriers * 15 kHz) in the frequency domain. In the time domain, in order to facilitate the analysis of the results, a PRB has a constant duration of 0.5 ms (7 consecutive OFDM symbols). Consequently, 2 PRBs can be allocated in the time domain for each sub-frame (similar to 4G frame, cf. Section 1.4)

The gNB forwards information to UEs via DCIs through the PDCCH. For instance, a DCI indicates to UEs the MCS to be used to decode messages, the PRBs which should carry their data and so on¹⁸. Each PRB can be allocated to any UEs with a specific MCS. Transmissions performed on different PRBs by different UEs are considered to have

Knowledge on Secrecy Multiuser Diversity », *in: IEEE Communications Letters* 19.9 (2015), pp. 1564–1567.

14. Taejoon, K. and Jong-Tae, L., « Capacity Analysis and Feedback Threshold Optimization in Fair Multiuser Diversity System », *in: IEEE Transactions on Vehicular Technology* 61.9 (2012), pp. 4189–4194.

15. 3GPP, « Max UEs/Subframe for Optimum E-UTRA DL Performance (5-20 MHz) », *in: 3rd Generation Partnership Project (3GPP), TSG-RAN WG1 R1-071354* (2007).

16. Kela, P., Turkka, J., and Costa, M., « Borderless Mobility in 5G Outdoor Ultra-Dense Networks », *in: IEEE Access* 3.3 (2015), pp. 1462–1476.

17. 3GPP, « 5G NR, Physical channels and modulation ».

18. *Ibid.*

independent channel state variations. According to classical hypothesis in scheduling, on each PRB, a MCS adapted to the channel state (between the gNB and the selected UE) is assumed. Full knowledge of radio conditions is supposed to be available at the transmitter¹⁹. Thanks to SNR measurement of the signal sent by each UE (for instance, after CSI report), the gNB is thus, able to estimate their channel quality for a given PRB. This requires that UEs have to transmit their control information successively on each PRB once for several frames (in order for the gNB to successfully refresh the channel state information).

The channel gain between the gNB and the UE k on the PRB n in the frequency domain is given by:

$$G_{k,n} = h 10^{\frac{X\sigma}{10}} \left(\frac{d_{ref}}{d_k} \right)^\alpha \quad (2.1)$$

where h represents the Rayleigh multi-path fading, which is modeled by an exponential distribution, X is a standard Gaussian random variable, σ is the standard derivation of shadowing in dB, d_{ref} is the reference distance, d_k is the distance between UE k and gNB while α is path loss exponent. The SNR computation of UE k on PRB n associated to gNB is given by:

$$\gamma_{k,n} = \frac{P_n G_{k,n}}{B_{sub} N_0} \quad (2.2)$$

where parameter P_n is the the transmitted power on PRB n . Parameter N_0 is the thermal noise power density and parameter B_{sub} is the sub-carrier spacing²⁰. To compute the spectral efficiency $\eta_{k,n}$ of UE k on PRB n associated to gNB, the Shannon's formula is used such as:

$$\eta_{k,n} = \log_2 \left(1 + \frac{\gamma_{k,n}^i}{\Gamma} \right) \quad (2.3)$$

where parameter Γ is a SNR correction factor that takes into account the difference between the information-theoretic performances and the practical implementation of the

19. Li, Y.G., Seshadri, N., and Ariyavisitakul, S., « Channel Estimation for OFDM Systems with Transmitter Diversity in Mobile Wireless Channels », *in: IEEE J. Sel. Areas Commun.* 17.3 (1999), pp. 461–471.

20. Ezzaouia, M. et al., « Autonomous and Dynamic Inter-Cell Interference Coordination Techniques for Future Wireless Networks », *in: IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, Oct. 2017.

MCS²¹ defined as follows:

$$\Gamma = -\frac{\ln(5 E)}{1.5} \quad (2.4)$$

where E is a BER target. In order to reduce the simulation time (which can be very long due to the realistic traffic of UEs), we choose to model a sub-system for the 5G NR frame. The frame duration is reduced to 5 ms but scheduling decisions are taken for a slot so it does not influence the mode of operation of schedulers. In the frequency domain, only 50 PRBs are available. It is similar to a LTE frame with a channel bandwidth of 10 MHz. Thus, the benefits of the opportunistic resource allocation can be correctly studied and cannot be attributed to an increase of the number of sub-carriers (in 5G the channel bandwidth is larger than in LTE, so the throughput is consequently increased only due a larger number of PRBs).

In the following, user traffic model is a Youtube traffic²² of 150 kbps. After an initial burst phase, this traffic continues with a throttling procedure that alternates between chunk transmission phases and sleeping phases. For full buffer sources, the user traffic model is the same but there is no sleeping phase.

The power consumption of a UE is computed as follows: let be A_k the mode of the UE k ($A_k = 1$ when the UE k is in active mode, while $A_k = 0$ when UE is in sleep mode). The supplementary power necessary to wake up the UE k from the sleep mode to the active mode (i.e. to supply power to the radio module) is denoted C_k . C_{nk} represents the power required to receive and decode data on a n^{th} allocated PRB (i.e. power required for the antenna transmission). The amount of power required to be active (to switch the radio module on) is higher than the power needed to receive data on an additional PRB if UE is already set in active mode. Consequently, $C_k > C_{nk}$. In the following of this work, C_k and C_{nk} values are considered respectively equal to 110.2 mW and 46.8 mW, for all k in accordance with measured hardware consumption²³.

These parameters are described in table 2.1.

21. Seo, H. and Lee, B.G., « A proportional-fair Power Allocation Scheme for Fair and Efficient Multiuser OFDM Systems », *in: IEEE Global Telecommunications Conference*, Nov. 2004, pp. 3737–3741.

22. Horvath, G. and Fazekas, P., « Modelling of YouTube Traffic in High Speed Mobile Networks », *in: 21th European Wireless Conference*, Nov. 2015, pp. 1–6.

23. Gueguen, C. and Manini, Malo, « Dynamic Tradeoff between Energy and Throughput in Wireless 5G Networks », *in: Wireless Communications and Mobile Computing Journal* (2018), pp. 1–12.

Parameters	Definition
Cell radius	500 m
Number of PRBs	50
Number of sub-frame	5
Antenna transmit power (P_n)	43 dBm (20 W)
Standard deviation of shadowing (σ)	8 dB
Path-loss exponent (α)	3.5 (urban context)
Target BER (E)	5×10^{-5}
Sub-carrier spacing (B_{sub})	15 kHz
Thermal noise power density (N_0)	- 174 dBm/Hz
Simulation duration	500 000 frames

Table 2.1 – Simulations parameters.

2.3 Contribution: analysis of the multi-user diversity

The main objective of this section is to provide an in-depth review of MUD and associated mechanisms. This section highlights that MUD is an important characteristic to consider for resource allocation but very sensitive. According to their mode of operation, resource allocation solutions exploit MUD differently and therefore, may provide very diverse system performance. In this study, we also show that the simulation parameters used to evaluate resource allocation solutions have consequences on the MUD usage. In some cases, it could lead to misleading analysis of the performance evaluation.

2.3.1 On system performance

The main objective of this subsection is to enlighten the impact of MUD on system performance. To do so, three KPIs are studied: the spectral efficiency, the mean energy consumption of UEs and the mean packet delay. Note that the packet delay is delay between the creation of the packet and its reception by the UEs. Figures 2.3 and 2.4 illustrate the behavior of non-opportunistic and opportunistic schedulers when the MUD is regulated, respectively. For instance, PF(1) means PF is only allowed to schedule one UE per sub-frame while PF(full) is the normal behavior of PF. Preventing schedulers from using the whole MUD, allows to clearly identify the impact of MUD on system performance.

As non-opportunistic schedulers do not take into account the channel state of UEs in their resource allocation decision, they do not take benefits from the increase of MUD to enhance the spectral efficiency (Fig. 2.3(a)). They also provides the same QoS to

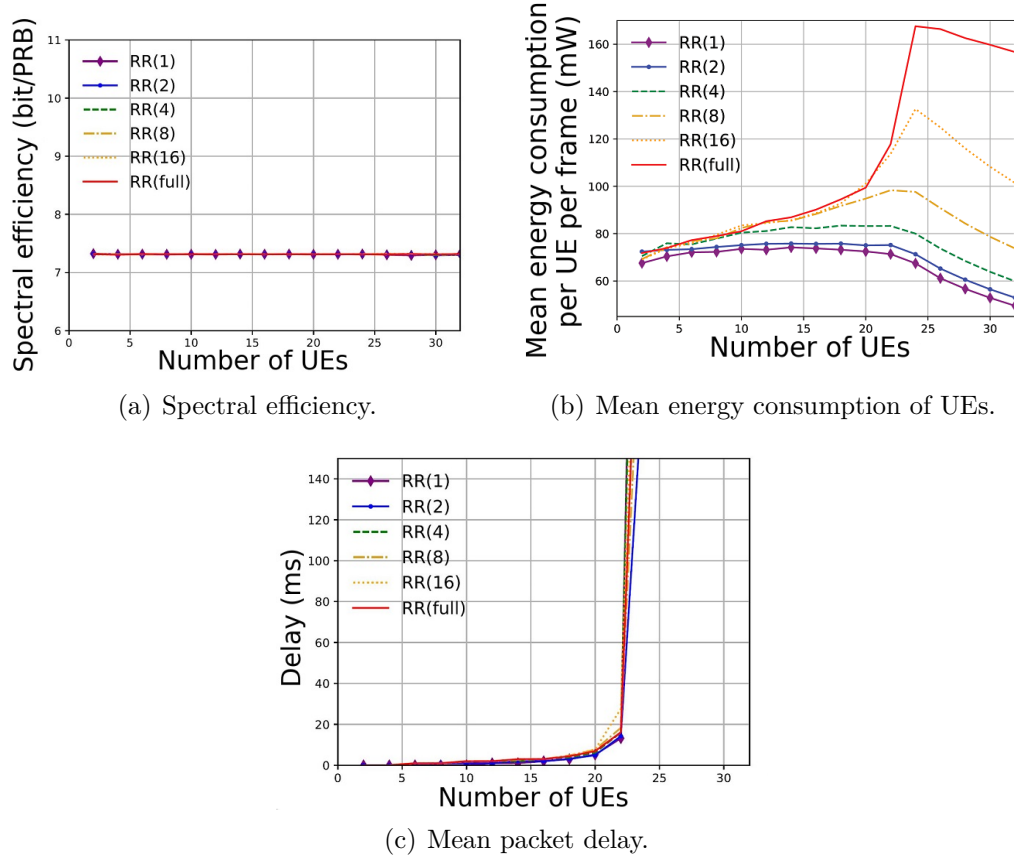


Figure 2.3 – Influence of multi-user diversity on non-opportunistic schedulers performance.

UEs regardless the set of UEs that can be scheduled on a given subframe (Fig. 2.3(c)). However, restraining the number of UEs per sub-frame is a well known manner to reduce the power consumption of UEs. It allows to compress the transmission time. For instance, with the Discontinuous Reception (DRX) technique, UEs get into sleeping mode during a certain period of time to avoid to listen the network when they have no PDCCH to decode²⁴. Its means that only a small number of UEs in the network are scheduled per TS. Using the same approach, Figure 2.3(b) shows that restraining the use of MUD results in reducing the power consumption of UEs. On the contrary, RR(full) schedules all the UEs in the network which leads to increase their mean power consumption.

Due to their opportunistic behavior, opportunistic schedulers take benefits from the increase of the MUD to enhance their spectral efficiency. This relies on the fact that among

²⁴. 3GPP, « Medium Access Control (MAC) Protocol Specification », in: *3rd Generation Partnership Project (3GPP), TS 38.321 v18.8.0* (Mar. 2022).

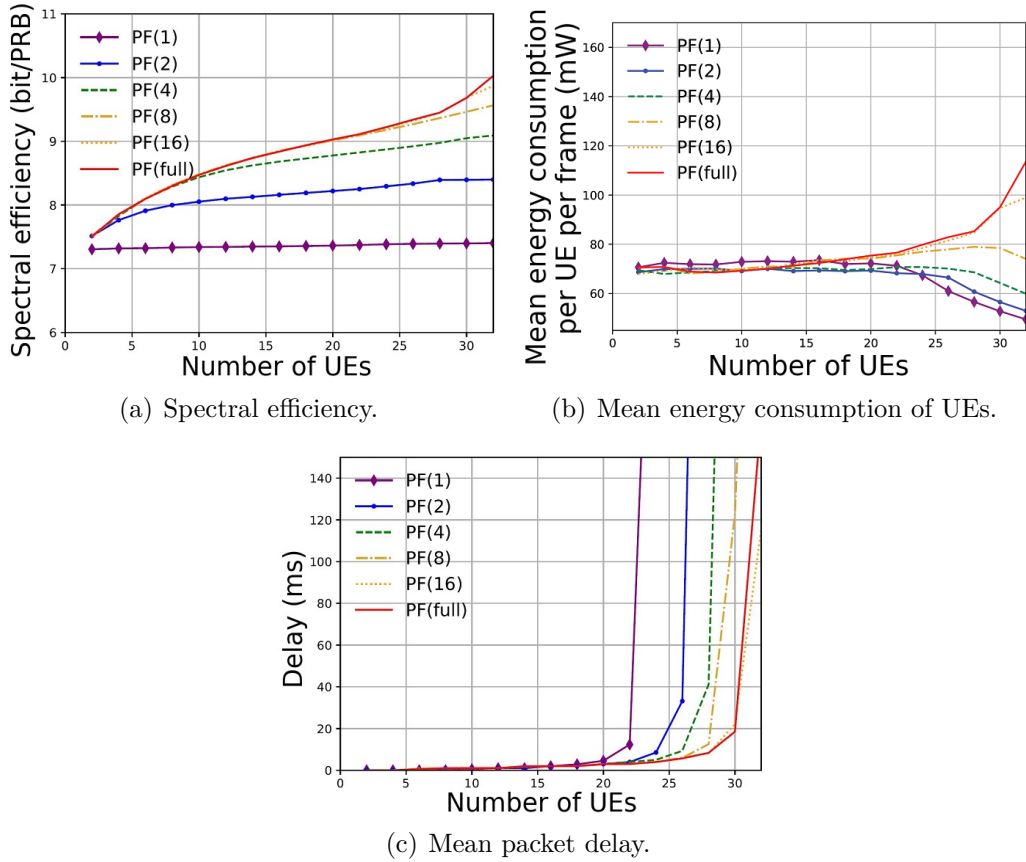


Figure 2.4 – Influence of multi-user diversity on opportunistic schedulers performance.

a sufficient number of UEs, some of them are likely to experience constructive fading at a given time. It means that the more the MUD increases, the more the set of UEs that experiences good combination of radio conditions is statistically rising. Opportunistic schedulers, hence, performs wiser allocations as MUD is growing.

This leads to highly impact the spectral efficiency (Fig. 2.4(a)) and the QoS (Fig. 2.4(c)) depending on MUD usage. As aforementioned for non-opportunistic schedulers, MUD has also an influence on the power consumption of UEs for opportunistic schedulers (Fig. 2.4(b)).

MUD is thus, an important mechanism in the resource allocation process that can significantly impact the system performance. Depending on their modes of operations, resource allocation solutions do not exploit MUD similarly. This dissimilitude in the use of MUD is one of the main criterion which explains that performances provided by resource allocation solutions highly differ. By wisely using MUD, resource allocation

solutions can significantly increase their performance. Techniques such as the DRX reduce unintentionally the use of the MUD, leading to reduce the power consumption of UEs at the expense of a lower system capacity. According to simulation results, we could propose a wise scheduling solution which reduces, for instance, the MUD usage when radio resources are abundant (at low traffic load) in order to make energy saving, while increasing the MUD usage as the traffic load rises to optimize the system capacity (further explained in Chapter 3).

2.3.2 On the design of resource allocation solutions

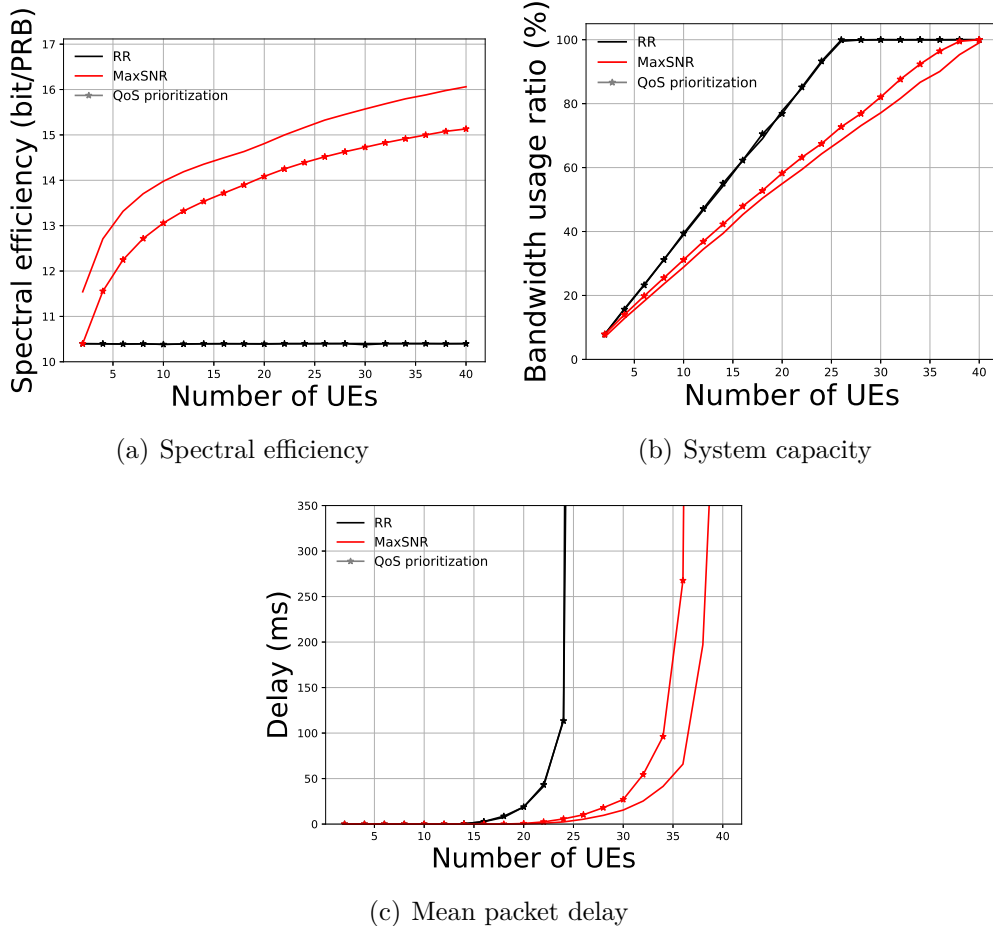


Figure 2.5 – Influence of QoS classes prioritization on system performance

In 5G, three main usage scenarios and applications have been envisaged: eMBB,

mMTC and uRLLC²⁵. Following this global trend, new resource allocation solutions that sort UEs according to their speed²⁶, distance from the gNB, QoS requirements and so on, have emerged. For instance, in ICI management, UEs are often split into two classes: UEs in cell overlapping areas and UEs in the inner of the cell²⁷. In the scheduling research field, UEs are often sorted according to their application requirements^{28 29}. Regardless the meaning of such distribution into classes, it means that UEs of a same classe are often scheduled together, leading to reduce the MUD. This is all the more important as the number of classes is large.

In order to analyze the impact of such divisions on MUD, three KPIs are studied: the spectral efficiency, the bandwidth usage ratio and the mean packet delay (Fig. 2.5). In this scenario, there are two classes of UEs based on their QoS requirements following 3GPP specifications³⁰. The first group has a packet delay budget of 100 ms, a 5G QoS Identifier (5QI) set to 1 and the traffic acts as a conversational voice. The last group has a packet delay budget of 300 ms, a 5QI set to 4 and the traffic acts as a non-conversational video (buffered streaming³¹). Note that these 5QI values are also concordant with those of the QoS Class Identifier (QCI) in LTE. Following the solutions aforementioned, these two classes are scheduled separately and the first group has the priority in the scheduling process.

Due to the non-opportunistic behavior of the RR and its mode of operation, the QoS prioritization has no incidence on spectral efficiency (Fig. 2.5(a)), system capacity (Fig. 2.5(b)) and mean packet delay (Fig. 2.5(c)). MaxSNR is opportunistic and takes benefits from MUD which is the driving force of any opportunistic schedulers. However, by splitting UEs into QoS classes, the set of UEs to schedule is thus reduced and the subset of UEs with constructive fading is even more restricted. This leads MaxSNR to provide a bad spectral efficiency, system capacity and mean packet delay with the QoS

25. 3GPP, « 5G: Study on scenarios and requirements for next generation access technologies », *in: 3rd Generation Partnership Project (3GPP), TR 38.913 version 14.3.0 Release 14* (Oct. 2017).

26. Saffar, I. et al., « Deep Learning Based Speed Profiling for Mobile Users in 5G Cellular Networks », *in: IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–7.

27. Shami, T.M. et al., « User-centric JT-CoMP clustering in a 5G cell-less architecture », *in: IEEE Indoor and Mobile Radio Communications (PIMRC)*, Nov. 2018, pp. 177–191.

28. Nasralla, M.M., « A Hybrid Downlink Scheduling Approach for Multi-Traffic Classes in LTE Wireless Systems », *in: IEEE Access* 8 (2020), pp. 82173–82186.

29. Kumar, A., Abdelhadi, A., and Clancy, C., « A Delay Optimal Multiclass Packet Scheduler for General M2M Uplink », *in: IEEE Systems Journal* 13.4 (2019), pp. 3815–3826.

30. 3GPP, « Policy and Charging Control Architecture », *in: 3rd Generation Partnership Project (3GPP), TS 23.203 R15 v15.5.0* (2019).

31. Horvath and Fazekas, *op. cit.*

prioritization.

These results highlight the importance of MUD in the design of resource allocation solutions. Dividing the set of UEs into QoS classes should be done cautiously or be subjected to induce sub-optimal performances.

2.3.3 On the impact of simulation parameters

Previous sections show that MUD has a strong influence on system performance and overlooking MUD can lead to the design of sub-optimal solutions.

Beyond that, the manner to carry out a performance evaluation can also impact MUD. For instance, 3GPP defines several scenarios to evaluate a solution according to the context (indoor, dense urban, rural and so on)³². In these scenarios, 3GPP recommends the following service profiles for UEs: full buffer sources or non-full buffer sources. Although recommended scenarios are different according to the service profile, the main goal of this subsection is to highlight that the MUD usage can be distorted according to the type of sources.

In this subsection, two groups of UEs are considered: the first is close to the gNB (100 m) while the second is further away (300 m). Performance evaluation focuses on three schedulers: RR, MaxSNR and PF. Figure 2.6 illustrates the spectral efficiency of these solutions for a given traffic load depending on the service profile (either non-full buffer or full buffer sources).

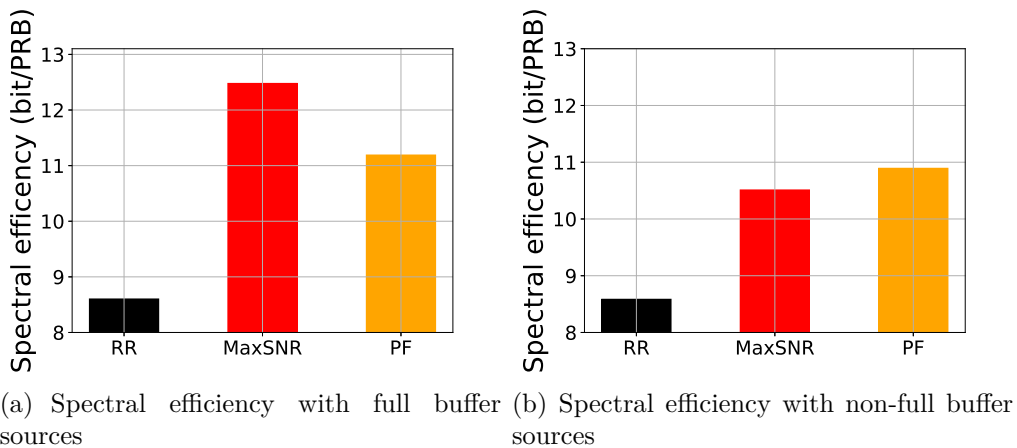


Figure 2.6 – Spectral efficiency according to the service profile

32. 3GPP, « 5G: Study on scenarios and requirements for next generation access technologies ».

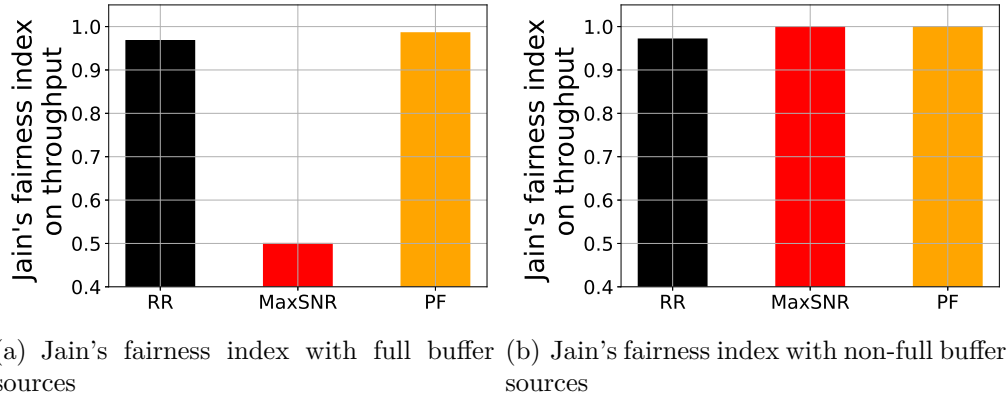


Figure 2.7 – Jain's fairness index on throughput according to the service profile

Regardless the type of sources, as the RR is non-opportunistic, it provides the same performance (Figs. 2.6(a), 2.6(b)) : 8,6 bit/PRBs.

With full buffer sources, UEs always have bits to transmit. Consequently, as MaxSNR allocates resources to the UE with the best radio conditions, it fully takes benefits from this service profile and mostly serves UEs close to the gNB. This allows MaxSNR to enhance its spectral efficiency (Fig. 2.6(a)). On the contrary, PF is designed to provide fairness regarding the distance and allocates as much resources to the two groups of UEs, leading to a lower spectral efficiency.

With non-full buffer sources, UEs have a more realistic model of traffic and do not transmit bits constantly. This leads MaxSNR to serve UEs far from the gNB when UEs closer have no remaining packet in their buffers. This mode of operation induces that UEs far from the gNB gain access to resources only after UEs closer have been served. Consequently, UEs far from the gNB might have sub-optimal radio conditions when they transmit. On the contrary, PF allocates resources to UEs when their radio conditions are good regardless the distance. Unlike MaxSNR, PF is designed to allocate resources on a long-time scale and fully take benefits from the MUD. This leads PF to provide a better spectral efficiency than MaxSNR (Fig. 2.6(b)).

The analysis of these results is accentuated by Figure 2.7 which illustrates the Jain's fairness index on throughput for the two service profiles. Note that, our Jain's fairness index is computed between the two groups of UEs. It means that the lowest value of this metric is 0.5.

As aforementioned, with full buffer sources MaxSNR only serves UEs close to the gNB, providing a good spectral efficiency at the expense of unfairness regarding the distance

of UEs (Fig. 2.7(a)). However, with non-full buffer sources as the traffic is variable and the system is not overloaded, UEs far from the gNB are served by MaxSNR, leading to a better fairness according to the distance of UEs (Fig. 2.7(b)). On the contrary, the mode of operation of RR and PF obliges them to equally allocate resources regardless of the service profile.

Beyond the fact that a given opportunistic scheduler can provide different spectral efficiencies and fairness depending on the service profile, the most important result to notice is that performances provided by MaxSNR and PF are antagonist in the two cases. With full buffer sources, MaxSNR provides the best spectral efficiency while with non-full buffer sources it is PF that provides the best performances. This is mostly related to the MUD usage of MaxSNR. On the one hand, UEs always transmit with full buffer sources and MaxSNR can only serve those which are close to the gNB. On the other hand, with non-full buffer sources, MaxSNR is obliged to serve UEs far from the gNB after UEs closer (which have on average, a better SNR), leading to virtually split the set of MUD into two parts.

This can induce a misleading interpretation of simulation results and more generally, to wrong conclusions based on the performance evaluation. The type of service profile should not be neglected and full buffer sources should be used cautiously or used to evaluate resource allocation solutions which do not rely on the MUD.

2.4 Discussion

In realistic environments, the MUD gain is highly related to the estimation of the channel state. In particular, the gNB relies on the CSI report to get feedback on the radio conditions of UEs' channel.

The CSI is described by several components such as: Channel Quality Indicator (CQI), Precoding Matrix Indicator (PMI), Channel State Information-Reference Signal (CSI-RS) resource indicator and so on³³. Usually, CSI reports are sent periodically by UEs to the gNB but the gNB can explicitly trigger a CSI report request thanks to a particular DCI (format 0₁ and 0₂). Two signals are used for the CSI report. The first is the Synchronization Signal Block (SSB) which is used for initial access. It means that the network is always transmitting this signal and no additional overhead will be caused

³³. 3GPP, « NR, Physical Layer Procedures for Data », in: *3rd Generation Partnership Project (3GPP), TS 38.214* (2021).

by a CSI report. However, the main drawback is that SSB are in limited number (20 PRB in the frequency domain) so a small part of the bandwidth is covered. The second signal is the CSI-RS. This signal allows a good flexibility in frequency and time domains but induces overhead. Consequently, the estimation of the radio conditions is a balance between having enough feedback to perform a wise resource allocation and limiting the overhead.

Although the commonly agreed hypothesis in scheduling field is that the radio conditions are available at any time and frequency, in the literature, some authors take into account imperfect CSI³⁴ or limited feedback³⁵ in their solution design. The MUD gain is thus reduced.

In order to address the challenge of reducing the overhead, feedback methods, feedback information and feedback reduction have sparked a lot of interest³⁶. In particular, different CQI feedback methods can be used such as: reporting the M best CQI bands, bitmap techniques, set of orthogonal functions and so on³⁷.

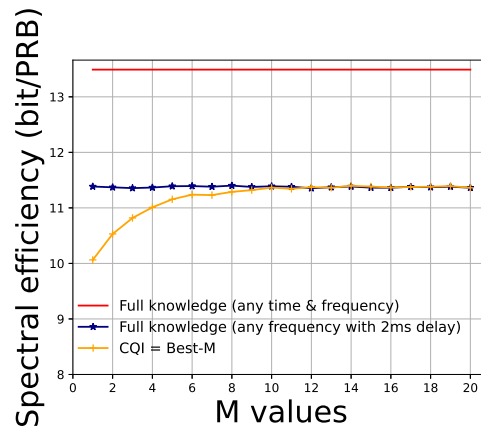


Figure 2.8 – Influence of Channel Quality Indicator on spectral efficiency.

Figure 2.8 illustrates the variations of the MUD gain according to the number M of bands reported. Three solutions are evaluated with a MaxSNR scheduler: the first is a

34. Libo, J. et al., « Toward Optimal Resource Scheduling for Internet of Things Under Imperfect CSI », *in: IEEE Internet of Things Journal* 7.3 (2020), pp. 1572–1581.

35. Nguyen, A.H. and Rao, B.D., « CDF Scheduling Methods for Finite Rate Multiuser Systems With Limited Feedback », *in: IEEE Transactions on Wireless Communications* 14.6 (2015), pp. 3086–3096.

36. Love, D.J. et al., « An overview of limited feedback in wireless communication systems », *in: IEEE Journal on Selected Areas in Communications* 26.8 (2008), pp. 1341–1365.

37. Kolehmainen, N. et al., « Channel Quality Indication Reporting Schemes for UTRAN Long Term Evolution Downlink », *in: IEEE Vehicular Technology Conference*, Nov. 2008, pp. 2522–2526.

ideal and complete knowledge of the radio conditions (classical assumption in scheduling field) in red, the second corresponds to a full knowledge at any frequency with 2 ms delay between the CQI computation and the reception of CSI report (which is a classical assumption³⁸) in blue, while the last corresponds to the best M CQI band scheme (with also 2 ms delay) in orange. According to the performance evaluation, full knowledge at any time and frequency provides the best spectral efficiency and is constant as it does not rely on a number M of frequency bands reported. The full knowledge solution with 2 ms delay provides a worse spectral efficiency than the full knowledge solution due to the time between the CQI computation and the reception of CSI report which can lead to imperfect CSI. Regarding the best M CQI bands, it is interesting to note that the spectral efficiency increases as the number M of bands grows until a threshold (around $M=10$). At this point, the MUD gain remains constant at the expense of additional and unnecessary overhead. This threshold is highly related with the number of sub-carriers allocated to each UE. It means that it depends on the size of the bandwidth in the frequency domain but also on the traffic load. In³⁹, authors attempt to characterize this threshold according to a number of UEs per subframe. However, they use PF as the scheduler, the service profile is full buffer sources and users are uniformly distributed over the cell area. As explained in the previous section, such parameters can lead to misleading interpretation so the conclusion on the maximum number of UEs per subframe is quite debatable.

CSI report thus plays an important role in the resource allocation process and in the MUD gain that stems from. Other mechanisms such as the DRX can also have an incidence on the MUD gain. By compressing the transmission time and restraining the number of UEs to be scheduled in order to save energy, the MUD is highly reduced which also leads to reduce the MUD gain. Consequently, the MUD gain does not only rely on how resource allocation solutions exploit the MUD but is also confronted to limitations of realistic environments.

2.5 Conclusion

The MUD is a strong characteristic that mostly influences the spectral efficiency, system capacity and power consumption of UEs. Splitting UEs into classes according to their profiles (distance, speed, QoS requirements and so on) should be done wisely because

38. *Ibid.*

39. 3GPP, « Max UEs/Subframe for Optimum E-UTRA DL Performance (5-20 MHz) ».

it always reduces MUD. As a result, the performances of opportunistic schedulers can be strongly affected. Simulation scenarios provided by 3GPP are very instructive and useful as long as all parameters and effects induced by such parameters are mastered and adapted to the context. In particular, the service profile can be either full buffer or non-full buffer sources which have a strong influence on performance evaluation. In some cases, it could lead to a misleading analysis of the scheduler performance due to unsuitable service profile to the context. Although the MUD is an important characteristic to provide good system performance, fully taking benefits from the MUD is quite unachievable in a realistic environment especially due to imperfect and incomplete CSI.

SCHEDULING

In the gNB, the scheduler allocates PRBs to UEs according to their needs and the available bandwidth (cf. Section.1.2.3). This step is thus crucial to guarantee an adequate QoS to UEs and to increase the system capacity. Based on the previous results (cf. Chapter 1), this chapter investigates the benefits of the opportunistic scheduling that compresses the transmission time to reduce the energy consumption on UE-side.

This chapter starts by introducing the second contribution of this thesis. The new proposed solution known as Fair Energy efficient scheduler for high system Capacity (FEC), relies on a QoS metric that allows to dynamically change its objective. At low traffic load, FEC reduces the energy consumption of UEs by compressing the transmission time. When the traffic load rises, FEC slightly drops the energy efficiency objective and starts to increase the spectral efficiency as the network requires more capacity.

Finally, this chapter ends with the introduction of the third contribution of this thesis. The Adaptative Multi User Diversity meta Scheduler (AMUDS) is an extension of the FEC's principle. It compresses the transmission time by reducing the MUD. Unlike FEC, AMUDS is a meta scheduler that is compatible with the most acknowledged schedulers. It thus provides to any of this schedulers a energy efficient feature while keeping their intrinsic properties.

3.1 Introduction

The scheduling is a subtle balance between giving enough radio resources to users to ensure their QoS and limiting the bandwidth waste to increase the system capacity. According to Section 1.2.3, this task is all the more difficult as the decision making process has to be made quickly and repeated regularly (in the order of a millisecond). In addition, the number of mobiles is continuously growing and applications are increasingly demanding in terms of throughput and delay constraints. This requires to heighten uninterrupted efforts to increase spectral efficiency while ensuring service differentiation and

fairness. User mobility also requires to reduce energy consumption in order to extend battery lifetime and in consequence to fight against global warming.

Thus, this topic sparks off a lot of interest and is well investigated in the literature. Advances in this field open new prospects to help the wireless network to reach stringent delay constraints, to optimize the transmission efficiency and to mitigate the energy consumption.

Conventional access methods like RR^{1 2 3} and opportunistic schedulers such as MaxSNR^{4 5} or PF and PF-based^{6 7 8 9 10} have already been introduced in Chapter 1. They constitute the set of schedulers often used as a standard of comparison in the literature and used in numerical simulation that implements 3GPP standard (e.g. LENA module of NS3¹¹). Although they are still used by network operators they can be improved especially on the energy consumption angle.

In the wider context to reduce the impact of ICT equipment on the environment and to reduce the network operator's expense, many studies have emerged on EE of mobile networks. Historically, the focus has been put on reducing energy consumption of RAN as it is widely acknowledged that it constitutes the most energy consuming part of mobile networks¹². Work is now ongoing to provide energy efficient solutions for the whole 5G network, particularly for network slices¹³. 3GPP also defined several KPIs regarding the

1. Simon, C. and Leus, G., « Round-Robin Scheduling for Time-Varying Channels with Limited Feedback », *in: IEEE 10th Workshop on Signal Processing Advances in Wireless Communications*, 2009, pp. 231–234.

2. Kuurne and Miettinen, *op. cit.*

3. Minelli et al., *op. cit.*

4. Wong, C. Y. and Cheng, R. S., « Multiuser OFDM with Adaptive Subcarrier, Bit, and Power Allocation », *in: IEEE J. Sel. Areas Commun.* (1999).

5. Bechir Dadi, M. and Belgacem Chibani, R., « Scheduling Performance's Study for LTE Downlink System », *in: International Conference on Green Energy Conversion Systems (GECS)*, 2017, pp. 1–4.

6. Viswanath, P., Tse, D. N. C., and Laroia, R., « Opportunistic Beamforming Using Dumb Antennas », *in: IEEE Transactions on Information Theory* 48 (June 2002), pp. 1277–1294.

7. Minelli et al., *op. cit.*

8. Hamouda and Bouhlef, *op. cit.*

9. Ge, Jin, and Leung, *op. cit.*

10. Masson, M., Altman, Z., and Altman, E., « Multi-User Collaborative Scheduling in 5G Massive MIMO Heterogeneous Networks », *in: IFIP Networking Conference*, 2020, pp. 584–588.

11. Patriciello, N. et al., « An E2E Simulator for 5G NR Networks », *in: Elsevier Simulation Modelling Practice and Theory (SIMPAT)*, vol. 96 (2019).

12. ETSI, « Assessment of mobile network energy efficiency », *in: ETSI Standard, 203 228 v1.3.1* (2020).

13. 3GPP, « Energy efficiency of 5G », *in: 3rd Generation Partnership Project (3GPP), TS 28.310 release 17* (Mar. 2021).

energy consumption for 5G networks such as the UE's battery lifetime¹⁴. Following this global trend, many solutions have emerged to reduce the energy consumption of UEs. For instance, the Discontinuous Reception (DRX) allows UEs to have long periods of sleep time and be active only when they have a DCI to decode on PDCCH¹⁵. This allows to compress the transmission time and thus, to reduce the energy consumption of UEs. Similarly, new schedulers focused on reducing the UE energy consumption have been proposed^{16 17} but often result in a trade-off between the spectral efficiency and the energy consumption of UE.

The novelty of this contribution differs from the literature in that it considers that spectral and energy efficiencies can be both optimized but at different times. Unlike the literature, this is achieved not via a trade-off but an optimization that depends on the QoS of users allowing to favor either the spectral or the energy efficiency. According to the results in Chapter 1, the new scheduler FEC has been designed following this analysis:

- In low loaded context, the available radio resources are plentiful and the system can easily meet user needs. Since guaranteeing high QoS is easily achievable due to a large surplus of available radio resources, the focus should be put on energy rather than on the system throughput.
- In highly loaded and overloaded system, the radio resources are particularly valued and the network meets high difficulties to satisfy all users. Since it could be considered that satisfying user delay constraints and throughput requirements is more important than reducing energy consumption at all costs, the lack of available PRB requires that schedulers focus on spectral efficiency in order to preserve QoS and reach high system capacity.

Consequently, one of the major issue here is to find an accurate metric able to define the current objective of the scheduler according to the network state. We used the same metric than the Weighted Fair Opportunistic (WFO) which is a QoS-oriented scheduler¹⁸. It is based on the Packet Delay Outage Ratio (PDOR) of each service flow. PDOR is defined as the ratio of packets that do not meet the delay threshold constraint (i.e. which

14. 3GPP, « 5G: Study on scenarios and requirements for next generation access technologies », *in: 3rd Generation Partnership Project (3GPP), TR 38.913 version 14.3.0 Release 14* (2017).

15. *Idem*, « Medium Access Control (MAC) Protocol Specification ».

16. Gueguen, C., « Opportunistic Energy Aware Scheduler for Wireless Networks », *in: IEEE Int. 77th Vehicular Technology Conference*, 2013, pp. 1–5.

17. Gueguen and Manini, *op. cit.*

18. Gueguen, C. and Baey, S., « Weighted Fair Opportunistic Scheduling for Multimedia QoS Support in Multiuser OFDM Wireless Networks », *in: IEEE Int. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2008.

are out of delay) compared to the total number of packets transmitted. In our proposal, until it is possible (i.e. as long as mobiles experience acceptable QoS and have a good PDOR), the focus is on energy minimization. If the QoS experienced by mobiles begins to decrease, FEC proportionally increases the focus on spectral efficiency in order to preserve it. Moreover, the solution is built in a cross layer approach that offers other advantages. Physical layer information allows to make opportunistic scheduling which ensures high system capacity. Higher layer parameters considered with the QoS metric of the FEC allow to provide, in addition, high fairness and QoS differentiation. The performance evaluation shows that FEC scheduler provides efficient global system energy consumption close to OEA without harming the system capacity that reaches MaxSNR performance. This is done while enabling good QoS differentiation and fairness close to the specialized scheduler on these metrics (e.g. WFO). This chapter is organized as follows. The first section introduces related works then, the two contributions on scheduling of this thesis are presented, FEC and AMUDS, respectively. This chapter ends with the conclusion.

3.2 Related work

Considering the number of scheduling solutions available in the literature, we provide a concise but representative study by using the most acknowledged schedulers, namely: Round Robin (RR) and Maximum Signal-to-Noise Ratio (MaxSNR) schedulers. We also provide more specialized schedulers, namely: Weighted Fair Opportunistic (WFO), Opportunistic Energy Aware (OEA) and Dynamic Trade-off (DT) schedulers. These five schedulers are representative of each specialization: non-opportunistic (RR), opportunistic (MaxSNR), QoS-oriented (WFO), energy consumption oriented (OEA), trade-off between spectral and energy efficiency (DT).

3.2.1 Round Robin (RR)

Original schedulers in wireless networks such as Round-Robin (RR)¹⁹ or Random Access (RA)²⁰ were inherited from wired networks or derived from operating systems. Although these solutions are somewhat outdated as they do not take into account radio conditions in their scheduling process, they continue to be studied for their ease of

19. Kuurne and Miettinen, *op. cit.*

20. Gokhan and Lang, *op. cit.*

implementation²¹ and used in numerous scheduler performance studies as a standard of comparison²².

RR scheduler is one of the most acknowledged non opportunistic schedulers. RR allocates resources to UEs one after another in a systematic and circular manner regardless radio conditions. This provides a poor spectral efficiency and a high UEs energy consumption as this mode of operation prevents from compressing the transmission time. Although the Round Robin may not take into account channel attenuation, it provides a certain degree of fairness as all the UEs have the same probability to access the radio channel at a very low computational cost. It is one of the simplest schedulers.

3.2.2 Maximum Signal-to-Noise Ratio (MaxSNR)

One of the most acknowledged opportunistic schedulers is the Maximum Signal-to-Noise Ratio (MaxSNR)²³ for its ability to increase the system capacity.

MaxSNR allocates a PRB n to a UE m among the set of k UEs such as:

$$m = \underset{k}{\operatorname{argmax}}(\eta_{k,n}) \quad (3.1)$$

This scheduler is focused on optimizing the spectral efficiency and increasing the system capacity. However, UEs far from their gNB experience a high path loss and have on average, a poorer SNR than UEs located in the inner of the cell. This leads MaxSNR to give priority to UEs close to their gNB, providing unfair allocations according to the distance between UEs and the gNB. UEs are to some extent, sorted and favored depending on their proximity to their gNB. The MaxSNR consequently, does not completely take benefits from the multi-user diversity and fail to provide an adequate QoS especially for cell-edge UEs with tight delay constraints.

3.2.3 Proportional Fair (PF)

Proportional Fair (PF) scheduler has been designed in order to solve drawbacks of MaxSNR²⁴. This solution gives priority to the UEs with the best ratio between short and long term throughput, leading to an enhanced spectral efficiency on long time scale

21. Prakash and Chaitali, *op. cit.*

22. Minelli et al., *op. cit.*

23. Tabatabaee and Tassiulas, *op. cit.*

24. Yao, *op. cit.*

and fairer allocations between inner (cell) UEs and cell-edge UEs. PF is still subject to improvement and various PF-based schedulers^{25 26 27} have been proposed since.

PF allocates a PRB n to an UE m among the set of k UEs such as

$$m = \underset{k}{\operatorname{argmax}} \left(\frac{\eta_{k,n}}{CF_k} \right) \quad (3.2)$$

Where CF_k is a compensation factor related to the distance of the UE to the gNB. In fact, most of the time, CF_k is the mean $\eta_{k,n}$. It means that the UE is selected not only when its radio conditions are great but above all, because its short term throughput is better than its average long term throughput. Unlike MaxSNR, PF does not segregate UEs according to their distance to the gNB which allows to use the multi-user diversity in its entirety. This leads to an enhanced spectral efficiency on the long term and a better fairness between UEs regarding the distance in comparison with MaxSNR.

3.2.4 Weighted Fair Opportunistic (WFO)

Weighted Fair Opportunistic (WFO)²⁸ is based on the PDOR which allow to ensure high fairness and service differentiation. It also outperforms other well-known fair opportunistic schedulers like PF or MaxSNR.

WFO allocates a PRB n to a UE m among the set of k UEs such as

$$m = \underset{k}{\operatorname{argmax}} \left(\frac{\eta_{k,n}}{f(PDOR_k)} \right) \quad (3.3)$$

with:

$$f(PDOR_k) = 1 + \beta * PDOR_k^\sigma. \quad (3.4)$$

It means that the UE is selected not only when its radio conditions are great but also because its PDOR is rising (meaning the QoS of UEs is decreasing).

25. Hamouda and Bouhlef, *op. cit.*

26. Ge, Jin, and Leung, *op. cit.*

27. Elhadad, M.I., El-Rabaie, M., and Abd-Elnaby, M., « Capacity enhanced scheduler for LTE Downlink System Based on PF Algorithm », *in: Fourth International Japan-Egypt Conference on Electronics, Communications and Computers (JEC-ECC)*, 2016, pp. 1–5.

28. Gueguen and Baey, *op. cit.*

3.2.5 Opportunistic Energy Aware (OEA)

The Opportunistic Energy Aware²⁹ is a scheduler that focuses on reducing the energy consumption of UEs. Despite being opportunistic by taking into account radio conditions, the OEA mainly aims to compress the transmission time by reducing the multi-user diversity in order to lower the energy consumed by UEs. This scheduler allocates resources to the UE with the best short term throughput and give a high priority to UEs already activated during the Time Slot (TS).

Let be A_k the mode of the UE k ($A_k = 1$ when the UE k is in active mode, while $A_k = 0$ when it is in sleep mode). The supplementary energy necessary to wake up the mobile k from the sleep mode to the active mode (i.e. to supply energy to the radio module) is denoted C_k . C_{nk} represents the energy required to receive and decode data on a n^{th} PRB (i.e. energy required for the antenna transmission). The amount of energy required to be active (to switch the radio module on) is higher than the energy needed to transmit on an additional resource unit if the mobile is already set in active mode. Consequently, $C_k > C_{nk}$. In the following of this work, C_k and C_{nk} values are considered respectively equal to 110.2 mW and 46.8 mW, for all k in accordance with measured hardware consumption³⁰. OEA allocates a PRB n to a UE m among the set of k UEs such as :

$$m = \operatorname{argmax}_k \left(\frac{\eta_{k,n}}{A_k * C_{nk} + (1 - A_k) * (C_k + C_{nk})} \right) \quad (3.5)$$

According to this formula, when the UE k is in active mode (i.e. $A_k = 1$), the energetic cost to transmit is only the cost C_{nk} . By contrast, when the UE k is in sleep mode (i.e. $A_k = 0$), the energetic cost to transmit is $C_k + C_{nk}$. An UE already set in active mode during a TS has a greater priority than UEs in sleep mode. A few number of UEs are thus awoken to transmit during the sub-frame. This allows to fully take benefits from the energy cost used to turn on radio module of the UE. However, the OEA will be compelled to slash the multi-user diversity and the few UE selected inescapably have sub-carriers on which they have a destructive multipath fading leading to unwise allocations in term of transmission efficiency. Consequently, the OEA reduces the energy consumption of UEs at the expense of a poorer QoS provided to UEs and a weaker spectral efficiency optimization than classical opportunistic schedulers.

29. Gueguen, *op. cit.*

30. Gueguen and Manini, *op. cit.*

3.2.6 Dynamic Trade-off (DT)

The Dynamic Trade-off (DT) scheduler has the ability to switch of behavior depending of the traffic load context and leads to a very efficient trade-off between energy consumption and spectral efficiency³¹

DT allocates a sub-carrier n to a UE m among the set of k UEs such as :

$$m = \operatorname{argmax}_k \left(\frac{\eta_{k,n}}{A_k * C_{nk} + (1 - A_k) * \left(\frac{C_k}{MD} + C_{nk} \right)} \right) \quad (3.6)$$

with:

$$MD_x = C + \beta x^\alpha \quad (3.7)$$

Where MD is a function of the bandwidth usage ratio x . This formula is really similar to OEA's principle but MD function allows to perform a trade-off between spectral and energy efficiencies according to the load of the system. At low traffic load, the value of MD is low meaning the cost C_k is dominant in the ratio so the cost to awake a new user is high. This leads to a trade-off between spectral and energy efficiencies. DT thus compresses the transmission time and awakes only few UEs per TS. On the contrary, when the system is loaded, the value of MD is high and the cost C_k to awake an UE is negligible. Therefore, the focus is put on only optimizing the spectral efficiency. Even though DT is really competitive compared to specialized schedulers (e.g. versus OEA at low traffic load or versus MaxSNR at high traffic load), this solution fails to provide service differentiation, full fairness and adequate QoS in many context particularly when mobile have different delay constraints or throughput requirements. In addition, DT adjusts its priority (energy versus system capacity) thanks to the system bandwidth usage ratio that does not take into account the significant importance of flow variabilities and application constraints. This conducts its dynamic trade-off to sometimes be inaccurate (triggered too early while no QoS problem are met or too late when QoS is already widely affected).

3.3 Scheduler to reduce energy consumption and increase system capacity

In this section, we propose a new scheduler named FEC that is based on weights that set the dynamic priorities in order to process to efficient radio resource allocation. These

31. *Ibid.*

weights are designed to achieve three important goals which are separately detailed in the following:

- Maximization of spectral efficiency and system capacity,
- Fairness guarantee,
- Minimization of energy consumption.

Then the algorithm flow chart and the merging of the previous weights in the balanced FEC solution are presented.

3.3.1 System description

In this section, the system description is explained by the introduction of the three important goals achieved by FEC.

System capacity maximization

A major objective of FEC scheduler is to optimize the system throughput. This is performed in a MAC/PHY opportunistic approach. Data communication reliability requirements of UEs are enforced considering each one independently, adapting the modulation and the transmit power to the UE specific channel state. During the allocation process at each frame, the radio resource allocation algorithm computes the maximum number of bits $q_{k,n}$ that can be potentially on PRB n if assigned to UE k while guaranteeing to respect its Bit Error Rate target (BER_k) such as:

$$q_{k,n} \leq \left\lfloor \log_2 \left(1 + \frac{3P \times T \times \left(\frac{1}{D_k}\right)^\gamma \times \alpha_{k,n}^2}{2N_0 \left[ERFC^{-1}\left(\frac{BER_k}{2}\right)\right]^2} \right) \right\rfloor, \quad (3.8)$$

with:

- P the power of the transmission,
- T the duration of an OFDM symbol,
- N_0 the spectral density of noise,
- D_k the distance between the gNB and UE k ,
- $\alpha_{k,n}^2$ the flat fading experienced by this UE on sub-carrier n ,
- erfc the complementary error function such as: $\text{erfc } z = 1 - \text{erf } z$ with erf z the error function (also called the Gauss error function)³².

32. Proakis, J.G. and Salehi, M., *Digital Communications*, New York: McGraw-Hill, 2007.

$\alpha_{k,n}$ is Rayleigh distributed such as its expectation is equal to 1. Due to multi-path fading, $\alpha_{k,n}^2$, the potential number of bits $q_{k,n}$ that a UE can send (or receive) on a radio resource unit fluctuates over the time. The exponent γ corresponds to the experienced path loss and shadowing. It generally varies from 2 to 4 considering environment density level. In the following, in order to consider dense urban context, γ is assumed equal to 3.5.

The supported modulation orders are assumed QAM and limited such as q belongs to the set $S = \{0, 2, 4, \dots, q_{max}\}$. Afterward, the maximum number of bits $\eta_{k,n}$ that can be transmitted on a PRB n for the UE k is:

$$\eta_{k,n} = \max \{q \in S, q \leq q_{k,n}\}. \quad (3.9)$$

Fairness guarantee

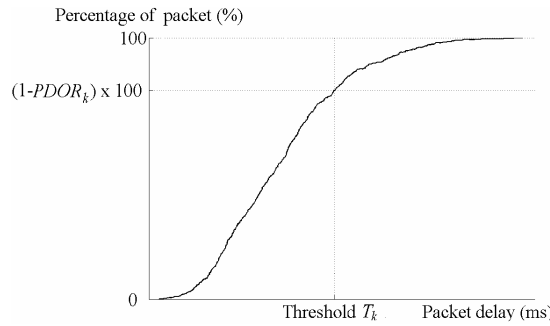


Figure 3.1 – Example packet delay CDF and experienced PDOR.

One of the major goals of the FEC scheduler design is to successfully support all multimedia transmission services, including a wide range of services such as VoIP, video-conference, file transfer, streaming and so on. This requires the coexistence of non real time traffic as well as delay sensitive flows with for example lower delay constraints but with tight data integrity targets. To be able to deal with the heterogeneous and various QoS requirements of multimedia services, the proposed FEC solution manages the QoS in a MAC/high-layers opportunistic approach. This relies on a generic approach of QoS management that implies that each traffic stream is considered as a service flow that has a specific QoS profile and possesses its own transmission buffer. Data integrity and delay which characterizes the QoS requirements of a service flow defined the QoS profile. In the following, we assume that one UE has one service flow. Data integrity requirements are specified by $BER_{target,k}$ for UE k . Data integrity can be considered as guaranteed by the physical layer by adapting the modulation scheme and the transmit power to the mobile

specific channel state. On the contrary, delay requirements are specified at the packet level and their achievements are directly impacted by the scheduler.

Clearly characterizing the delay requirements is necessary. We are convinced that the major constraint that can be considered in user communication is the limitation of the occurrences of large delay values. On this issue, the concept of delay outage can be applied. A delay outage is experienced by a packet of the service flow of UE k when its delay is greater than its application specific threshold denoted T_k . The Packet Delay Outage Ratio ($PDOR_k$) of each service flow of UE k is defined as the ratio of packets that do not meet the delay threshold constraint T_k (out of delay) compared to the total number of packets sent. Figure 3.1 shows an example of cumulative distribution of packet delay of a service flow at a given time instant. One of the objectives of the FEC scheduler is to regulate the experienced PDOR all along the lifetime of the service flow such as each PDOR value stays low and fairly distributed. This allows to ensure the satisfaction of the delay requirements at a short time scale as well as the service differentiation.

Energy consumption minimization

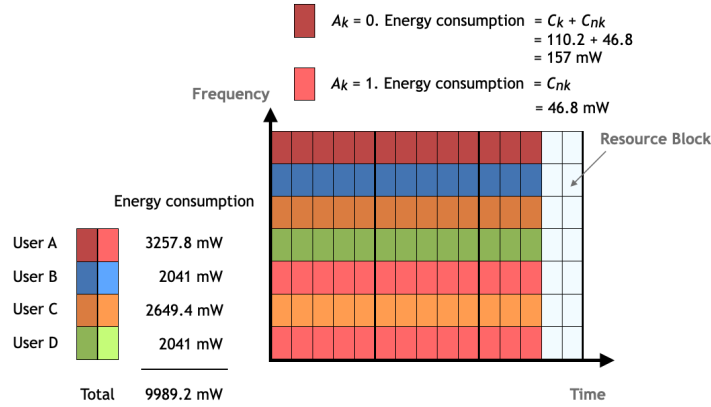
The last important goal of FEC is to provide efficient energy radio resource allocation in addition to the system capacity and fairness increase. Current opportunistic resource mapping (as MaxSNR, PF or WFO for instance) basically overexploits multi-user diversity. Since a flat fading is experienced during a frame on a given PRB³³, a same UE often experiences the greatest channel condition on this one. This induces that classical opportunistic schedulers often allocate the same set of sub-carriers to the same UE during the frame duration. In addition, the probability to have different UEs selected on different PRB is high. Consequently, during a same frame, many selected UEs can not be set in sleep mode for a low number of PRBs allocated³⁴, leading to "horizontal allocations".

Figure 3.2 is a conceptual representation of the PDCCH that shows each UE with their allocated PRB depending on the type of scheduler used (opportunistic or opportunistic but energy oriented).

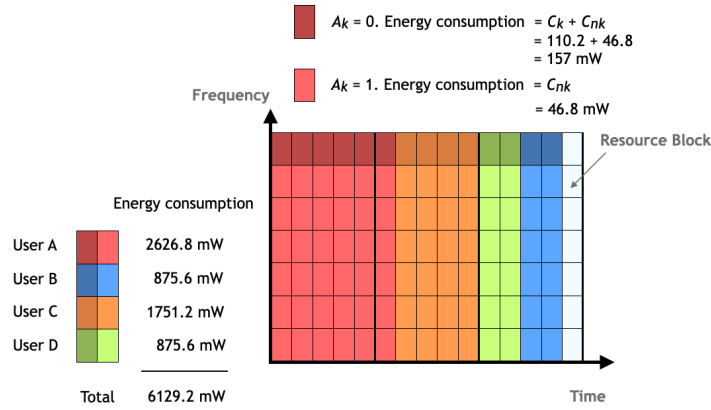
Figure 3.2(a) illustrates the resource allocation performed by classical opportunistic schedulers (such as MaxSNR). In this example, we consider that UE A (in red) has the

33. Truman, T.E. and Brodersen, R.W., « A Measurement-Based Characterization of the Time Variation of an Indoor Wireless Channel », in: *Proc. IEEE Int. Universal Personal Communications Record (ICUPC)*, 1997.

34. UEs may consume a lot of power to transmit/receive few bits during a long time (with many allocated PRB during the frame but few on the same Time Slot).



(a) Classical and horizontal opportunistic (e.g. MaxSNR) management of PRBs



(b) Vertical Opportunistic Energy Aware (OEA) management of PRBs

Figure 3.2 – Conceptual scheduling strategies

best radio channel conditions on the first PRB (first TS and first set of sub-carriers). Consequently, MaxSNR allocates this PRB to UE A. As it is the first time that this UE is awake during this TS, UE A has to turn on its radio module. Thus, its energy consumption depends of the awoken cost (C_k) and the cost related to data reception ($C_{n,k}$), represented by the dark red color. On the contrary, on the set of sub-carriers 6 and 7 of the first TS, UE A is already awake and its energy consumption is only related to the cost of data reception (represented by the light red color). As aforementioned, the multi path fading is assumed consistent for the frame duration. As a consequence, UE A obtain the same sets of sub-carriers during the frame. This leads to "horizontal allocations" where many different UEs are awake during the same TS allowing to take benefits from the multiuser diversity at the expense of a higher energy consumption.

On the contrary, Figure 3.2(b) illustrates OEA resource allocation behavior. As aforementioned, OEA compress the transmission time to reduce energy consumption of UEs. For instance, as UE A has the best radio channel conditions on the first PRB, all the next PRBs are allocated to UE A until no more data has to be received. This leads to "vertical allocations", allowing to reduce the energy consumption of UE at the expense of a lower spectral efficiency than classical opportunistic schedulers (due to a lower multiuser diversity usage, cf. Chapter 2).

The FEC scheduler integrates a modified version of the energy efficient OEA solution³⁵, keeping its energy benefits while correcting its lacks in terms of system capacity and fairness. To achieve this goal, FEC extends the classical OEA opportunistic cross-layer approach to obtain a new vertical opportunistic resource mapping. When a user is in active mode, FEC tries, like OEA, to take benefit from its activation to compress its duration of activity and to transmit more data per "used" time slots. Thus, FEC scheduler highly increases sleeping mode duration and energy preservation. Originally, OEA scheduler used an "Energy Transmission Cost" (ETC_k) parameter (in Watt). It is based on the energy cost of service flow of UE k to transmit on a PRB:

$$ETC_k = A_k * Cn_k + (1 - A_k) * (C_k + Cn_k). \quad (3.10)$$

The parameter ETC_k is used in OEA scheduler and highly contribute on its energy efficiency. However, it has the negative side effect to widely limit the usage done of the multi-user diversity. This makes OEA unable to ensure optimized system capacity. In order to keep OEA energy minimization qualities while fixing this throughput and fairness issues, FEC integrates a modified ETC_k parameter called "Context Aware Objectives Prioritization" parameter $CAOP_k$:

$$CAOP_k = A_k * Cn_k + (1 - A_k) * \left(\frac{C_k}{f(PDOR_k)} + Cn_k \right), \quad (3.11)$$

where $f(PDOR_k)$ is:

$$f(PDOR_k) = 1 + \beta * PDOR_k^\sigma. \quad (3.12)$$

The parameter σ allows having the adequate sensitivity and reactivity to PDOR fluctuations in order to ensure high fairness at a short time scale. Extended studies have already been done to calibrate this function. A value of σ equal to 3 ensures the best trade-off between average performances and short term fairness and it is the value considered in the following³⁶. In addition, β is a normalization parameter. It ensures that $CAOP_k$ and

35. Gueguen, *op. cit.*

36. Gueguen, C. and Baey, S., « A Fair Opportunistic Access Scheme for Multiuser OFDM Wireless

$\eta_{k,n}$ are in the same order of magnitude. Given that $PDOR_k$ has an order of magnitude 10^{-2} , β is set to $10^{2\sigma}$.

The higher $PDOR_k$ is, the more the system reduces $CAOP_k$ values and consequently the cost to activate a new UE. The number of active users will increase, intensifying the multi-user usage and consequently making the spectral efficiency higher at the expense of the energy consumption (maximum $PDOR_k$ value (100%) makes $CAOP_k$ constant and induces FEC similar to a MaxSNR resource allocation). At the opposite, low $PDOR_k$ values make FEC to decrease the number of active UE in a same time. This results in an energy consumption reduction at the expense of the multi-user diversity usage. This provides a radio resource allocation relatively similar to OEA resource allocation algorithm.

3.3.2 Contribution: Fair, Energy efficient and high system Capacity scheduler (FEC)

The contribution of this work is to present a new scheduler that combines all previously described parameters that also uses the dynamic $PDOR_k$ parameter to adapt priority to QoS experienced by users and consequently to the context. This section starts with an introduction of the mode of operation of the proposed solution and ends with an overhead and complexity discussion.

FEC merging of priorities

FEC scheduler allocates the PRB n to the UE k which has the greatest $FEC_{k,n}$ value such as:

$$FEC_{k,n} = \frac{\eta_{k,n}}{A_k * Cn_k + (1 - A_k) * \left(\frac{C_k}{f(PDOR_k)} + Cn_k \right)}. \quad (3.13)$$

As shown in Figure 3.3, the probability for a UE to receive PRBs is highly relied on the magnitude of its $FEC_{k,n}$. Considering $\eta_{k,n}$ allows to increase system capacity avoiding unprofitable radio resource allocation. Thanks to $f(PDOR_k)$ function, FEC also ensures service differentiation: if a UE experienced QoS decrease due to traffic burst, hard delay constraints or high BER requirements, it has proportionally more chance to be activated in the scheduling process with FEC. By adjusting the multi-user diversity usage thanks to

Networks », in: *Journal on Wireless Communications and Networking. European Association for Signal Processing (EURASIP). Special issue: Fairness in Radio Resource Management for Wireless Network* (Feb. 2009).

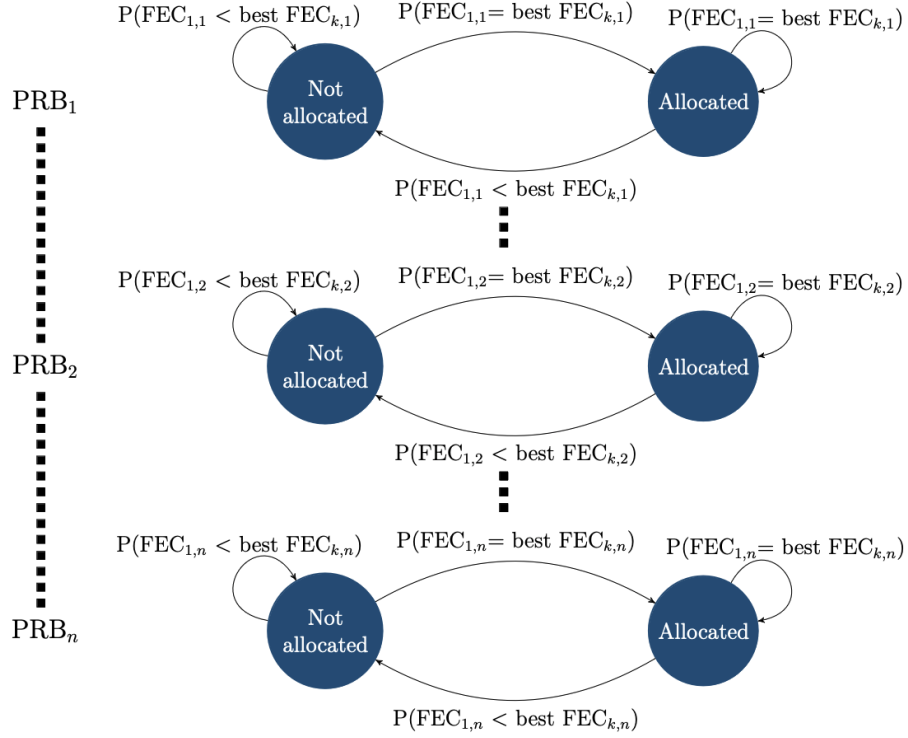


Figure 3.3 – Probability allocation of the Markov chain for UE 1 on PRB {1...n} associated to FEC.

a good function of $f(PDOR_k)$, FEC adequately adjusts the number of activated users per slot to the minimum possible in order to always make efficient energy management while respecting the QoS requirements. If the system experiences difficulties, FEC increases the multi-user diversity thanks to its ability to consider the raised values of $PDOR_k$ which urging it to obtain a better spectral efficiency. This avoids users QoS to decrease while guaranteeing a relative fairness and service differentiation. This also adequately improves the MUD usage of the scheduler and consequently the system capacity when required by the load context or by higher layer user application constraints. The other parameters allow to reduce energy waste.

FEC overhead and complexity discussion

FEC scheduler needs some inputs to work, such as:

- UE buffer occupancy,
- UE SNR (like MaxSNR, OEA, PF WFO and DT solution),
- User PDOR (like WFO),

- UE activity A_k (like OEA and DT).

Since buffer occupancy and SNR are used by all efficient solutions (MaxSNR, WFO, PF, DT...), FEC scheduler has no complexity addition on these points. PDOR is a simple integer to forward regularly to the gNB but this can be done at a larger scale than the scheduling period. Collecting this input is consequently feasible and does not provide significant overhead. A_k is directly determined by the scheduler in its decision process and occurs no overhead. Considering the algorithm complexity, FEC scheduling requires the sorting of $FEC_{k,n}$ for each UE k and all sub-carriers n as it is computed by MaxSNR with $\eta_{k,n}$. Consequently, the proposed FEC algorithm can be considered to have a complexity close to MaxSNR. The only supplementary complexity is the computation of each $FEC_{k,n}$ that requires 2 divisions, 2 multiplications and 2 additions more than MaxSNR but with integers (and sometime just with 0 and 1) that can be considered widely acceptable.

3.3.3 Performance evaluation

In this section, the proposed Fair Energy efficient with high system Capacity scheduler (FEC) is compared to 5 schedulers. They are well known in the literature and/or highly efficient in at least one objective that is in the scope of action of our proposed solution (i.e. spectral efficiency, energy consumption and QoS).

Round Robin (RR): This simple scheme allows to compare the proposed solution with a classical state of the art non-opportunistic scheduler³⁷.

Maximum Signal to Noise Ratio (MaxSNR): This opportunistic scheduler is widely acknowledged to its ability to widely increase the capacity of wireless networks³⁸.

Weighted Fair Opportunistic (WFO): This solution is able to ensure high fairness and service differentiation. It also outperforms other well-known fair opportunistic schedulers like Proportional Fair (i.e.³⁹).

Opportunistic Energy Aware (OEA): This user selection strategy is focused on energy consumption minimization⁴⁰. Contrary to T-MAC solution, it keeps some benefits of the multi-user diversity usage and represents a more challenging point of comparison (these solutions are compared in the publication of DT⁴¹).

Dynamic Trade-off (DT): This scheduler has the ability to switch of behavior

37. Minelli et al., *op. cit.*

38. Wong and Cheng, *op. cit.*

39. Gueguen and Baey, *op. cit.*

40. Gueguen, *op. cit.*

41. Gueguen and Manini, *op. cit.*

depending of the traffic load context and leads to a very efficient trade-off between energy consumption and spectral efficiency⁴².

Discrete event simulations are used to obtain performance evaluation. The global simulation setup is given in table 3.1:

Table 3.1 – Global simulation setup.

Parameters	Values
Number of sub-carriers n_{max}	32
Number of time slots t_{max}	10
Wake up cost C_k	110.2 <i>mW</i>
Transmission cost C_k	46.8 <i>mW</i>
Target BER	10^{-3}
Average $\alpha_{k,n}$ value	1
Sub-carrier spacing	15 <i>kHz</i>
Cell radius	500 <i>m</i>
Simulation duration	$5 \cdot 10^5$ frames

It is also considered that all UEs run realistic Variable Bit Rate applications that generates high volume of data with important sporadic and tight delay requirements which significantly complicates the task of the resource allocation algorithm. Each UE has a traffic composed of an video and voice streams⁴³ but with different required throughputs and/or delay constraints according to the considered scenario.

Scenarios and Key Performance Indicators

Four simulation scenarii are shown in the presented performance evaluation. First, the behavior of the schedulers when UEs occupy different geographical positions are analyzed. The second scenario underlines the performance of the resource allocation algorithms when UEs have heterogeneous bit rate requirements. In the third scenario, QoS differentiation concerning applications with different delay requirements is evaluated. The fourth simulation scenario gather the previous ones and considers UEs with all three: heterogeneous geographical positions, bit rate and delay requirements.

42. *Ibid.*

43. Tanwir, S. and Perros, H., « A Survey of VBR Video Traffic Models », *in: IEEE Communications Surveys Tutorials*, 2013; Tamimi, A.K.A., Jain, R., and So-In, C., « Modeling and Prediction of High Definition Video Traffic: A Real-World Case Study », *in: Second International Conferences on Advances in Multimedia*, 2010; Heyman, D.P., « The GBAR Source Model for VBR Videoconferences », *in: IEEE/ACM Transactions on Networking* (1997), pp. 554–560.

This study focuses on six KPI to evaluate each solution:

- **The spectral efficiency** is the mean number of bits received on each PRB used (bits/PRB).
- **The bandwidth usage ratio** is the ratio between the number of RUs used by a solution and the total number of RUs available (in %). When this ratio reaches 100% it means that the system is congested.
- **The mean energy consumption** per UE is the mean energy (in mW) that a UE consumes on average.
- **The mean PDOR** is the mean PDOR of a UE (in %).
- **The mean packet delay** is the mean delay to transmit one packet (in ms).
- **The Jain's fairness index** is the the well known Jain's fairness index It is computed on PDOR metric.

Scenario 1

It is widely acknowledged that, in wireless networks context, the UEs close to the access point usually obtain better QoS than UEs more distant due to their higher mean spectral efficiency. In this scenario, a first half of UEs is situated close to the access point and the second half is twice farther from the access point to study the impact of heterogeneous distance on the scheduling performances (fairness behavior, capacity to optimize the usage of the multi-user diversity to increase system capacity, etc.). The other parameters are fixed for all the UEs and described in Table 3.2. In addition, performance evaluations, on each criterion, are carried out studying the influence of the traffic load (underlining the different scheduler behaviors in low traffic load and high traffic load contexts). The network traffic load increases in our simulations by adding users 2 by 2 (each time, 1 close user and 1 far user).

Table 3.2 – First scenario setup.

Group	Distance (d_k)	Delay threshold (T_k)	Data rate (D_k)
1	100 m	80 ms	150 Kbps
2	200 m	80 ms	150 Kbps

Spectral efficiency and throughput. Figure 3.4(a) shows, for each scheduler, the spectral efficiency obtained with different traffic load in the system. Since RR is not opportunistic, it does not take any benefit of multi-user diversity and its spectral

efficiency is constant⁴⁴ and low. Energy focused scheduler (OEA), significantly limits the multi-user diversity usage in their allocation process offering lightly better results. At the opposite, MaxSNR is highly opportunist and grants an important spectral efficiency gain. However, as explained in Section 3.3.1, MaxSNR has a severe lack on fairness and is not able to take all the advantages of the multi-user diversity and is slightly outperformed in highly loaded context by WFO and DT that are fairer⁴⁵.

Thanks to its dynamic CAOP parameter based on the PDOR values, FEC has a lesser spectral efficiency in low traffic load context using a moderate usage of the multi-user diversity concentrating its efforts on energy (Fig. 3.4(a)). Nonetheless, when it appears necessary (when energy focused schedulers approach congestion (Fig. 3.4(b))), its CAOP factor adequately increases. This induces FEC raises its usage of the multiuser diversity which improves the spectral efficiency at the same level than MaxSNR (allowing to reach the same high overall system capacity limit (Fig. 3.4(b))).

Energy consumption.

The abilities of each scheduler to be more or less energy efficient can be studied thanks to the analysis of Figure 3.4(c). Clearly, the Round Robin scheduler is the solution that provides the worst results. With RR that is not opportunistic, a non negligible amount of PRBs are allocated with a poor spectral efficiency profitability (i.e. in term of bits per PRB). This results in high energy and PRB waste. In addition, RR allocation pattern is regular and cyclic. Consequently, with RR, many UEs are simultaneously activated and this exacerbates the energy waste since more UEs have to pay the high radio transmission activation price C_k . Focusing on the RR curves, we can notice that with more than 20 UEs, the RR curve increases more slowly. The explanation is that after this threshold, the cell is overloaded. RR does not success to provide enough PRBs to each user to reach their application requirements. Due to the lack of radio resource unit, UEs are more often obliged to wait a long time in sleep mode even with data to transmit. Consequently, each user consumes in overall less energy over the time after the system congestion threshold is crossed.

Another approach is OEA resource allocation strategy. Its concept is to drastically

44. When system congestion occurs, RR spectral efficiency slightly increases due to the fact that close UEs pick up the PRB previously allocated to far UEs in order to achieved an equal ratio of 50% assigned to each group.

45. However when WFO reaches its provided system capacity limit (24 UEs), it tries in vain to help far UEs (increasing their priority) which reduce its ability to take full benefit of the multiuser diversity usage. In addition, in highly overloaded system (more than 30 UEs), all UEs' PDOR are equal to 1 and WFO scheduling becomes similar to MaxSNR which explain a same spectral efficiency

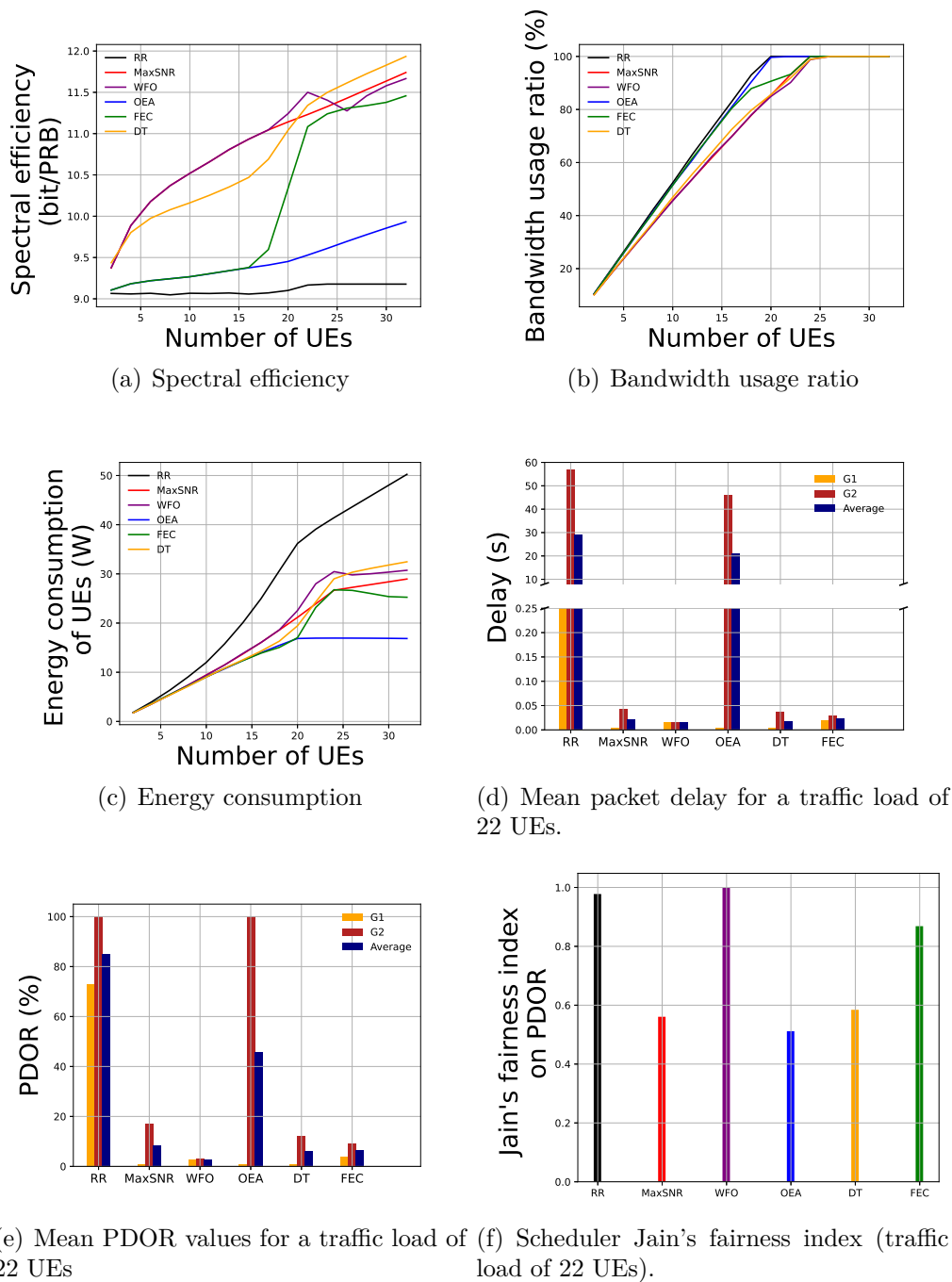


Figure 3.4 – Scenario 1 (distance influence on the schedulers performances)

limit the usage of the multi-user diversity to a low value. This induces that this solution ensures low energy consumption whatever the traffic load considered. However, this

result must be put into perspective since OEA keeps to focus on energy consumption minimization even when traffic loads is high. This stubborn radio resource management leads OEA to quickly reach congestion (Fig. 3.4(b)) and to provide high delay occurrences (Fig. 3.4(d)).

Contrary to OEA, WFO fully exploits the multiuser diversity thanks to its extended opportunistic approach and extends the system capacity (Fig. 3.4(b)). Not focusing on energy consumption, WFO is outperformed on this criterion by OEA. However, it outperforms RR (Fig. 3.4(c)) thanks to a strongly better spectral efficiency (Fig. 3.4(a)). About WFO performance, it is important to notice that when the system is overloaded (after 24 UEs (Fig. 3.4(b))), WFO unfruitfully tries to ensure fairness helping far UEs (increasing their priority). Activating less often close UEs and more the far UEs, the number of activated UEs is reduced and consequently more UEs are set in sleep mode which stabilizes WFO energy consumption. Regarding MaxSNR, we can notice that MaxSNR has slightly better energy results than WFO and particularly when system is overloaded. Indeed, when the system is overloaded, this scheduler has a propensity to segregate a part of the UEs (far from the gNB) and consequently it obtains reduced benefits of multiuser diversity usage. This is a weakness in term of spectral efficiency and fairness but an advantage to increase user sleep duration and reduce energy consumption.

Considering underloaded contexts (i.e. number of UEs inferior to 18 in this scenario), due to large excess of available radio resource units (Fig. 3.4(b)), guaranteeing high QoS is easily achievable by DT and FEC. Thanks to their system of weights that dynamically adjust the objectives priority to the context, they consequently focus their priorities on energy (Fig. 3.4(c)) rather than system throughput (Fig. 3.4(a)). However, considering the network traffic load, DT does not fully optimizes the trade-off. FEC, thanks to its CAOP parameter based on QoE measured is more accurate. It outperforms DT and offers energy gain equal to OEA until this stay possible.

Considering highly loaded context (number of UEs between 18 and 24), the lack of available PRBs (Fig. 3.4(b)) requires schedulers to focus more on system capacity than energy consumption which then becomes a lower priority in order to preserve QoS. In this context, DT and FEC decide to slightly sacrifice energy in order to sustain the network viability and then favor high spectral efficiency proportionally to the traffic load with DT and proportionally to UEs QoE difficulties with FEC. However, performing the energy-throughput transition according to the evolution of the traffic load is not accurate and leads the DT to make its transition too early in this scenario. Thanks to

is CAOP parameter based on PDOR, FEC begins its energy-throughput transition just when difficulties really occur and provides greater and longer energy gain (Fig. 4.4(d)).

In overloaded context (number of UEs superior to 24), since all UEs experience high delay and consequently high PDOR, FEC scheduling becomes close to MaxSNR which is close to be the best strategy in this specific case. Consequently, these schedulers offer slightly the same results in term of spectral efficiency (Fig. 3.4(a)), system capacity (Fig. 3.4(b)) and energy (Fig. 3.4(c)).

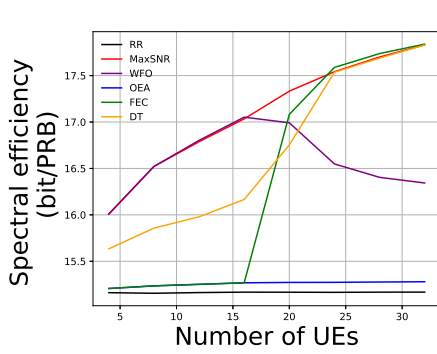
Delay and fairness. An important QoS KPI is the packet latency. Figure 3.4(d) shows the mean packet delay in the system in milliseconds and figure 3.4(e) the mean PDOR when system load is high (26 UEs) which potentially makes UEs experience some difficulties. Note that 2 groups emerged:

- First, RR and OEA that have the worst results since they quickly experienced congestion. Since they had a low spectral efficiency (Fig. 3.4(a)), they do not success to support a significant amount of traffic load with acceptable QoS.
- Secondary, MaxSNR, WFO, DT and FEC (each one being opportunistic algorithm solution) are able to better endure higher load increase with satisfactory delay and PDOR.

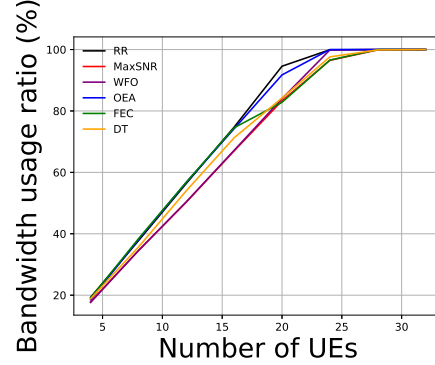
Concerning fairness about distance, results show (Fig. 3.4(d), 3.4(e) and 3.4(f)) that RR, OEA, MaxSNR and DT significantly penalize UEs far from the gNB (group 2). OEA and MaxSNR are the most unfair due to their pure opportunistic approach that blindly favors UEs with good SNR. RR and DT ensure a same amount of PRBs that slightly improves this behavior without completely solving the problem. Indeed, giving each UE the same amount of PRBs does not mean to allow them to transmit the same amount of bits as it depends on their position (due to the path loss attenuation on SNR). WFO is conceived in order to ensure high fairness and consequently offers the best results on this metric. However, the proposed FEC scheduler is close to WFO results (between WFO and DT) ensuring good fairness (Fig. 3.4(f)) with low delay and PDOR values for all groups as well as high system capacity. However, this is achieved by the FEC with, in addition, high energy gain.

Scenario 2

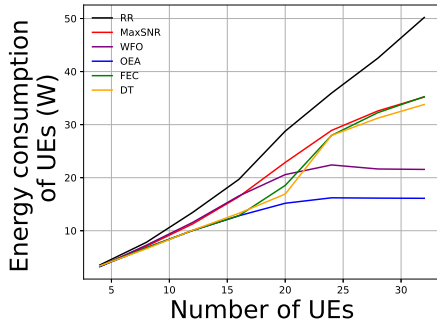
Here, UEs are divided in two groups that differ only by their data rate requirements as described in Table 3.3. This allows to study the ability of each scheduler to manage mobiles using applications with various data rate. Results are shown in figure 3.5.



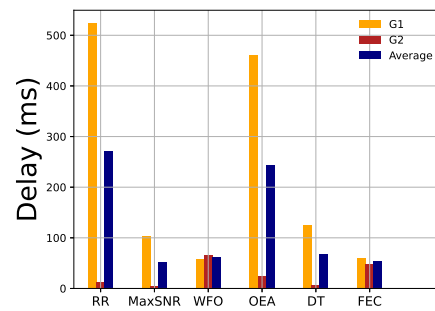
(a) Spectral efficiency.



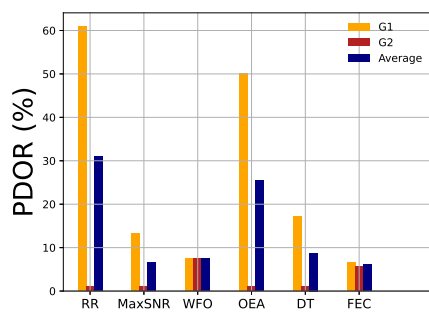
(b) Bandwidth usage ratio.



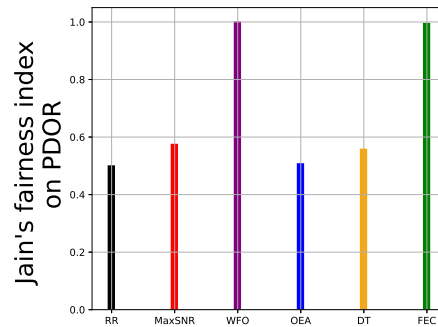
(c) Energy consumption.



(d) Mean packet delay for a traffic load of 24 users.



(e) Mean PDOR values for a traffic load of 24 users.



(f) Scheduler Jain's fairness index (traffic load of 24 users).

Figure 3.5 – Scenario 2 (performance with heterogeneous bit rate sources).

Table 3.3 – Second scenario setup.

Group	Distance (d_k)	Delay threshold (T_k)	Data rate (D_k)
1	100 m	250 ms	300 Kbps
2	100 m	250 ms	100 Kbps

As expected, RR provides the worst performances due to its non opportunistic PRBs management. Much energy is wasted, system capacity limit is low and users that require high throughput fail to meet their delay requirements. Limiting the usage of multiuser diversity, OEA offers large energy gain but provides poor system capacity (slightly superior than with RR), QoS and fairness. On the contrary, with MaxSNR, spectral efficiency is increased at the expense of energy and fairness. WFO, which is able to ensure service differentiation, successfully corrects this lack of fairness while staying close to MaxSNR other performances (on spectral efficiency⁴⁶ and system capacity). Thanks to their abilities to adjust the objectives priority according to the context, DT and FEC provide low energy consumption in underloaded context (Fig. 3.5(c)) since UEs satisfaction is easily met. When traffic load increases, they make their energy-throughput transition in order to raise the system capacity (Fig. 3.5(a)) and hold QoS requirement⁴⁷. However, since DT is unable to ensure service differentiation, it fails to guarantee high fairness and UEs of group 1 experience difficulties to meet their QoS requirements. On the contrary, FEC succeeds to ensure an high fairness (Fig. 3.5(f)) and high QoS (Fig. 3.5(d) and 3.5(e)) and provides performances close to WFO for all groups.

Scenario 3

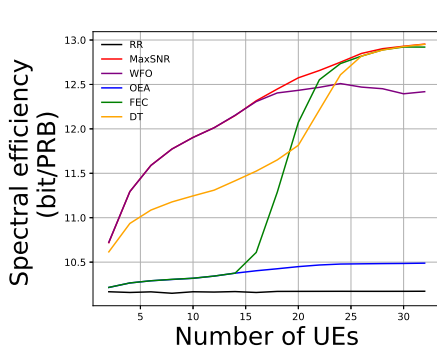
In this scenario the influence of heterogeneous delay requirements on the scheduling performances is studied. In this simulation scenario, UEs are split into two groups that differ only by their delay requirements (cf. Table 3.4).

Results are shown in Figure 3.6 and underline the same tendency for each schedulers than in Scenario 2. However, they allow some new analysis/discussions:

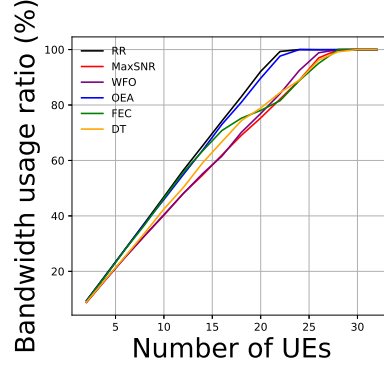
- DT triggers its “energy-throughput” transition based on system capacity traffic

⁴⁶. In overloaded context the spectral efficiency of WFO decreases since UEs with high throughput requirements experience huge difficulties to ensure their QoS, giving them disproportionate priority. This makes WFO using a truncated multiuser diversity. This decreases its spectral efficiency but also its global energy consumption.

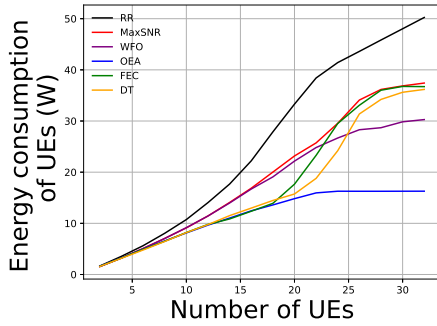
⁴⁷. In overloaded context DT offers the same spectral efficiency than MaxSNR since, in this scenario, all UEs are at the same distance from the access point and no supplementary of multiuser diversity can be used by DT compared to MaxSNR (contrary to the scenario 1).



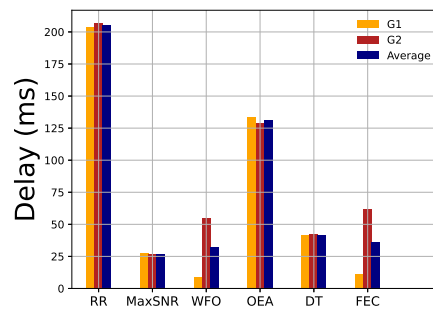
(a) Spectral efficiency.



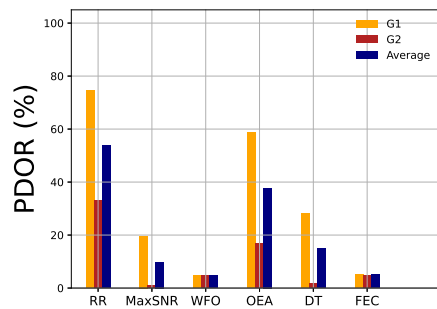
(b) Bandwidth usage ratio



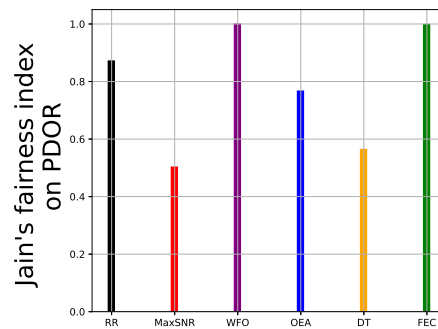
(c) Energy consumption.



(d) Mean packet delay for a traffic load of 20 UEs



(e) Mean PDOR values for a traffic load of 20 UEs



(f) Scheduler Jain's fairness index (traffic load of 20 UEs)

Figure 3.6 – Scenario 3 (performance with heterogeneous delay constraints)

Table 3.4 – Third scenario setup.

Group	Distance (d_k)	Delay threshold (T_k)	Data rate (D_k)
1	100 m	50 ms	150 Kbps
2	100 m	250 ms	150 Kbps

- load measures. This information, although interesting, is not accurate and makes DT to initiate its transition too late since UEs experienced difficulties to ensure their QoS requirements since a while (contrary to the scenario 1 where it was too early). Consequently, in this scenario and like OEA scheduler, DT offers better energy consumption than FEC (Fig. 3.6(c)) but at the expense of a higher PDOR (Fig. 3.6(e)). In addition, FEC offers strong service differentiation close to WFO results and widely outperforms DT offering a same PDOR level for all groups whatever their delay constraints are (Fig. 3.6(e) and 3.6(f)).
- This scenario underlines that many un-equivalent levels of fairness can be offered by a scheduler to its UEs. The basic approach is to consider as fair an allocation of a same number of PRBs to each UEs (like with RR) but this does not guarantee a same throughput and is far to be a relevant definition. More evolved approaches consider fair to offer a same packet delay to all UEs. However, the greatest level of fairness can be considered as to ensure to all UEs a same degree of satisfaction whatever their distances, application throughput requirements or delay constraints. Consequently, in this scenario, ensuring high fairness between UEs with different delay requirements is not equal to ensure to all UEs a same packet delay like provided by MaxSNR and DT (Fig. 3.6(d)). In this scenario, the objectives are to guarantee a packet delay for the user of group 1 inferior to 50 ms and for UEs of group 2 inferior to 250 ms. Fairness can be adequately measured as providing a same ratio of packets in delay outage (i.e. a same PDOR) for all groups. Following this analysis, only WFO and FEC schedulers are able to offer a good fairness in this scenario (Fig. 3.6(e) and Fig. 3.6(f)).

Scenario 4

Previously, the behavior of each scheduler has been studied in simple contexts considering one criterion at a time to better understanding their influences on the performances. In order to compile these ones, this section shows a study of the performance of the evaluated protocols in a more general context. 8 heterogeneous groups of mobiles are considered

here as described in Table 3.5.

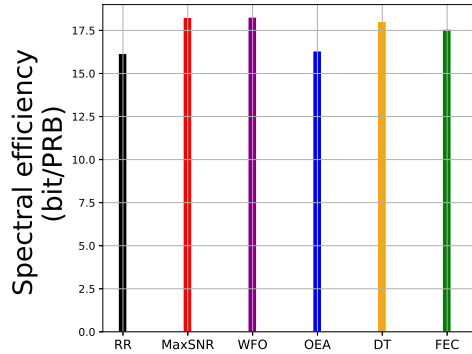
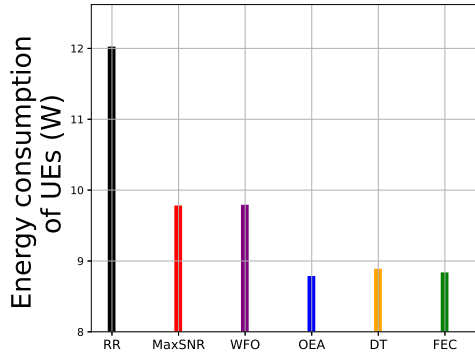
Table 3.5 – Fourth scenario setup.

Group	Distance (d_k)	Delay threshold (T_k)	Data rate (D_k)
1	50 m	250 ms	100 Kbps
2	100 m	250 ms	100 Kbps
3	50 m	250 ms	300 Kbps
4	100 m	250 ms	300 Kbps
5	50 m	80 ms	300 Kbps
6	100 m	80 ms	300 Kbps
7	50 m	80 ms	100 Kbps
8	100 m	80 ms	100 Kbps

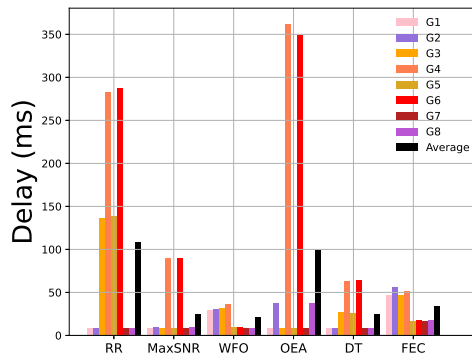
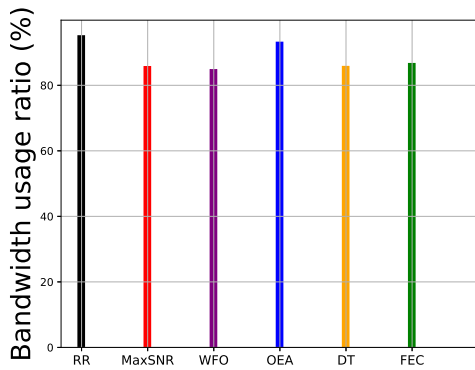
Figure 3.7(a) shows the global energy consumed with each scheduler when traffic load represents around 50% of the system capacity limit. As expected, RR is the most energy greedy since this resource allocation provides inefficient spectral efficiency and horizontal resource allocation which induces that UEs are activated on many time slots to transmit few bits. MaxSNR and WFO offer better results due to their opportunistic behavior which provide better spectral efficiency. However, they are respectively widely outperformed by DT, FEC and OEA that try to have a more vertical resource allocation which highly compress user time activity and consequently significantly increase user time sleep duration.

Figure 3.7(b) and 3.7(c) respectively show the spectral efficiency and bandwidth usage ratio for each scheduler in high traffic load context. RR provides, again, poor results since it does not take into account the wireless transmission conditions in its resource allocation process. Too focused on energy objective, OEA fails to ensure significant improvement on these metrics. At the opposite, MaxSNR and WFO that overexploit the multiuser diversity provide huge benefits: high spectral efficiency (Figure 3.7(b)) and low bandwidth usage ratio (i.e. high system capacity limit (Figure 3.7(c))). DT and FEC are doing their "energy to system capacity" transition in order to support the traffic load growth and reach close results to MaxSNR.

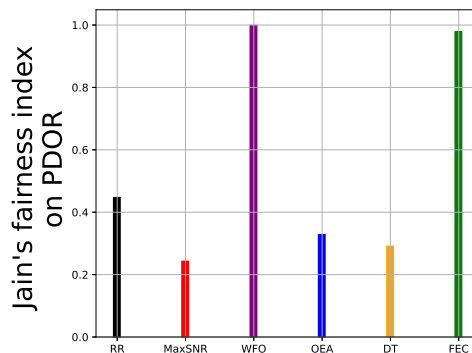
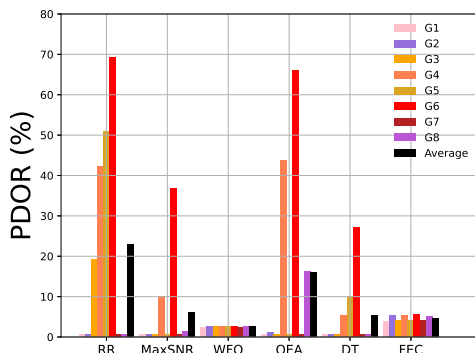
Figure 3.7(d), 3.7(e) and 3.7(f) highlight the inabilities of RR, MaxSNR, OEA and DT to provide service differentiation. When delay requirements are stringent and/or the path loss is high and/or the source bit rate is raised, mobiles experienced severe difficulties (particularly the group 6). On the contrary, WFO and FEC success to differentiate UEs according to their profile and position. They guarantee to each group the respect of



(a) Energy consumption for a traffic load of 16 users (b) Spectral efficiency for a traffic load of 32 users



(c) Bandwidth usage ratio for a traffic load of 32 users (d) Mean packet delay for a traffic load of 32 users



(e) Mean PDOR values for a traffic load of 32 users (f) Scheduler Jain's fairness index (traffic load of 32 users)

Figure 3.7 – Scenario 4 (global scheduling performances analysis)

their heterogeneous delay requirements (Fig. 3.7(d)), withdrawing negative impact of distance (Groups 1, 3, 5, 7 are no more penalized) and consequently offer a same amount of PDOR/QoE to each group (Fig. 3.7(e)) as well as a good fairness (Fig. 3.7(f)).

3.3.4 Conclusion

Jointly guaranteeing high system capacity, high fairness, high QoE and low system energy consumption is very challenging in wireless networks. Specialized solutions as MaxSNR, WFO or OEA have been well designed to meet one of these criteria but by failing others. Alternative works (like DT) rely on the usage of tradeoffs but provide un-optimized/averaged performance. This work proposes a new paradigm: the magnitude of importance of these objectives could vary according to the context. In very low traffic load context, radio resource units are abundant and energy minimization should be the main objective. With the rise of the network traffic load, more focus/priority should be done on spectral efficiency and system capacity increase in an adequate "energy-throughput" management. In high traffic load, the spectral efficiency increase should become the primary objective in order to continue to satisfy the maximum number of users and energy minimization priority must be relegated. In this work, a new access scheme called FEC scheduler has been proposed. It allows to tune users' priorities and the multiuser usage benefit according to users' requirements and networks difficulties. Performance evaluation shows that the proposed solution provides together a level of fairness, energy efficiency and system capacity close to the acknowledged specialized schedulers on these 3 major performance indicators: without penalizing the system capacity (reaching almost the same spectral efficiency than MaxSNR), it provides efficient global system energy consumption (very close to OEA energy specialized scheduler) while ensuring high level of user satisfaction, QoS differentiation and fairness (close to WFO fair specialized scheduler). This results in a scheduler that widely outperforms existing solution trying to ensure these 3 KPIs together as DT solution.

3.4 Meta scheduler to reduce UE's energy consumption

Based on the results obtained in Chapter 2 and in the previous section, the principle of varying the MUD to increase/decrease the spectral and energy efficiencies can be extended

to create a meta-scheduler.

To the best of our knowledge, the denomination of 'meta-scheduler' in wireless networks, has been used once and refers to choose the most suitable scheduler (among a set of defined schedulers) according to the context⁴⁸. In this work, we define the meta-scheduler as a fully fledged process of the resource allocation chain, acting just before the classical scheduling. Rather than designing a scheduler solving a specific sub-objective, the energy efficiency concern is grasped in a more comprehensive way by the proposal of a meta-scheduler compatible with the most acknowledged schedulers. The Adaptive Multi User Diversity meta Scheduler (AMUDS) relies on the same analysis than FEC: when the system is under loaded, delay requirements of UEs are easily reachable and optimizing the spectral efficiency is not necessary. Then, the focus could be wisely put on the minimization of the energy consumption. As the traffic load rises, resources tend to become scarce and limited and enhancing the transmission efficiency is required since QoS of UEs must be ensured.

The novelty of AMUDS relies on the way that these goals are achieved. Unlike FEC, AMUDS is not a scheduler but directly influences the opportunistic behavior of the most acknowledged schedulers either by increasing or decreasing their multi-user diversity usage depending on the cell traffic load. The proposed solution is thus, able to control and monitor the number of active UEs per TS allowing to reduce the energy consumption or to enhance the spectral efficiency according to the context. One of the main benefits of the proposed solution is its ability to be compatible and to cooperate with the most acknowledged schedulers. By extending the principle of compressing the transmission time to any of these schedulers (such as: DRX, OEA, FEC, DT), AMUDS is able to provide energy efficient solutions even on non energy-focused schedulers. As AMUDS only slightly modifies these scheduler behaviors, they keep their intrinsic properties and provide the same performance when the traffic load rises. The proposed solution is also easier to implement for a network operator than a whole scheduler focused on a single priority and can be adapted with ease to the already existing solution.

The main contributions of this work are the following:

- **Meta-scheduler:** proposal of a new element of the resource allocation chain compatible with existing schedulers to address the challenge of reducing the energy consumption of UEs.

48. Song, J., Veciana, G., and Shakkottai, S., « Meta-Scheduling for the Wireless Downlink through Learning with Bandit Feedback », in: *18th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, 2020.

- **Complexity guarantee:** The proposed solution requires no additional information than classical schedulers, namely: Channel Quality Indicator (CQI) to adapt the modulation and user data to compute the transport block size and the number of radio resources allocated. Regarding the additional computation, only a delay derivative function is used to act as a simple hysteresis function.
- **Simulation results:** AMUDS has been tested under 2 main scenarios. The first scenario is quite simple which allows to have a fully understanding of the proposed solution while the second scenario tests AMUDS with different positions of UEs in the cell.

3.4.1 System Description

This study focuses on the radio resource allocation problem for the set of UEs located in the coverage zone of a gNB. A centralized approach is assumed since this allows efficient opportunistic scheduling approaches. The physical layer is considered to operate using a Time Division Duplex mode (TDD) which allows a good compatibility with the OFDM based transmission mode⁴⁹. The global available bandwidth is split into sub-frequency bands called sub-carriers. Radio resources are distributed in time domain in frames themselves split in Time Slots (TS) of constant duration (which is an integer multiple of the OFDM symbol duration). Each frame duration is assumed equal to a value inferior to the coherence time of the channel, allowing transmission on each sub-frequency to experience flat fading during each frame. Full knowledge of radio conditions is supposed to be available at the receiver⁵⁰. Thanks to SNR measurement of the signal sent by each UE (for instance during the uplink contention subframe), the gNB is thus, able to estimate their channel state attenuation at a given TS on each sub-carrier. According to⁵¹, channel state can be assumed stable on a scale of 50 ms. This requires that UEs have to transmit their control information alternatively on each sub-carrier once every frames (in order for the gNB to successfully refresh the channel state information). A PRB, defined as a (15 sub-carriers, time slot) pair, can be allocated to any UEs with a specific modulation order. Transmissions performed on different PRBs by different UEs have independent channel state variations. On each PRB, a modulation order adapted to the channel state (between the gNB and the selected mobile) is assumed. This provides the flexible resource

49. Kela, Turkka, and Costa, *op. cit.*

50. Li, Seshadri, and Ariyavisitakul, *op. cit.*

51. Truman and Brodersen, *op. cit.*

allocation allowing opportunistic scheduling.

The channel gain between the gNB and the UE k on the sub-carrier n is given by:

$$G_{k,n} = h 10^{\frac{X\sigma}{10}} \left(\frac{d_{ref}}{d_k} \right)^\alpha \quad (3.14)$$

where h represents the Rayleigh multi-path fading, which is modeled by an exponential distribution, X is a standard Gaussian random variable, σ is the standard derivation of shadowing in dB, d_{ref} is the reference distance, d_k is the distance between UE k and gNB while α is path loss exponent. The SNR computation of UE k on sub-carrier n associated to gNB is given by:

$$\gamma_{k,n} = \frac{P_n G_{k,n}}{B_{sub} N_0} \quad (3.15)$$

where parameter P_n is the transmitted power on sub-carrier n of gNB. Parameter N_0 is the thermal noise power density and parameter B_{sub} is the sub-carrier spacing⁵². To compute the spectral efficiency $\eta_{k,n}$ of UE k on sub-carrier n associated to gNB, the Shannon's formula is used such as:

$$\eta_{k,n} = \log_2 \left(1 + \frac{\gamma_{k,n}}{\Gamma} \right) \quad (3.16)$$

where parameter Γ is a SNR correction factor that takes into account the difference between the information-theoretic performances and the practical implementation of the MCS⁵³ defined as follows:

$$\Gamma = -\frac{\ln(5E)}{1.5} \quad (3.17)$$

where E is a BER Target.

3.4.2 Contribution : Adaptative Multi-User Diversity Meta Scheduler (AMUDS)

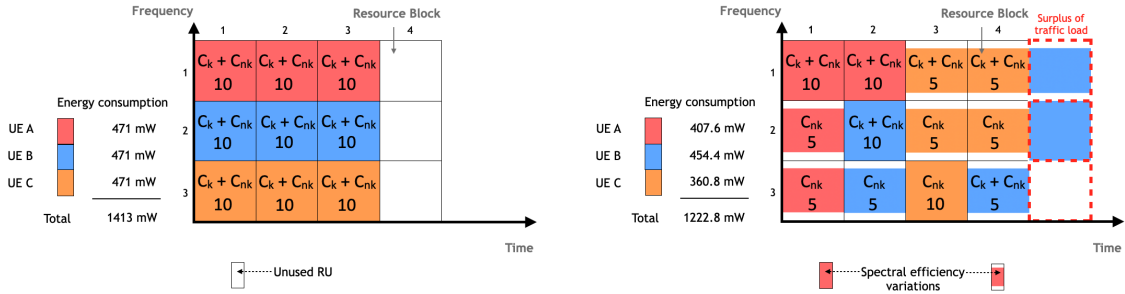
The Adaptative Multi User Diversity meta Scheduler has mainly two objectives:

- To reduce the energy consumption of UEs without harming their QoS nor the system capacity.
- To be compatible with the most acknowledged schedulers while keeping their intrinsic properties such as system capacity maximization, fairness and so on.

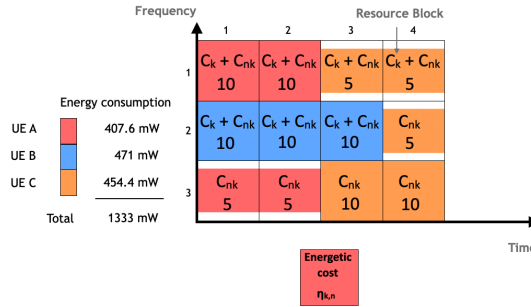
52. Ezzaouia, *op. cit.*

53. Seo and Lee, *op. cit.*

To achieve these goals, the AMUDS aims to coordinate with the scheduler of the gNB by influencing its multi-user diversity usage. The modification of this parameter has significant impacts on the performance provided by a scheduler and more generally by the network. This can be particularly seen on UEs energy consumption, system capacity, spectral efficiency and QoS of UE (for more details cf. Section 2.3).



(a) $MUD = 3$: good spectral efficiency and QoS (b) $MUD = 1$: good energy consumption but some PRBs are unused leading to a sub optimal energy consumption



(c) $MUD = 2$: wise bandwidth usage leading to an accurate trade-off between energy consumption and spectral efficiency

Figure 3.8 – Illustration of MaxSNR allocations for different MUD values.

According to the results of the previous section and section 2.3, dynamically restraining the multi-user diversity according to the traffic load of the cell seems to be appropriate and efficient. However, considering that traffic load of UEs highly varies over the time, designing a solution only based on the average traffic load is sub-optimal and inaccurate. Indeed, it does not allow to take into account traffic load spikes induced by the UEs' heterogeneous needs and applications at a given time. The proposed solution is hence, based on a multi-user diversity parameter updated at the beginning of each frame according to the estimated traffic load for this given frame. This parameter, denoted MUD , is

computed before the resource allocation process in order to restrain or increase the use of the multi-user diversity.

To determine the appropriate value of MUD parameter at a given time is the crucial issue of the proposed solution as different MUD values can induce opposite performance. The figure 3.8 attempts to bring this phenomena to light with MaxSNR allocation depending on three MUD values. For this example, it is considered that three UEs are in the cell and they have the same traffic load of 300 bits. UEs are embodied by a specific color in the frames. The color filling rate illustrates the degree of spectral efficiency. The more the PRB is filled, better is the spectral efficiency. A PRB is denoted by the pair (number of set of sub-carriers, number of time-slot). For instance, PRB (1,2) is the PRB on the first line and second column of the frame. It is assumed that red, blue, and orange UEs have a better spectral efficiency on the set of sub-carriers 1, 2 and 3, respectively. On their favorite PRBs, UEs have a throughput of 100 bit/s while on others they have only 50 bit/s. Consequently, UEs need to transmit on three PRBs of their favorite sub-carrier or on six others PRBs in order to empty their buffers. MaxSNR allocates PRBs to UEs TS per TS.

The figure 3.8(a) illustrates the classical behavior of the MaxSNR which can use all the multi-user diversity available (where $K = 3$). As the MaxSNR takes into account the $\eta_{k,n}$, red, blue and orange UEs obtain their three favorites PRBs allowing them to have a great spectral efficiency. Although this MUD value may allow to optimize the spectral efficiency, some PRBs are unused and the bandwidth usage could be improved in this case.

The figure 3.8(b) illustrates MaxSNR behavior with a MUD value set to 1. This MUD value indicates that the focus is put on compressing the transmission time by allowing only one UE to transmit per TS in order to reduce the energy consumption. Since the red UE empties its buffer after transmitting on PRB (1,2), the blue UE is allowed to transmit. On the third TS, the orange mobile is selected to transmit and gets the entire TS. Since it ends its transmission on PRB (2,4), the blue UE is allowed to transmit on PRB (3,4). From this point, no more PRBs are available for this frame. Blue UE has transmitted once on its favorite set of sub-carriers and two times on others. Consequently, the surplus of traffic load for this UE is equal to 100 bits, being one PRB on a depreciate set of sub-carriers and one PRB on its favorite (as the scheduler allocates PRBs TS per TS). Although this MUD value may allow to reduce the energy consumption, the bandwidth usage is once again sub-optimal considering the surplus of traffic load.

The figure 3.8(c) illustrates MaxSNR behavior with a MUD value set to 2. This value provides an accurate trade-off between enhancing the spectral efficiency ($MUD = K$) and reducing the energy consumption ($MUD = 1$). Hence, all the UEs end their transmission during the frame while they have on average, a lower energy consumption than the one provided on figure 3.8(a). This wise optimization of the bandwidth usage is the one that AMUDS permanently seeks by adjusting the MUD value according to the context.

MUD parameter relies on the expectation of the number of PRBs required by an UE k to emptied its packet buffer, denoted RUR_k . It is defined such as:

$$RUR_k = \frac{BO_k + TLe xp_k}{\eta avg_k} \quad (3.18)$$

where BO_k is the buffer occupancy of the UE k while ηavg_k is its average spectral efficiency η during the last 50 frames. $TLe xp_k$ is the estimation of the traffic load produced by the UE k during the frame. This estimation corresponds to the average traffic load of UE k during the last 50 frames. Using the RUR_k criterion, AMUDS varies the value of MUD parameter such as:

$$MUD = \begin{cases} MUD + 1 & \text{if } \sum_{(i=1, i \in S)}^K RUR_k > FS, MUD < K \\ MUD - 1 & \text{if } \sum_{(i=1, i \in S)}^K RUR_k + \epsilon < FS, MUD > 1 \\ MUD & \text{else} \end{cases} \quad (3.19)$$

where parameter S is the set of K UEs linked to the gNB and FS is the frame size defined as the number of PRBs (number of TS per number of sub-carriers).

We also provide an example of a function that prevents AMUDS from switching the MUD value too often. Parameter ϵ is a number of PRBs, correlated with the evolution of the average delay during the last 50 frames such as:

$$\epsilon = \begin{cases} \epsilon + 1 & \text{if } \dot{f}(Delay) > 0 \\ \epsilon - 1 & \text{if } \dot{f}(Delay) < 0, \epsilon > 0 \\ \epsilon & \text{else} \end{cases} \quad (3.20)$$

AMUDS adopts a non-greedy strategy by increasing or decreasing by one the value of MUD parameter. This allows to reduce the impact of incertitude linked to an imprecise Channel State Information (CSI) or a wrong anticipation of the evolution of the traffic load in the cell. In addition, the increase of the multi-user diversity with opportunistic

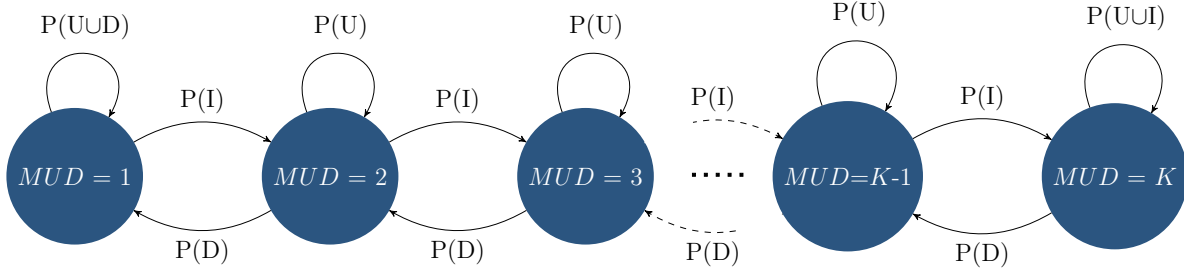


Figure 3.9 – Markov chain illustrating the probabilities that the value of MUD parameter varies.

schedulers makes the spectral efficiency to rise which may also cause CSI to be imprecise. On the one hand, AMUDS directly increases the value of MUD parameter when the sum of RUR_k is greater than FS . This allows to opportunistic schedulers to increase their spectral efficiency in order to handle a more important traffic load at a given time. On the other hand, AMUDS decreases the MUD parameter value when the sum of RUR_k is lower than FS . It means that the system is able to handle the estimated current and future traffic load. AMUDS thus strongly encourage the scheduler to compress the transmission time in order to reduce the energy consumption of UEs. However, decreasing the spectral efficiency can be risky and requires the use of an adjustment variable (denoted ϵ in this study). The main role of this parameter is to take into account the average delay of the system in order to potentially compensate miss computation of metrics such as CSI, estimation of an UE's traffic load and so on. It also prevents the MUD parameter from oscillating between two values too frequently. When the traffic load of the cell remains almost constant and application requirements of UEs are met, AMUDS keeps the same value of MUD parameter (Equation 3.19). Note that a given value v of the MUD parameter means that the scheduler is only able to schedule a number v of UEs during the same TS but UEs selected can change from one TS to another.

Figure 3.9 represents a discrete-time Markov chain where states are the MUD values and each transition is representative of a new frame. The probabilities that the value of the MUD parameter may be Increased by one, Decreased by one or may remain Unchanged are denoted by $P(I)$, $P(D)$ and $P(U)$, respectively. The probability for a scheduler to be authorized to used a certain amount of the multi-user diversity is highly relied on the magnitude of the current RUR_k value. According to Equation 3.19, the more users have new packets incoming in buffer BO_k adding to the older TL_{exp_k} and the less is their spectral efficiency η_k , higher is the amount of multi-user diversity used by a scheduler. The AMUDS solution, by adjusting the multi-user diversity usage thanks to RUR_k criterion

offers to each scheduler (that provides different η_k and buffer management) the accurate multi-user diversity to maximize the mobile transmission compression while guaranteeing the total amount of traffic can be contained in the frame in order to always preserve QoS.

Figure 3.10 illustrates the operation mode of AMUDS. Before the resource allocation process of a given frame, two situations can happen:

- There are no data available from the previous frames. AMUDS consequently, cannot adjust MUD parameter to the context of the cell, leading to set MUD and ϵ values to K and 0, respectively. These initialization values allow to await that the AMUDS gathers data and learns from the context.
- Data are available from previous frames. AMUDS computes RUR_k for all k as well as ϵ in order to adjust MUD parameter value. From this step, depending on $P(D)$, $P(I)$ and $P(U)$, the value of MUD parameter is either decreased, increased or remains unchanged.

Then, the scheduler operates with the multi-user diversity allowed by the proposed solution until the end of the current frame.

Variations of MUD parameter allow to optimize the bandwidth usage. At low traffic load, the network easily meets application requirements of UEs and a lot of PRBs are unused. Taking benefit from this unwise bandwidth usage, AMUDS forces schedulers to compress the transmission time. As the energetic cost C_k is paid more seldom by UEs, their mean energy consumption is reduced at the expense of a poorer spectral efficiency (which is not important at a low traffic load). As the traffic load rises, AMUDS proportionally and gently increases the value of MUD parameter in order to obtain a trade-off between spectral and energy efficiencies. Opportunistic schedulers can hence, take advantage of a better multi-user diversity to increase their spectral efficiency while keeping to compress the transmission time as much as possible. When the system congestion is close, AMUDS pushes the energy efficiency into the background to only be focused on the spectral efficiency optimization in order to ensure a decent QoS to UEs.

Thanks to this dynamical and opportunistic behavior, AMUDS provides energy efficiency on the most acknowledged schedulers without harming the system capacity neither the QoS. In addition, AMUDS cooperates with schedulers without drastically distorting their mode of operation. This allows to keep their intrinsic properties such as system capacity maximization (for MaxSNR), fairness regarding the distance (for PF), a same number of resource allocated for all the UEs (for RR) and so on. Regarding the operation mode of the proposed solution, data required to perform AMUDS are often already known

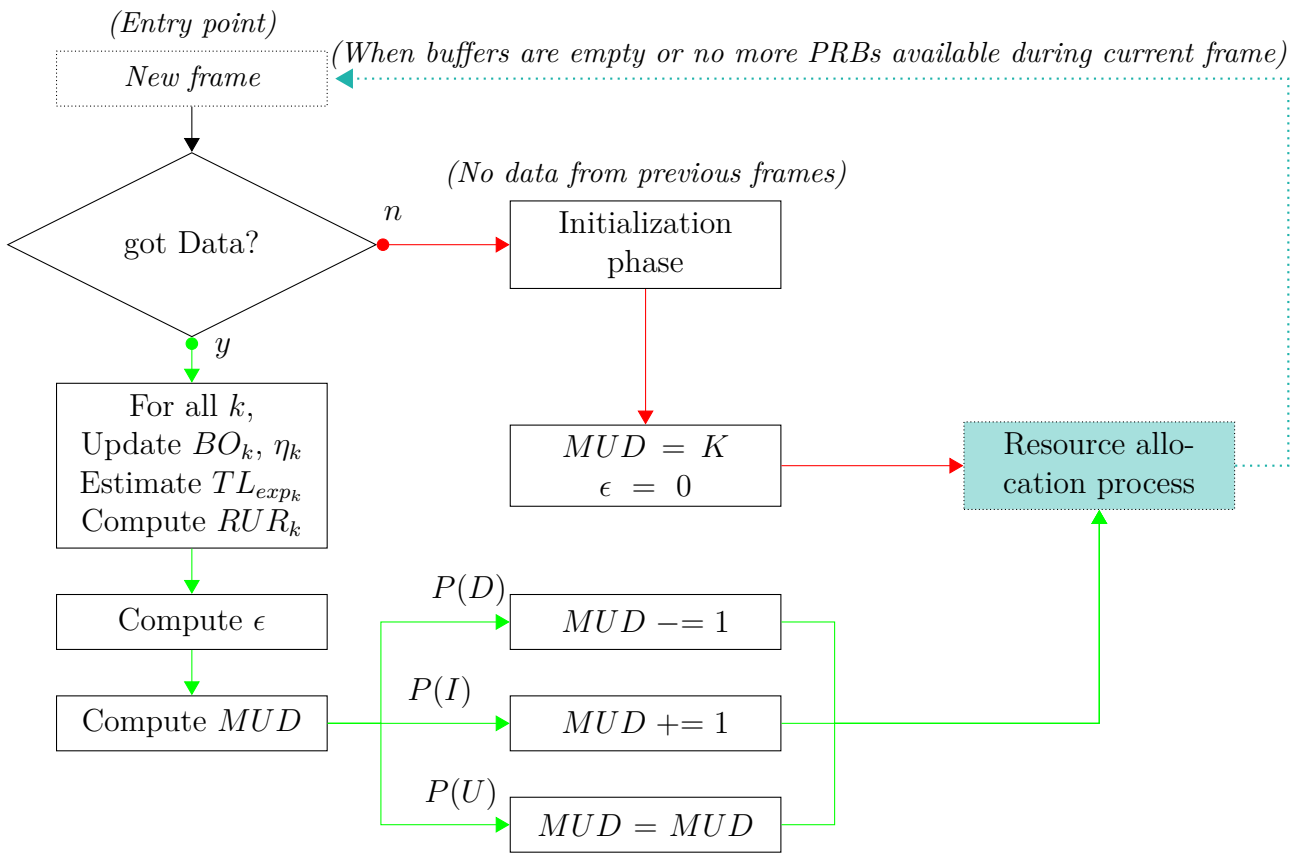


Figure 3.10 – Flowchart of the AMUDS.

and used by schedulers which leads to almost no additional signaling (and complexity). Hence, the proposed solution can be implemented with ease, particularly for a network operator as it is less complex than a whole scheduler and it can adapt to the already existing solution.

Parameters	Value
Cell Radius	500 m
Number of sub-carriers	32
Number of Time Slots	5
RRH transmit power	20 W (43 dBm)
Standard deviation of shadowing	$\sigma = 8$ dB
Path-loss exponent (α)	3.5 (urban context)
Target BER	5×10^{-5}
Subcarrier spacing	15 kHz
Thermal noise power density (N_0)	- 174 dBm/Hz
Simulation duration	500 000 frames

Table 3.6 – Simulations parameters.

3.4.3 Performance evaluation

In this section, performance of the proposed solution are evaluated with Round-Robin, MaxSNR and Proportional Fair schedulers. Whatever the method used to provide performance evaluation, whether it is analytical, in testbed or numerical, to obtain reliable results is not trivial. We choose to provide numerical simulations as the main benefit is to be able to test a various number of solutions and scenarii inexpensively. The simulator is built according to section 3.4.1.

In the performance evaluation, two scenarii are provided. The first is a proof-of-concept where all the UEs are located at the same distance to their gNB. This allows to reduce phenomena that can make the analysis too complex. In the second scenarii, UEs are located at different distances from their gNB. For both scenarii, UEs run realistic applications based on Youtube Streaming⁵⁴. This significantly but realistically complicates the task of schedulers with tight delay constraints and heterogeneous traffic load spikes. Simulation parameters are described in table 3.6.

54. Horvath and Fazekas, *op. cit.*

Scenarios and Key Performance Indicators

Two scenarios are provided. The first is a proof of concept, where UEs are located at the same distance from the gNB. In the second scenario UEs are located at heterogeneous distances from the gNB.

This study focuses on five KPIs to evaluate each solution. These KPIs are identical to those in Section 3.3.3 except for the PDOR metric. Indeed, this new solution is focused on reducing the energy consumption and does not claim to provide fairness between UEs.

- **The bandwidth usage ratio** is the ratio between the number of RUs used by a solution and the total number of RUs available (in %). When this ratio reaches 100% it means that the system is congested.
- **The spectral efficiency** is the mean number of bits received on each RU used (bits/RU).
- **The mean energy consumption** per UE is the mean energy (in mW) that a UE consumes in average.
- **The energy gain** is the percentage of energy gain with a solution which cooperates with AMUD in comparison with the same solution without the cooperative work of the proposed solution.
- **The mean packet delay** is the mean delay to transmit one packet (in ms).

Scenario 1

The main objective of this scenario is to provide a simple context where UEs are located at the same distance from their gNB (100 m) and run the same Youtube streaming traffic⁵⁵. Considering that AMUDS does not perform service differentiation and in order to reduce the simulation time and the convergence delay, only one type of traffic is running for all UEs. This Youtube streaming traffic alternates between two main phases composed either of small traffic bursts or either constant but significant data volume. This allows to run realistic traffic that complicates the task of schedulers. This simulation setup allows to properly explain AMUDS influence on the most acknowledged schedulers without external phenomena interfering in the analysis.

Due to its non-opportunistic behavior, RR provides the poorest spectral efficiency and this metric remains unchanged with AMUDS (Fig. 3.11(a)). This leads to a poor sys-

55. *Ibid.*

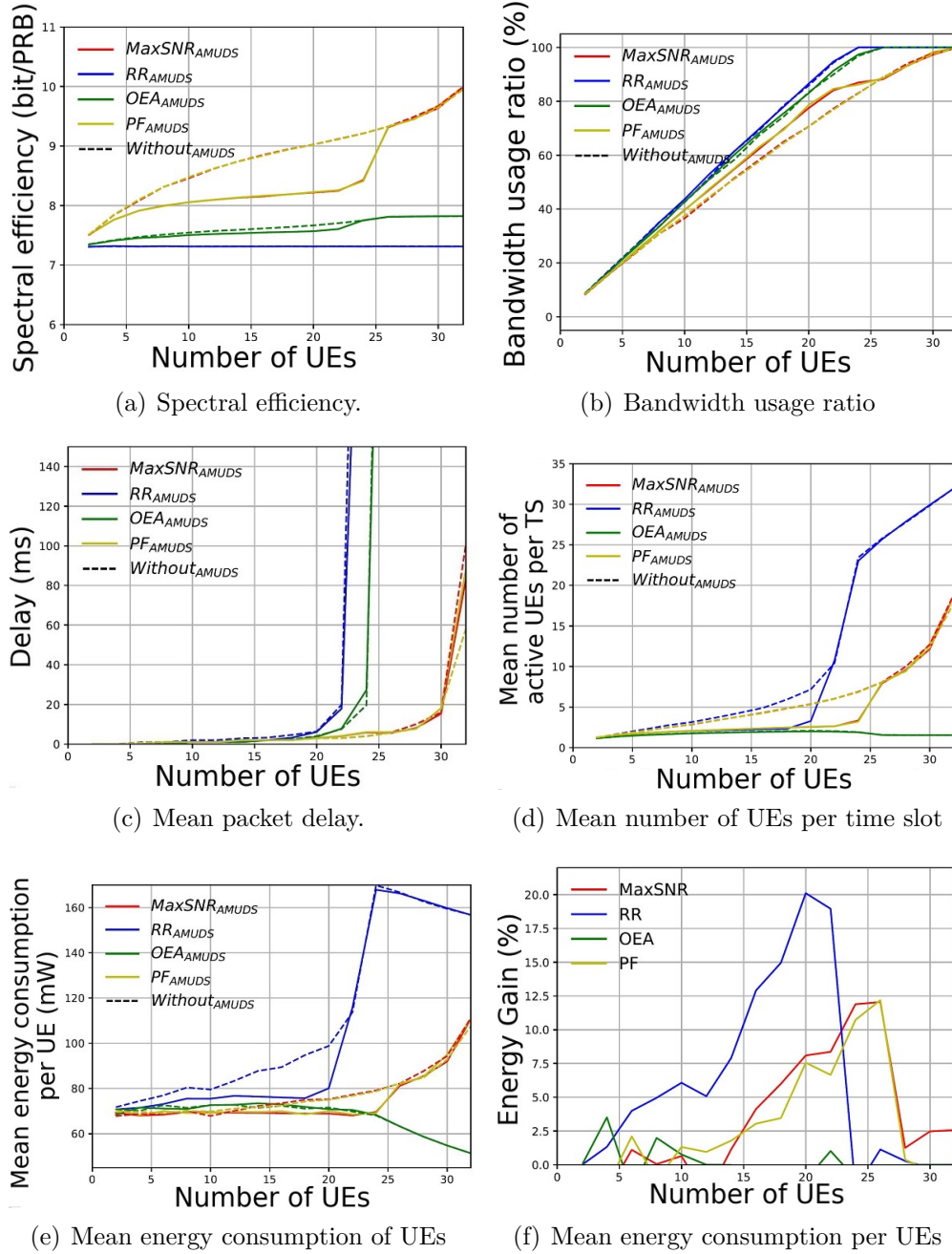


Figure 3.11 – Scenario 1

tem capacity (as the system congestion occurs at 24 UEs) and a poor QoS to UEs (Figs. 3.11(b),3.11(c)). Regarding the energy consumption of UEs, $RR_{withoutAMUDS}$ awakes a high number of UEs per TS due to its mode of operation (Fig. 3.11(d)), leading to a high mean energy consumption per UE (Fig. 3.11(e)). In contrast, RR_{AMUDS} cannot

completely use the multi-user diversity leading to a lower number of UEs awoken per TS (Fig. 3.11(d)). Hence, their energy consumption is highly reduced in comparison with $RR_{withoutAMUDS}$ before the congestion (Figs. 3.11(e),3.11(f)). When the system congestion occurs, AMUDS pushes the energy consumption into the background and allows RR to use the complete multi-user diversity leading to $RR_{withoutAMUDS}$ and RR_{AMUDS} providing the same performance. Note that to avoid content repetition, the impact of multi-user diversity on energy and spectral efficiencies is no longer detailed since it is already discussed in section 2.3 with the same simulation parameters.

OEA is an energy-focused scheduler which gives priority to UEs already awoken during a TS. OEA thus aims to reduce the energy consumption but slashes its multi-user diversity. Consequently, it provides a low energy consumption (Fig. 3.11(e)) at the expense of poor spectral efficiency, system capacity and QoS (Figs. 3.11(a),3.11(b), 3.11(c)). As the operation mode of OEA is to compress the transmission time, AMUDS has a weak impact on this scheduler. Although $OEA_{withoutAMUDS}$ and OEA_{AMUDS} may almost provide the same performance, OEA_{AMUDS} efficiently reduces the energy consumption thanks to a better usage of the bandwidth than $OEA_{withoutAMUDS}$ (Fig. 3.11(f)).

Regarding PF and MaxSNR schedulers, as UEs are located at the same distance from their gNB, they provide the same performance. $MaxSNR_{withoutAMUDS}$ and $PF_{withoutAMUDS}$ take benefit from the whole multi-user diversity (Fig. 3.11(d)) to increase their spectral efficiency as the traffic load rises (Fig. 3.11(a)). This provides the best system capacity and QoS among all the schedulers (Figs. 3.11(b),3.11(c)). $MaxSNR_{AMUDS}$ and PF_{AMUDS} in contrast, are prevented from using the multi-user diversity. Consequently, at low traffic load the focus is put on reducing the energy consumption (Figs. 3.11(e),3.11(f)) at the expense of the spectral efficiency optimization (Fig. 3.11(a)). At a traffic load of 24 UEs in the system, AMUDS estimates that the system will experience difficulties to ensure an adequate QoS to UEs and it swaps of priority by increasing the multi-user diversity (Fig. 3.11(d)). AMUDS operates its transition phase and makes a trade-off between energy efficiency and spectral efficiency. After 26 UEs in the system, the focus is put on optimizing the spectral efficiency as the system is closed to be congested. This leads MaxSNR and PF with and without AMUDS to have the same usage of the multi-user diversity and they almost provide the same performance regarding spectral efficiency, system capacity and QoS.

To summarize, this scenario emphasizes that solutions running the proposed solution make a better usage of the bandwidth by reducing the number of active UEs per TS when

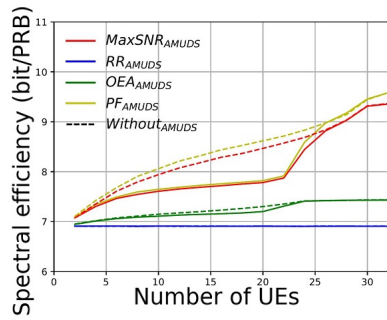
resources are abundant in order to reduce the energy consumption (at low traffic load). When resources tend to become scarce and limited, AMUDS operates a swap of priority and the focus is put on optimizing the spectral efficiency. Consequently, AMUDS is able to provide energy efficiency on these schedulers without harming their performance on system capacity and QoS.

Scenario 2

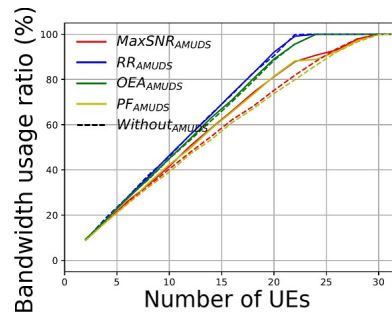
In this scenario, UEs are split into two groups. The first is located at 100 meters to the gNB while the second is located 1.5 times further.

RR and OEA have the same behavior than in first scenario. By not taking into account radio condition, RR provides a poor spectral efficiency, system capacity and QoS (Figs. 3.12(a),3.12(b),3.12(e)). RR_{AMUDS} is restrained from using the whole multi-user diversity at low traffic load leading to a better mean energy consumption per UE and energy gain in comparison with $RR_{withoutAMUDS}$ (Figs.3.12(c),3.12(d)). Due to its objective to compress the transmission time, OEA reduces the energy consumption (Fig. 3.12(c)) at the expense of a lower spectral efficiency than classical opportunistic schedulers (Fig. 3.12(a)) leading to a poor system capacity and QoS (Figs.3.12(b),3.12(e)). As aforementioned in previous scenario, AMUDS has a weak impact on this kind of schedulers due to the use of the same strategy to reduce the energy consumption (i.e. restraining the multi-user diversity to compress the transmission time).

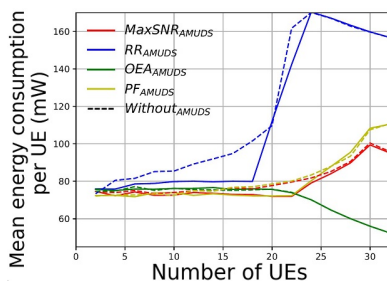
Regarding $PF_{withoutAMUDS}$ and $MaxSNR_{withoutAMUDS}$, they provide unlike performance as UEs are located at different distances from their gNB. MaxSNR segregates UEs regarding their distance and allocates resources first and foremost to those closest to their gNB (as they have on average, a better throughput). Consequently, MaxSNR favors the short term radio conditions and UEs far from their gNB can get access to resources when their radio condition are not optimal. In contrast, PF allocates resources according to the best ratio between short and long term throughput, leading to a better spectral efficiency than MaxSNR on the long term (Fig. 3.12(a)). This also can be highlighted by the mean packet delay per group. MaxSNR is highly unfair on this metric as UEs far from their gNB experience a much more higher delay (Fig. 3.12(g)) than UEs close (Fig. 3.12(f)). On the opposite, PF provides a better fairness between these groups. PF_{AMUDS} and $MaxSNR_{AMUDS}$ are prevented from using the whole multi-user diversity at low traffic load. This leads to efficiently reduce the energy consumption (Fig. 3.12(d)) at the expense of a lower spectral efficiency.



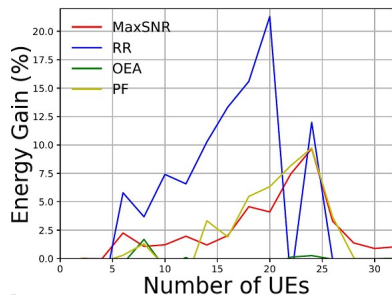
(a) Spectral efficiency.



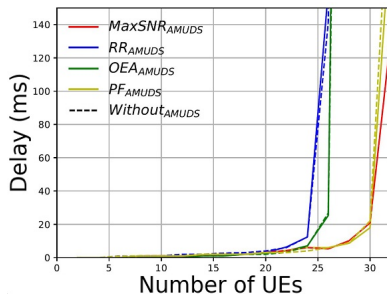
(b) Bandwidth usage ratio



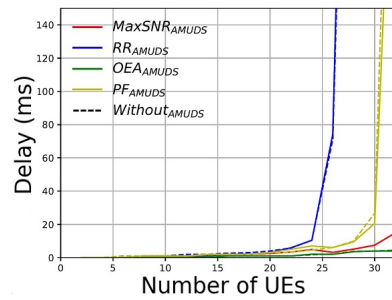
(c) Mean energy consumption of UEs



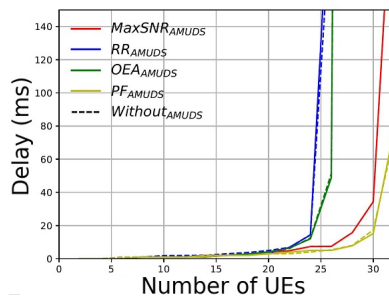
(d) Mean energy consumption per UEs



(e) Mean packet delay.



(f) Mean packet delay of UEs located at 100 m



(g) Mean packet delay of UEs located at 150 m

Figure 3.12 – Scenario 2.

When the system starts to experience difficulties to ensure an adequate QoS for UEs, PF_{AMUDS} and $MaxSNR_{AMUDS}$ swap of priority to enhance the spectral efficiency and system capacity (Fig. 3.12(b)) while keeping the same level of fairness among UEs than $PF_{withoutAMUDS}$ and $MaxSNR_{withoutAMUDS}$ regarding the mean packet delay (Figs. 3.12(f),3.12(g)).

To summarize, this scenario emphasizes that AMUDS is able to provide energy efficient solutions on the most acknowledged schedulers while keeping their intrinsic properties. For instance, $MaxSNR$ keeps favoring UEs close to their gNB while PF provides a better fairness regarding the distance. Moreover, schedulers performance remains unchanged with or without AMUDS when the system is close to be congested.

3.4.4 Conclusion

Global warming issue and the ever-increasing number of devices make energy efficiency of mobile networks a major issue. This study focuses on the reduction of UE energy consumption through a meta-scheduler solution, compatible with the most acknowledged schedulers. The main objective of AMUDS is to enable each scheduler to reduce the energy consumption of UEs without any performance loss regarding the system capacity and the QoS. A wise radio resource management is thus required and AMUDS permanently seeks for this optimization of the bandwidth by adjusting a multi-user diversity parameter according to the context of the cell. When the network effortlessly copes with traffic load and application requirements of UEs are easily met, AMUDS prevents schedulers from using the whole multi-user diversity. This allows to compress the transmission time in order to reduce the energy consumption. As the traffic load rises, AMUDS increases the multi-user diversity usage to enhance the spectral efficiency and the system capacity. This only induces a slight modification of the scheduler behaviors, allowing to keep their intrinsic properties without harming the system capacity neither than the QoS they should provide. Consequently, the proposed solution is able to provide to the most acknowledged schedulers an energy efficient feature without altering their performance.

3.5 Conclusion

Based on the results obtained in Chapter 2, this chapter emphasizes that MUD can be used as a parameter to design new efficient resource allocation solutions. Preventing

schedulers to use the whole MUD allows to compress the transmission time to save energy on UE-side. On the contrary, using the whole MUD allows to optimize the spectral efficiency to increase the overall system capacity.

FEC is a new scheduler that dynamically changes of objective according to the PDOR. Consequently, this scheduler is able to save energy when the QoS requirement of UEs are easily met. When the traffic load rises, FEC slightly changes back of objective and starts to increase the system capacity.

Inspired by the mode of operation of the FEC, AMUDS is a meta scheduler that provides a EE feature to the most acknowledged schedulers. Therefore, schedulers such as MaxSNR or PF for instance, could be energy efficient while keeping their intrinsic properties and performance.

These new solutions improve the proposals of the literature and enlighten that is possible to save energy without any major loss of QoS or system capacity.

ROUTING

This chapter corresponds with the fourth contribution of the thesis. The main contribution of this chapter lies in the design of a new routing solution for 5G wireless networks. The design of such solution is challenging because its goals are various. Beyond the requirements of wired routing solutions (e.g. reliability, convergence time...) an efficient 5G routing solution also has to take into account the channel quality between nodes and reduce the signaling in order to avoid the bandwidth waste. Based on the previous works, this chapter investigates the benefits of an opportunistic approach in routing field. The main objective of the proposed solution is to find the path that provides the shortest packet delay for the considered communication. Inspired by the Little's Law, finding this path is made possible by taking into account the link and nodes states.

4.1 Introduction

Routing in wireless multihop networks raises a lot of interest. Advances in this field open the path to new features in ad-hoc and hybrid networks. On the one hand, finding the appropriate path to route packets in terms of throughput, delay and bottleneck is a crucial issue to ensure the QoS requirements of UEs. On the other hand, routing solutions can help to manage more efficiently the traffic load by using UEs as relay to offload macro cells¹.

An efficient routing solution should be designed to identify the optimal path, that is the set of links that provide overall the best throughput and lowest latency. On a given path, the link with minimal throughput strongly jeopardizes the global Quality of Service (QoS) and Quality of Experience (QoE). As a result, many papers define the optimal path as the one whose bottleneck link is the least constraining. However, if routing protocols

1. Gueguen, C., Rachedi, A., and Guizani, M., « Incentive Scheduler Algorithm for Cooperation and Coverage Extension in Wireless Networks », *in: IEEE Transactions on Vehicular Technology* 62.2 (Feb. 2013), pp. 797–808, ISSN: 0018-9545.

build their decision only considering the path bottleneck value, the probability to over-exploit the same link is high and could result in a high packet delay risk. In addition, the lack of load balancing will induce low system link profitability reducing the global system capacity. Numerous solutions have been proposed for wired networks. However, these proposals are not applicable to wireless networks because of the particularities of radio wave propagation. These require adaptive routing strategies that are able to dynamically take into account the variability of the link throughput as well as the router load, resulting in the continuous selection of a high throughput and low loaded path.

As aforementioned in Section 1.1.2, a significant signal attenuation observed in wireless networks is multipath fading. As a consequence, channel state varies quickly across time, every few milliseconds². On a longer time scale, the channel state also varies due to path loss and shadowing if nodes are mobile. The achievable throughput is thus affected because the modulation scheme of the transmission must be adapted. In the following, we define the short term Link State Information (LSI_{short}) as the measure of the rapidly-changing throughput values due to multipath fading. LSI_{short} values are computed on a short time scale, as opposed to average Link State Information (LSI_{avg}) which is the arithmetic mean of the short term values collected over a larger time scale.

In order to take multipath fading into account, algorithms based on metrics like hop number or the mean Signal to Noise Ratio (SNR) are not efficient because they identify the mean best path in the best case. For instance, in Optimized Link State Routing (OLSR), the rate at which routing information is sent between nodes is in the order of the second^{3 4}. As aforementioned, multipath fading happens on a much shorter time scale. The mean best path is not systematically the short term best path because the latter changes quickly over time. This is why the well-known Ad-hoc On-demand Distance Vector (AODV)^{5 6} is also non optimal in terms of throughput. The path selected by AODV is the one providing the best throughput at a given time, overlooking variations of radio conditions and leading to sub-optimal performance.

2. Goldsmith, A., *Wireless Communications*, Cambridge University Press, 2005.

3. Clausen, T. et al., « RFC7181: The Optimized Link State Routing Protocol Version 2 », *in: IETF-Proposed Standard RFC 7681* (2014).

4. Ferronato, J. J. and Trentin, M. A. S., « Analysis of Routing Protocols OLSR, AODV and ZRP in Real Urban Vehicular Scenario with Density Variation », *in: IEEE Latin America Transactions* 15.9 (2017), pp. 1727–1734.

5. Das, S., Perkins, C., and Belding-Royer, E., *Ad hoc On-Demand Distance Vector (AODV) Routing*, RFC 3561, July 2003, DOI: 10.17487/RFC3561, URL: <https://rfc-editor.org/rfc/rfc3561.txt>.

6. Singh, D. and Ghosh, A. C., « Mobility-Aware Relay Selection in 5G D2D Communication Using Stochastic Model », *in: IEEE Transactions on Vehicular Technology* 68.3 (2019), pp. 2837–2849.

Recent technical advances give access to measures of the instantaneous radio propagation conditions, like LSI_{short} . This allows to design a solution that adapts the path between a source and a destination, as a function of multipath fading, taking inspiration from advances in scheduling protocols^{7 8 9}. For instance, if a node is experiencing strong interference at some point, an opportunistic algorithm can decide to route traffic to an alternative node instead.

In the literature, opportunistic routing refers to packet forwarding solutions based on geolocalisation¹⁰ or algorithms that use an increased number of nodes to transmit data¹¹. The first aims at taking profit of user locations to gain efficiency in the routing of packets. This can benefit to the network coverage but without any delay bound guarantee. The second exploits the radio links diversity to increase the overall communication reliability. Extremely Opportunistic Routing (ExOR)¹² is an emblematic instance of this family of solutions. ExOR combines MAC and routing functionalities. A first packet is sent to multiple nodes of the multihop wireless networks. Based on this transmission experience, the best node is then elected to forward a batch of packets. With this method, long distance but lossy links are advantageously exploited while it would have been discarded by classic routing algorithms. Performance evaluation over a large 802.11 testbed shows significant throughput gain for most links using ExOR compared to classic routing solutions. A major drawback is however traffic increase and subsequent congestion in the network due to packet transmission duplication. This was not properly evaluated in the testbed. Athanasopoulou proposes a back pressure adaptative algorithm¹³. Routing and scheduling components are decoupled in the algorithm by designing a probabilistic routing table that is used to route packets per destination queues. Mao minimizes the overall network energy consumption working on a smart management of the forwarder list with priorities¹⁴. Static and dynamic transmission power management strategies are investigated to elaborate a performant energy opportunistic routing policy. Darehshoorzadeh

7. Wong and Cheng, *op. cit.*

8. Viswanath, Tse, and Laroia, *op. cit.*

9. Gueguen and Baey, *op. cit.*

10. Costantino, G. et al., « LoSeRO: a Locality Sensitive Routing Protocol in Opportunistic Networks », *in: Proceedings of the 31st Annual ACM Symposium on Applied Computing*, 2016, pp. 644–650.

11. Biswas, S. and Morris, R., « ExOR: Opportunistic Multi-hop Routing for Wireless Networks », *in: SIGCOMM Comput. Commun. Rev.* 35 (2015), pp. 133–144.

12. *Ibid.*

13. Athanasopoulou, E. et al., « Back-Pressure-Based Packet-by-Packet Adaptive Routing in Communication Networks », *in: IEEE/ACM Transactions on Networking* 21 (Feb. 2013), pp. 244–257.

14. Mao, X. et al., « Energy-Efficient Opportunistic Routing in Wireless Sensor Networks », *in: IEEE Transactions on Parallel and Distributed Systems* 22 (Nov. 2011), pp. 1934–1942.

proposes a generic Markov model to evaluate candidate selection algorithms¹⁵. The necessary inputs are the candidate list of each node, the link delivery probability, and the maximum number of retransmissions in each node.

All these routing algorithms were primarily designed for ad hoc networks and base their decisions on average SNR values, assuming a relatively SNR stability in the medium term time scale. However, a substantial gain may be drawn taking into account the short term SNR fluctuations that occur in wireless networks. We think that a fully "opportunistic" routing solution in the sense of "opportunistic" scheduling solutions, taking inspiration in Maximum Signal to Noise Ratio (MaxSNR)^{16 17}, Proportional Fair (PF)^{18 19} or Weighted Fair Opportunistic (WFO)²⁰, has a high potential. Taking profit of SNR short term variations in wireless networks, these resource allocation algorithms optimize the system capacity, with a strong impact on packet delay and user satisfaction. The potential of fully opportunistic routing is argued in²¹ through a preliminary theoretical study. A global framework is proposed here to develop routing algorithms but assuming a perfect knowledge of instantaneous SNR values. This is rather theoretical but provided an upper bound. Practical routing solutions should be designed to work even with missing values of SNR, which would be more realistic. Moreover, an in-depth benchmarking with classical routing remains to be done. Link State Opportunistic Routing (LSOR)²² has been conceived to incorporate the short term SNR measures in the path identification decision. As a result, LSOR dynamically adapts the optimal path at each time instant as a function of radio conditions. LSOR is designed to adapt to the various granularity of channel state information available in practical implementations. This leads to improve performance at the expense of a tendency to over-exploit the same link. In²³, authors

15. Darehshoorzadeh, A., Cerda-Alabern, L., and Pla, V., « Back-Pressure-Based Packet-by-Packet Adaptive Routing in Communication Networks », *in: Elsevier Computer Networks* 55 (2011), pp. 2886–2898.

16. Wong and Cheng, *op. cit.*

17. Ostovari, P., Wu, J., and Khreishah, A., « Cooperative Internet Access Using Helper Nodes and Opportunistic Scheduling », *in: IEEE Transactions on Vehicular Technology* 66 (July 2017), pp. 6439–6448.

18. Viswanath, Tse, and Laroia, *op. cit.*

19. Liu, F., Riihijärvi, J., and Petrova, M., « Analysis of Proportional Fair Scheduling Under Bursty On-Off Traffic », *in: IEEE Communications Letters* 21 (May 2017), pp. 1175–1178.

20. Gueguen and Baey, *op. cit.*

21. Chen, W. et al., « Opportunistic Routing and Scheduling for Wireless Networks », *in: IEEE Transactions on Wireless Communications* (Jan. 2017).

22. Gueguen, C., Fabian, P., and Lagrange, X., « Link State Opportunistic Routing for Multihop Wireless Networks », *in: Wireless Networks* (Feb. 2019).

23. Ma, L. and Denko, M. K., « A Routing Metric for Load-Balancing in Wireless Mesh Networks »,

propose a Load Balancing Algorithm using Weighted Cumulative Expected Transmission Time metric (WCETT-LBA) that takes into account mean buffer occupancy but does not consider multipath fading short term fluctuations. This achieves high load-balancing but does not optimize throughput.

In this chapter, we propose to extend and merge LSOR and WCETT-LBA and we present a new Opportunistic Buffer Occupancy Routing (OBOR) solution. Like LSOR, it takes into account the radio condition variations in order to always select an efficient path in terms of throughput but adds, in its decision algorithm, collected buffer occupancy in order to avoid best link over-exploitation. The result is a balanced protocol that better shares the traffic load in the whole network, reducing high packet delay occurrences, increasing link profitability and consequently system capacity. OBOR reaches better throughput and delay compared to state-of-the-art solutions. OBOR could also allow extending cell coverage by being combined with a non-conventional opportunistic scheduler such as in²⁴.

The chapter organization is as follows: the next section presents the state of the art. The third section describes our OBOR solution in details. OBOR performance and literature solutions are compared in section 4. The fifth section discusses overhead. Last section concludes this chapter.

4.2 Related work

Routing problems have been an important research field for years, leading to the implementation of several protocols. In this section, we chose to describe four well known relevant algorithms found in the literature. We dissociate opportunistic protocols from traditional ones such as AODV and OLSR. WCETT-LBA attempts to postpone network congestion using load balancing strategy but experiences the same issues (not considering short term SNR values). In contrast, opportunistic protocol such as LSOR fully exploits wireless network specificities which brings much better performance.

in: 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07), vol. 2, May 2007, pp. 409–414.

24. Gueguen, Rachedi, and Guizani, *op. cit.*

4.2.1 Optimized Link State Routing (OLSR)

OLSR is one of the non-opportunistic routing protocols. It uses routing tables to compute the path to destination. To keep the topology up to date, it sends control messages regularly²⁵. Instead of testing all its neighbors, a node chooses a few neighbors and sends them a control message. Those chosen neighbors are called multipoint relays²⁶. A node's multipoint relay will be the one responsible for forwarding control messages to other "normal" nodes. The normal nodes will just process the control messages and will not forward them. The protocol is indeed efficient in a scaling scenario on huge a topology thanks to its lower flooding and broadcasting compared to Link State Routing (LSR)²⁷. One other OLSR advantage is the absence of delay while establishing a route since all nodes already have routing tables (in steady system). To send a packet, the transmitting router simply has to look inside its routing table and to send the packet to its selected neighbor.

But as explained before, OLSR's main drawback is that it is not opportunistic and does not consider wireless network specificities like multipath fading²⁸. OLSR intentionally does not define how the paths are computed. This allows to provide flexibility according to the context where the routing algorithm is implemented. However most of the implementations take the number of hops as a metric. Regarding our simulation context, our implementation takes into account the average throughput values and the number of hops in order to not downgrade this solution due to our topology grid.

25. Jacquet, P. et al., « Optimized link state routing protocol for ad hoc networks », *in: Proceedings. IEEE International Multi Topic Conference, 2001. IEEE INMIC 2001. Technology for the 21st Century*. 2001, pp. 62–68.

26. Huhtonen, A., « Comparing AODV and OLSR routing protocols », *in: Telecommunications Software and Multimedia* (2004), pp. 1–9.

27. Eom, H., « Information-Dynamics-Conscious Development of Routing Software: A Case of Routing Software that Improves Link-State Routing Based on Future Link-Delay-Information Estimation », *in: The Computer Journal* (Mar. 2008), p. 2.

28. Note that OLSR is hardly improvable by taking short term throughput values into consideration. To create accurate routing tables, OLSR has to converge and it seems hardly realistic with link quality changing very fast over time. Even more, RFC 7181 defining OLSRv2 (Clausen et al., *op. cit.*) sets signaling frame exchange timer to about one second which is much longer than multipath fading variation time. It makes OLSR unable to consider those values.

4.2.2 Adhoc On-Demand Distance Vector (AODV)

AODV chooses a route by using flooding²⁹. Route Requests (RREQs) packets³⁰ are sent whenever a source needs to send data to a new destination. When a node (other than destination) receives a RREQ, it broadcasts it to all its neighbors. If a node receives a second RREQ³¹, this one will be ignored. The chosen route will be the one used by the first RREQ reaching destination, in the most common implementations of AODV. Then, the destination node will send back to the source a Route Response (RREP) packet using the chosen path. Meanwhile, every RREQ reaching destination will be ignored. One advantage of AODV is that the chosen route will be the best in terms of instantaneous throughput and delay at first. But there are three main drawbacks: the first is path load, the second is non-consideration of multipath fading and the last but not least, the high connection establishment delay. Concerning the buffer occupancy drawback (path load), one example can easily be found: during the route calculation, one node can be heavily loaded due to destructive interference, but on average this node is part of the best average route. Indeed, the RREQ will reach destination late due to buffer occupancy and the best average path will not be chosen. The selected route will be the best for a short moment but might not be effective afterwards (for instance when the fully loaded buffer will become empty). The second drawback is the non-consideration of multipath fading. It means the selected route will always be the same until path is cut. It will be exposed to throughput variations (e.g. multipath fading) and will not consider it. In the worst scenario, AODV could take the least efficient average route because it would be lucky at the path selection time due to favourable multipath fading conditions and it will stick to it. It could have terrible performance during the whole connection. Even if AODV would take the best average route, this one might have some low throughput values which constantly change because of multipath fading effect. It means that to get the best path, protocols must look over adaptive and dynamic solutions. Finally, the last drawback is the delay to select the route. Selecting a route can be long because it needs to wait for the RREQ to reach destination from source, and then wait again for the RREP to travel back to source. This is not suitable for real time applications.

29. Das, Perkins, and Belding-Royer, *op. cit.*

30. Eom, H., « Performance Comparison of AODV, DSDV, OLSR and DSR Routing Protocols in Mobile Ad Hoc Networks », *in: International Journal of Information Technology and Knowledge Management* (Mar. 2010), pp. 545–548.

31. toward the same destination and the next hop toward that destination has not been changed

4.2.3 WCETT-LBA

Weighted Cumulative Expected Transmission Time with Load Balancing Algorithm (WCETT-LBA) is based on a new metric trying to avoid selection of overloaded paths. However, WCETT-LBA considers only average router queue length (BO_{avg}) and mean transmission rate values (LSI_{avg}). Multipath fading is not considered, reducing selected path throughput efficiency. In addition, being based on average path selection router queue length does not guarantee low packet delay at long term. Considering short term Buffer Occupancy (BO_{short}) values would have been more profitable. Indeed, it is always more profitable to consider updated and accurate values than average values to optimize system efficiency.

4.2.4 Link State Opportunistic Routing (LSOR)

Link State Opportunistic Routing (LSOR) protocol takes into consideration the known multipath fading values in the route selection process. LSOR uses the same base as OLSR but with this supplementary information, short term SNR will be considered when selecting the best path. This path will be dynamically chosen and corrected whenever short term SNR will change due to multipath fading or node mobility. To be realistic, LSOR is designed with different levels of links knowledge making it usable in practice. LSOR heavily decreases packet delay and enhances system capacity in opposition to traditional (OLSR, AODV) and existing opportunistic (ExOR) routing solutions. However, without considering the router buffer occupancy load, the path selected by LSOR can experience high delays. Other paths with lower (but still efficient) radio conditions and with lower traffic load could provide better QoS/QoE.

4.2.5 Discussion

Literature solutions have their pros and cons. OLSR always knows the whole network topology but only considers mean LSI values such as link throughput. Even if this solution could select the “best” route, this would only be an average best route. The real best path over time varies quickly (on a time scale of 50 ms³²). Taking only the average best route and sticking to it will always be equal or less efficient than selecting the “real best” route based on short term throughput values. In other words, a protocol that selects the route

32. Truman and Brodersen, *op. cit.*

adapting to short scale radio conditions variation will always pick up the best real path over time, which consequently widely increases system throughput capacity and decreases packet delay.

AODV protocol is capable of finding the best instantaneous route in short term values, but will never change it, even with radio condition variations due to destructive interference or node mobility. It will lead to decrease system performance when a variation of radio conditions will occur since the first and last selected route will not statistically always be the best over time.

To conclude this state of the art section, existing solutions from literature are far from exploiting optimal throughput of the links of a topology. Indeed, they are widely outperformed by LSOR which takes into consideration LSI_{short} . However, continuously selecting the best path can cause congestion, drastically increasing packet delays. Consequently, in order to increase network performance, the focus can not be put only on the throughput but additionally on load level of each path. WCETT-LBA attempts to solve this issue. However, it is not designed in order to consider LSI and BO values with different knowledge levels (LSI_{avg} & BO_{avg} or LSI_{short} & BO_{short}) making it inefficient in practice. A possible improvement could be to consider both channel state information in order to provide efficient throughput and buffer occupancy information in order to avoid routers congestion. This could allow to exploit all routes proportionately to their throughput and traffic load.

4.3 Node and link state routing

The main goal of all routing solution is to find the best path toward each destination in order to maximize user satisfaction by providing them with a low delay for instance. This usually means selecting the best route in terms of throughput value. However, traditional solutions found in literature are not built in order to consider the multipath fading effect that can severely affect the performance of their decision. In addition special attention should be given to avoid best path over-exploitation in order to reduce the risk of congestion and the probability to experience high packet delays. OBOR, described in this section, is conceived in order to solve this problem.

4.3.1 Contribution: Opportunistic Buffer Occupancy Routing (OBOR)

The system description is explained through the introduction of the new routing solution.

OBOR manages the routing decision in order to find an efficient path in terms of throughput but with acceptable buffer occupancy. From queuing theory and particularly inspired by Little's law³³, the mean expected duration spent in a system (here, the mean packet delay) is equal to the average number of packets in the system divided by its mean throughput. For a specific link of a router, the metric used can be defined as:

$$C_{link(i,i+1)} = \frac{(1 + BO_i)}{LSI_{i,i+1}}, \quad (4.1)$$

where BO_i is the router i Buffer Occupancy and $LSI_{i,i+1}$ the known link throughput between the node i and $i + 1$. Considering the time taken to travel across a full built path from source to destination, the mean expected packet delay is consequently defined as:

$$C_{path(source,dest)} = \sum_{i=source}^{dest-1} \frac{(1 + BO_i)}{LSI_{i,i+1}} = \sum_{i=source}^{dest-1} C_{link(i,i+1)} \quad (4.2)$$

The objective of OBOR in order to guarantee high QoS and QoE is consequently to find the path j such as:

$$j = \underset{p}{\operatorname{argmin}} \left(\sum_{i=source}^{dest-1} \frac{(1 + BO_i)}{LSI_{i,i+1}} \right), \quad p = 1, \dots, P, \quad (4.3)$$

with p the path index and P the number of possible paths between the source and the destination.

In our opportunistic routing vision, route value relies on the short term values of the channel states and router buffer occupancy sizes. In the perfect scenario, we can consider a full knowledge of this values. However, some of them can change faster or could flood too much signaling to allow full data collection. To get a more realistic context, only a few of these values should be considered as known. We specifically built OBOR in order to adapt to varying degrees of knowledge. In the rest of the paper, we will define the number of hops where OBOR will collect short term LSI and BO values as the parameter k . OBOR

33. Little, J. and Graves, S, *Building Intuition*, Springer, 2008.

could be implemented based on many algorithms: we propose a solution inspired from Dijkstra's algorithm³⁴ to find the path with the minimum delay cost in the network. The metric used in (4.1) dynamically deals with the following rules according to the router knowledge: if short term LSI (LSI_{short}) and BO (BO_{short}) values can be collected and are known (the link is not farther than k hops away from the searching node), then they are considered in the OBOR route selection process. Otherwise, mean BO (BO_{avg}) and LSI (LSI_{avg}) values are considered.

Figure 4.1 is a representation of a possible form for the OBOR solution's flowchart with:

- i is one of the system node
- $C_{path(i)}$ is the link cost from the source to its neighbor i
- $C_{previous(i)}$ is the cost of all links on the path from the source to i
- A node checked is a node that have been chosen considering his $C_{previous(i)}$
- A visited node is a node which have already computed the algorithm during the path discovery

The algorithm is executed by every node, starts when a node wants to send a (data) packet and restarts every time short term throughput or router buffer occupancy values change at a neighbor less or equal than k hops away. Note that it is possible that some distant links on a path (where the short term LSI and BO values are obsolete or unknown) finally experience high multipath fading or sudden high router load. The decision is distributed. The emitter router will narrowly find the best possible route (primary path) considering the values that it is able to obtain at this node (as described above). Throughout the route, instantaneous channel quality and router queue occupancy evolve at each hop. It permits transitional nodes to rework the primary route if a better path is found (appearance of good SNR or low buffer occupancy at the considered hop, but that the source cannot know due to signaling delay restrictions). Consequently, intermediate nodes might be able to adjust the original selected route if they determine that another link grants a better end to the path. Applying these rules, packets will not always be sent on the original route and might be deviated to previously non-selected links that are revealed as better ways.

Considering this distributed approach, the packets managed by OBOR experience lower average delay than with AODV, OLSR, WCETT-LBA and LSOR that do not consider short term LSI and BO.

34. Brassard, G. and Bratley, P., *Fundamentals of Algorithmics*, Englewood Cliffs: Prentice Hall, 1996.

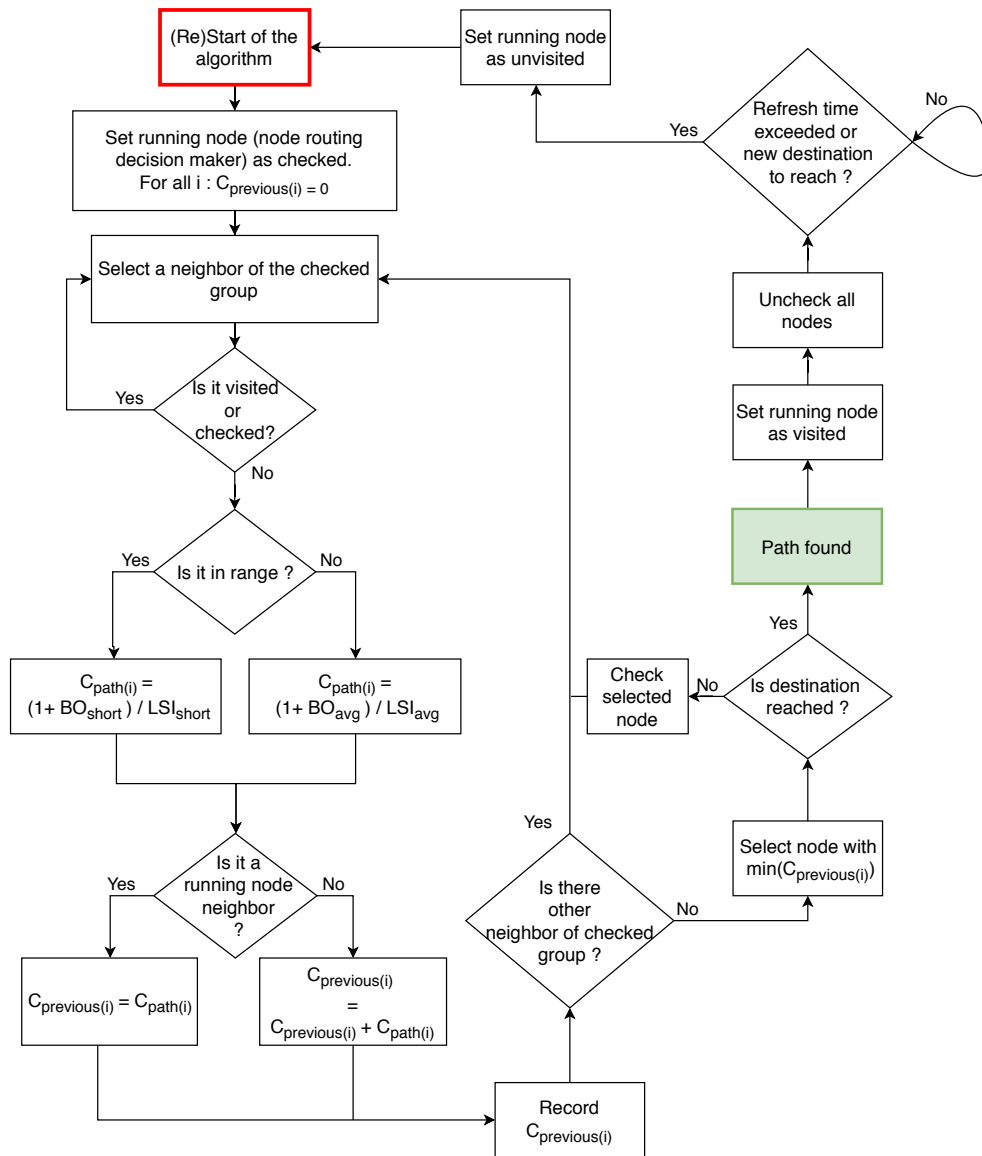


Figure 4.1 – The flowchart of the OBOR algorithm.

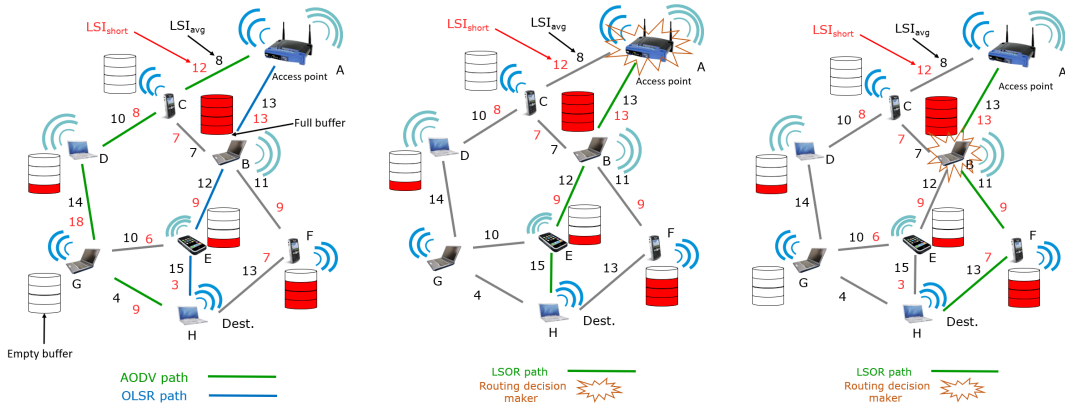
4.3.2 Operation modes of OLSR, AODV, LSOR and OBOR solutions

A perfect topology example is described on Figure 4.2. In this topology, data packets are sent by the source which is the router called "A" to destination named "H". Average links throughput values are written in black and short term values in red. Fig. 4.2(a) illustrates how OLSR and AODV would process the routing problem in this topology at $t = 0$ ms. As expected, OLSR selects the best average route while AODV chooses the best short term throughput route³⁵.

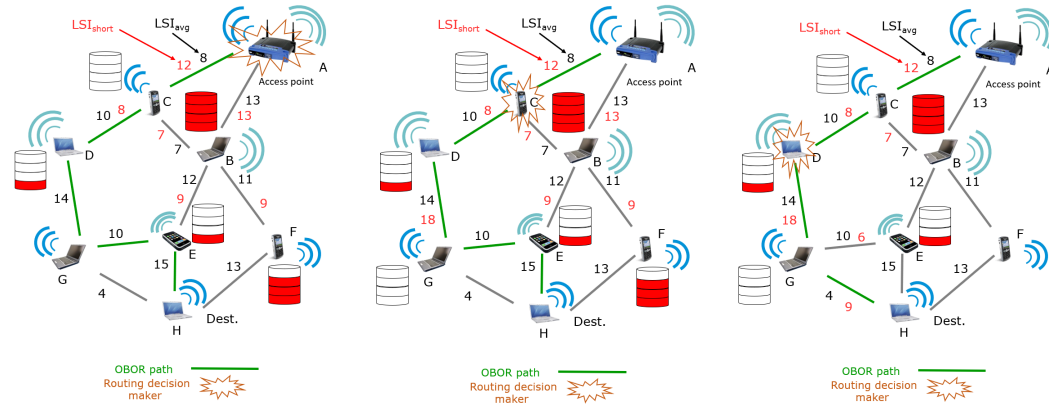
This example (cf. Fig. 4.2) depicts OLSR behaviour which could lead to select a link with a very low short term throughput value. Indeed, we can read a short term throughput value of 3 Mbps on the link (E-H), value which ignored from OLSR. By considering the average value of 15 Mbps instead of its search for the best route, it will select this link. Hence, at this very specific moment, AODV will have a better bottleneck value, 8 Mbps on the (C-D) link, when OLSR bottleneck will be quite lower. But then AODV problem will be the non-consideration of multipath fading, so it will not change its initial route in the future. Consequently, the last selected link of AODV (G-H) which has a low average value will presumably slow down data transmission in the future (see below).

For this example, parameters k is equal to two hops. It means the node using LSOR to find the best route to the destination will be able to know short term values on links up to two hops away. Short term values farther than 2 hops are considered obsolete (older than 50 ms), inaccurate (i.e. unreliable) and consequently, average throughput values will be taken into account to select the potential best route. Fig. 4.2(b) illustrates the information known by the access point with this assumption ($k = 2$) and the route originally selected by LSOR on the access point. Here, using the access point knowledge of the topology, it identifies that the best route is the one with a 9 Mbps bottleneck (A-B-E-H). Consequently, this node delivers packets to mobile B. Note that, in this case, the last link of the selected route has a low short term throughput value (Fig. 4.2(a)) and will affect system to a sub-optimal chosen route (because the access point does not know the short term value of the last link and uses the mean value). However, on Fig. 4.2(c), packets traveled to one hop closer to the destination and mobile B is now the processing node. Furthermore, B being closer to destination, it can compute the short

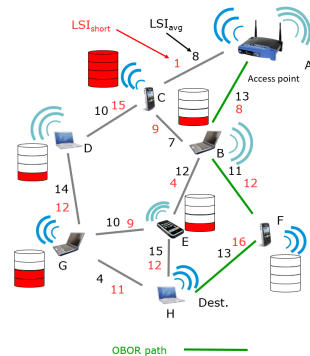
35. AODV floods the whole network with signaling RREQ packets until the first one reaches destination. It will establish the route that will regularly be the best in terms of throughput (if not overloaded) at this moment.



(a) Paths chosen by OLSR and AODV at $t = 0$. (b) Router A: runs LSOR when $k = 2$. The source chooses the next hop at $t = (0 + \epsilon)$ ms. (c) Router B: runs LSOR when $k = 2$. The second node chooses the next hop at $t = (0 + \epsilon)$ ms.



(d) Router A: runs OBOR when $k = 2$ at $t = 0$. (e) Router C: runs OBOR when $k = 2$ at $t = (0 + \epsilon)$ ms. (f) Router D: runs OBOR when $k = 2$ at $t = (0 + \epsilon)$ ms.



(g) OBOR when $k = 2$ at $t = (50 + \epsilon)$ ms. OBOR's path has been adapted to new link throughput and router load.

Figure 4.2 – An example of the paths chosen by OLSR, AODV, LSOR and OBOR when $k = 2$.

term throughput values. B node has got the information that E-H link has low throughput values due to multipath fading effect. Knowing this information, B adjusts the path and sends packets on the route (B-F-H) with a higher bottleneck than the original path selected by the access point. But even if it is profitable in terms of throughput, LSOR selected path relies on mobiles experiencing high traffic load (specifically nodes B and F). It will decrease drastically users' QoE because of the experienced high packet delay.

In order to solve this problem, OBOR considers, in addition to LSI_{short} , the BO_{short} values in the routing decision process. Fig. 4.2(d) highlights OBOR behaviour. The primary path (A-C-D-G-E-H) obtained by the access point's knowledge offers a good throughput limited by a bottleneck of 8 Mbps (a bit less than with LSOR) but it goes through routers with empty or very low buffer occupancy. At the next step, node C does not change the primary path (Fig. 4.2(e)). However, at the third step, node D which is near the destination and taking the opportunities of new LSI and BO information, identifies link (G-E) as worse than link (G-H) and adequately changes the primary path (Fig. 4.2(f)).

Let's take a snapshot (Fig. 4.2(g)) of the route selected by AODV and OBOR with parameter k equal to 2 but when radio conditions have changed, 50 ms later. AODV's route has not been modified (A-C-D-G-H) and the consequences are clear: (A-C) link, the bottleneck of the route, has dropped to 1 Mbps due to multipath fading and AODV is clearly missing the best path, even though it was the best at the path selection instant (also note that node C is overloaded). Hence, AODV shows great performance at the route selection and only occasionally thereafter. AODV's chosen path will most of the time not be optimal. OLSR also sticks to the original path (A-B-E-H), and the bottleneck (now B-E) will now offer 4 Mbps as throughput values, which can be considered as mediocre performance. On the other hand, OBOR detected a variation in the topology short term values and the algorithm has been run in order to select the new best route (A-B-F-H), experiencing a bottleneck of 8 Mbps and low router buffer occupancy for each node composing the path.

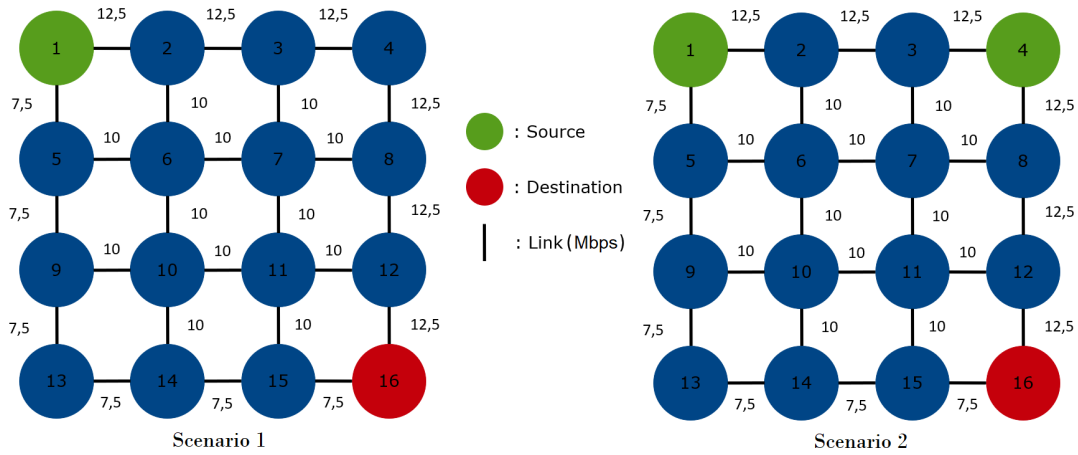


Figure 4.3 – Performance evaluation topologies.

4.3.3 Performance evaluation

Scenarios and Key Performance Indicators

This study focuses on four Key Performance Indicators (KPIs) to evaluate the performance of each solution:

- The system limit capacity is defined as the traffic load reachable (in Mbps) by each protocol before the system congestion. Higher the value is, higher the system is able to process packets.
- The mean packet delay is the mean delay (in ms) to transmit one packet, hence defined as the time between the creation and the arrival of the packet at the destination.
- The system delivery ratio is the ratio (in %) between packets sent and received, denoted as SDR.
- The energy consumption is defined as the mean energy consumed to send one bit between two nodes (in mJ/bit). In this study, we consider the consumption of a mobile while transmitting equals to a constant amount of energy per time tick of 157 mW³⁶.

36. Gueguen, C. and Merlhe, C., « Fair Energy Efficient Scheduler Providing High System Capacity for Wireless Networks », in: *SN Applied Sciences* 2.12 (Dec. 2020).

Scenario 1: single-source context

Figure 4.4 illustrates every protocol results within an asymmetric topology³⁷ for $n = 4$ (Fig. 4.3). As aforementioned, in the best cases, OLSR chooses the best mean route. The best mean route is distinctly identified as the top right one here in our topology. Concerning AODV, it usually chooses the one with the best short term LSI values (statistically the lowest packet delay) at route establishment and then never changes it over time. So the topology aspect at the transmission opening is truly crucial for AODV. It sometimes chooses the best average path, some other time the worst one or the medium one. Here we can see in Fig. 4.4 that OLSR is way better than AODV. AODV might choose the worst average path, which could only be good at the opening of transmission moment, while OLSR will always take the average best path that is better on the long term.

As explained in section 4.2.3, WCETT-LBA is an algorithm based on average values. It will select its path considering both average throughput and average router buffer occupancy. The changing route period of WCETT-LBA is about one second³⁸. It predominantly selects the top right path unless it is overloaded. When this path is overloaded, the algorithm selects another route with lower average LSI values that is statistically less efficient on the long term (longer than 50 ms) and, even if load balancing is allowed, this last route will experience increased difficulties (the top right path has more chances of being efficient on the long term (i.e. time scale of one second)). This explains why it is better than AODV but worse than OLSR (Fig. 4.4).

Concerning LSOR solution, at low traffic load, it provides the best performance regarding the delay (Fig. 4.4(b)) since it continuously chooses the best short term path in terms of throughput while no congestion risk occurs. However when the traffic load increases, the same selected path could be successively chosen and could become overloaded. This can significantly increase packet delay and cause the SDR to drop. Taking into account the radio condition specificities in its routing path selection, LSOR widely outperforms AODV, WCETT-LBA and OLSR in terms of system limit capacity by reaching a value of up to 12 Mbps (Fig. 4.4(a)).

Different assumptions about the depth of available knowledge k of short term LSI and BO have been tested as well as the refreshing rate frequency values. $OBOR_1$ means a node is able to use a short term LSI and short term BO value up to one hop. This allows

³⁷. The network topology is symmetric, however the link bandwidths are not. For instance, links on the top and right network borders have higher bandwidth

³⁸. Ma and Denko, *op. cit.*

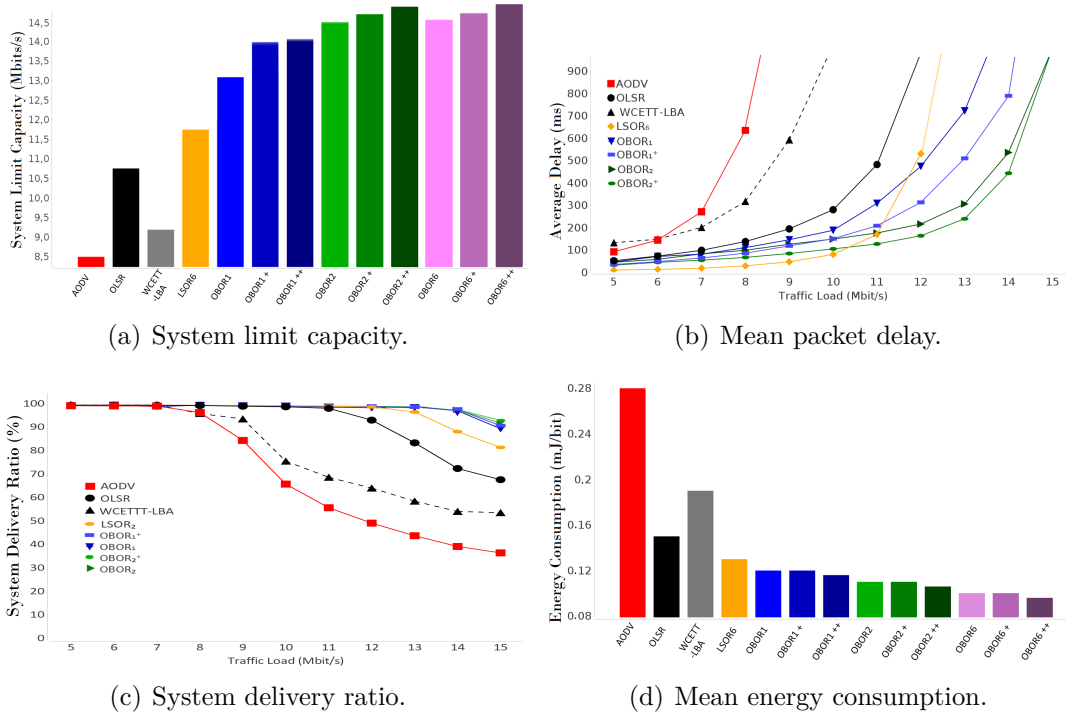


Figure 4.4 – Scenario 1 (single source).

each node to partially adapt both to the path load fluctuation and to multipath fading variations avoiding selecting some temporary overloaded path or bad path in terms of throughput. It drastically affects the system by increasing the whole throughput capacity, increasing the load before the congestion limit of the network by reducing delay (Fig. 4.4(b) and 4.4(a)). The more the acknowledgment of short term LSI and BO values (k) are, the more precise will be the adjustment of the route by every node to prevent overloaded nodes and multipath fading variations. It increases OBOR capacity by splitting single source load on the whole network (Fig. 4.4(a)). As anticipated, the larger k is, the lower will be the arriving packet delay (in every traffic load scenario) and the later the network congestion limit will be reached. In generic topology, for every n (total number of network nodes), values of k exist such that:

$$k \geq 2(n - 1), \quad (4.4)$$

$$k_{complete} = 2(n - 1). \quad (4.5)$$

A knowledge k as in (4.4) also assumes that all short term LSI and BO values of every

network's single links and nodes are known and available on any network node. In a small n topology this is possible but when n starts to increase, it becomes an idealistic assumption and unrealistic from a technical point of view. For the sake of completeness, we ran simulations for all scenarii applying every existing k values between 1 and $k_{complete}$. In those scenarii, n equals to 4 means that to get a complete knowledge, k must be higher or equal to 6. Consequently, $k = 6$ represents the full knowledge of short term LSI and BO in the topology. In these idealistic scenarii, OBOR protocol will systematically find the best route based on short term values. From this point, expanding the knowledge k over 6 will not change anything. However, to reach a total knowledge at $k_{complete}$, it might represent a high overhead and it is a strong assumption. In state-of-the-art radio access management research^{39 40 41 42 43}, the knowledge $k = 1$ assumption is extensively approved since short term LSI (i.e. short term Signal to Noise Ratio) are frequently analysed during the process of scheduling. Since the assumption $k = 1$ can always be treated as valid, $OBOR_1$ results illustrate the minimum guaranteed gain delivered by OBOR. Yet, it is important to note that $k \in \mathbb{N}^*$ and $k \in]1, 3]$ can also bring a discussion (Section 4.3.4) but his study is generally possible/feasible. Therefore, in the following, we focus our study to $k = 1$ and $k = 2$.

Results of the system limit capacity (Fig. 4.4(a)) illustrate OBOR performances for different k and RRF values. For instance, $OBOR_2^+$ means that this version of OBOR has a refresh rate up to 25 ms and a LSI knowledge of two hops. Short term LSI will be the same for 2 route decisions but the BO value can change drastically during 25 ms, even more with a high traffic load. This upgrade permits to be even more accurate about avoiding path overload and network congestion. As illustrate in Fig. 4.4(a), quicker is the refreshing time, slightly better are the results in term of system limit capacity. Concerning the impact of k values on OBOR performance, note that performance of $OBOR_2$ and $OBOR_6$ are very closed. Considering $OBOR_6$ can be seen as $OBOR_{optimum}$ in this topology, these results underline that full knowledge is not required to reach high performance. From this point, as $OBOR_2$ almost reaches $OBOR_{optimum}$ performance with a lower overhead (i.e more realistic) while delay and SDR KPIs are highly related to the system limit capacity, we chose to do not display $OBOR_6$ versions on these two last KPIs

39. Wong and Cheng, *op. cit.*

40. Ostovari, Wu, and Khreishah, *op. cit.*

41. Viswanath, Tse, and Laroia, *op. cit.*

42. Anchun, W. et al., « Dynamic resource management in the fourth generation wireless systems », in: *Proc. IEEE Int. Conference on Communication Technology (ICCT)*, vol. 2, April 2003, pp. 1095–1098.

43. Liu, Riihijärvi, and Petrova, *op. cit.*

for presentation matters.

Focusing on energy, Fig. 4.4(d) represents the energy consumed by each transmitted bit at 15 Mbps. Solutions that do not consider multipath fading never optimize throughput and the transmission takes a longer time. That is why OBOR and LSOR widely outperform traditional routing protocols (respectively OLSR, WCETT-LBA and AODV). In addition, making efficient load balancing allows to transmit the traffic load faster. More nodes are solicited but for a shorter duration that reduces their time of activity, highly greedy in energy consumption. OBOR takes into account the wireless radio condition like LSOR in order to continuously optimize path throughput but, adding buffer occupancy consideration, packet delay is reduced and less energy is consumed.

To recap these scenarii results, WCETT-LBA and AODV deliver good performance on a low traffic load but are widely outperformed by all other protocols with an increased traffic load. OLSR is better since it chooses the best mean path. However, as soon as link radio conditions fluctuate (i.e. short term LSIs differ from mean or long term LSIs), LSOR and OBOR clearly outperform the other solutions. This gap increases with k . $LSOR_6$ being an algorithm with a full vision of short term LSI values, it is clearly good with a low traffic load (Fig. 4.4(b)). With medium and high traffic loads, it has difficulties to keep a good delay and its delivery ratio drops (Fig. 4.4(c)). Even with the knowledge of only LSI_{short} and BO_{short} of one hop ($k = 1$), the less powerful version of OBOR ($OBOR_1$) still outperforms the best version of LSOR considering full knowledge ($LSOR_6$) after 12 Mbps. It can nearly reach 13.5 Mbps before congestion. Indeed, the capacity to balance the load across the network makes it an efficient algorithm to ensure a good scalability when time to increase traffic load comes. It avoids links over- and/or under-exploitation. In addition, the larger the knowledge level k is, the more efficient OBOR becomes (i.e. $OBOR_2 > OBOR_1$). The higher the routing decision period is, the better is the ability of OBOR to avoid congested paths, to reduce the delay and to handle traffic load.

its capacity to reduce delays and increase system capacity (i.e. $OBOR_2 < OBOR_2^+$).

Scenario 2: multi-sources context

Our second scenario is based on multi-sources topology (2 sources : node 1 and node 4 Fig. 4.3). The scenario allows to study the abilities of each solution to avoid selecting a same path for both sources and consequently link over-exploitation. Fig. 4.5 shows the results for this scenario. Both sources produce the same quantity of data. Traffic load information are the data created in the whole network.

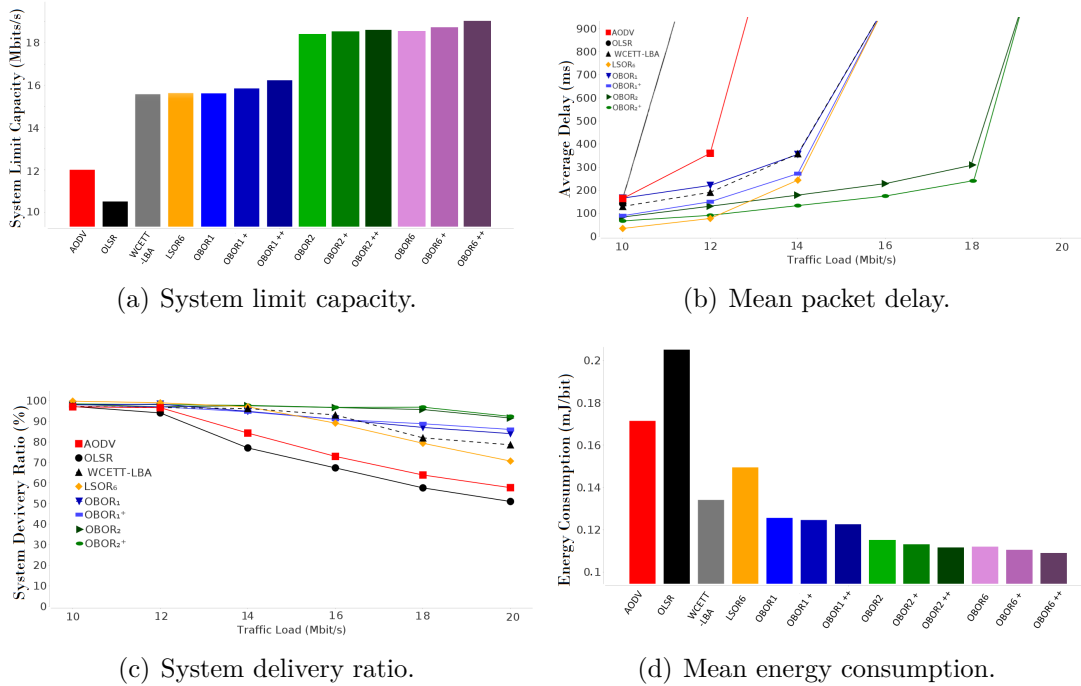


Figure 4.5 – Scenario 2 (multi-sources).

The first thing to notice is the OLSR case. As shown in Fig. 4.3, the average best path goes through both sources. Choosing this path, OLSR transferred the data from the first source to the second node which is also the second source that quickly leads to congestion (Fig. 4.5(a)) and Fig. 4.5(b)). With OLSR, system capacity is consequently the same for scenario 1 (Fig. 4.4(a)) and 2 (Fig. 4.5(a)).

With AODV, the selected paths depend on initial radio conditions. Sometimes the path of the first source does not cross the second source path, sometimes they share the exact same path, and sometimes they only have one or two nodes in common. Making involuntary load balancing, AODV provides in this scenario better performance than OLSR.

The same phenomenon appears with WCETT-LBA. As expected, taking into account average values of routing traffic load, it makes efficient load balancing and outperforms OLSR and AODV in terms of delay, packet delivery ratio and system capacity (Figs. 4.5(b), 4.5(c) and 4.5(a)).

Continuously choosing the best current path at each instant, LSOR keeps good performance on low traffic load situation. In addition since radio conditions vary quickly, LSOR also performs temporary load-balancing that helps to increase system capacity limit and

reduces delay (Figs. 4.5(b) and 4.5(a)). However, the top right path (i.e. mean best path Fig. 4.3), is often selected causing congestion when approaching a traffic load of 15-16 Mbps.

Taking the benefits of LSOR with LSI_{short} and additionally with the BO_{short} values, OBOR is able to reach WCETT-LBA performance using its lowest degree of knowledge ($OBOR_1$ ($k = 1$)). Indeed, with a range of 1 hop, $OBOR_1$ could sometimes engage the flow in a potential good path in terms of throughput and traffic load. When progressing forward, it then discovers too lately that the rest of the path is heavily overloaded or experiences high multipath fading and it could have a lack of choice to find better path. However, if the refresh routing time is slightly increased, $OBOR_1^+$ quickly sees that the selected path starts to be overloaded and the change of path reduces failure effects. This allows $OBOR_1^+$ to provide results very close to $LSOR_6$ while requiring a widely less amount of data to collect (and consequently, easier to implement contrary to $LSOR_6$). With a range of 2 hops, $OBOR_2$ is able to avoid the majority of these cases. Results show that $OBOR_2$ clearly outperforms all other solutions in terms of packet delivery ratio (Fig. 4.5(c)) and system capacity (Fig. 4.5(a)) providing a low packet delay even with high traffic load (Fig. 4.5(b)). As in the previous scenario, the results in Fig. 4.5(a) show that the differences between the performance of $OBOR_2$ and $OBOR_6$ are not significant.

To summarize, as expected, algorithms without any load balancing and path adjustment are heavily outperformed by the others solutions. WCETT-LBA at low traffic load has interesting results because of low variation of router buffer occupancy, however with a high traffic load, average values are not sufficient enough leading to congestion. Like with LSOR, there is a huge difference between one and two hop(s) of knowledge for OBOR but full knowledge is not required. The solutions can be classified according to system capacity limit (Fig 4.5(a)): OLSR (10.5 Mbps), AODV (12 Mbps), WCETT-LBA & $OBOR_1$ & $LSOR_6$ (15.75 Mbps), $OBOR_1^+$ (16 Mbps), $OBOR_1^{++}$ (16.25 Mbps), $OBOR_2$ (18.5 Mbps), $OBOR_2^+$ & $OBOR_6$ (18.6 Mbps), $OBOR_2^{++}$ & $OBOR_6^+$ (18.7 Mbps), $OBOR_6^{++}$ (19 Mbps). This represents a gain for $OBOR_2$ of 54 and 76 % respectively compared to AODV and OLSR.

4.3.4 Discussion about the overhead

Analytical estimation

It is possible to determine long term SNR (LSI_{avg}) values if the transmission power and either the Bit Error Rate (BER) or the Expected Transmission Count (ETX) are known. The short term throughput values (LSI_{short}) can be determined as well if enough transmissions are occurring on a link, which allows the implementation of OBOR. However, the time required to collect these measurements might be longer than the time during which they are relevant (because the channel conditions may evolve quickly). Thus, they can be obsolete reducing the resulting performance of OBOR. To address this issue, other methods can be used such as studied by 3GPP and in literature concerning medium access/radio resource allocation^{44 45 46}. Using channel model equations along with the measurement of received power while transmission occurs on a link yield to more reliable values compared to using a strategy based upon ETX or BER because these strategies are more instantaneous. Furthermore, those more reliable measurements can be accomplished using data packets being transmitted over the channel, reducing the need for dedicated signaling. The only information that is additionally required is the transmission power of the receiver (that the sender can add to a packet). However, if we assume this value rarely changes, the resulting overhead can be neglected.

It is important to note that even though the overhead of short term LSI values and BO is small, it will become significant if the LSI_{short} and BO_{short} of each links are sent to each node of the topology. This is why OBOR is able to work with partial knowledge of LSI_{short} and BO_{short} values. The extra overhead required by OBOR compared to OLSR for different values of k is as follows:

- $OBOR_1 \implies$ Each node only has knowledge of their links. There is no forwarding of LSI_{short} values to neighbors so no overhead to consider. Each router uses power measurements recorded in received data packets. BO_{short} values must still be forwarded, though. Designating N as the number of nodes in the topology and L as the number of links of a node, the node has to forward its data on L links to communicate with its neighbors. Assuming the worst case where every single node has L links and BO_{short} values are updated every T seconds, the total forwarding cost (in packet) is L for one node every T seconds. The global system overhead cost

44. Wong and Cheng, *op. cit.*

45. Viswanath, Tse, and Laroia, *op. cit.*

46. Gueguen and Baey, *op. cit.*

- is consequently $(L * N/50.10^{-3})$ packet/s for $OBOR_1$, $L * N/25.10^{-3}$ for $OBOR_1^+$ and $L * N/10.10^{-3}$ for $OBOR_1^{++}$.
- $OBOR_2 \implies$ The measured short term LSI and BO must be forwarded to each neighbor at one hop. Two cases are treated:
 - It forwards the new measured LSI values at every detected variation (called unsynchronized forwarding and which makes an upper bound). When it detects a variation of short term LSI, the node has to forward its data on $L - 1$ links to communicate to its neighbors. Sent packets can add BO_{short} information and the increased cost is inconsequential for $OBOR_2$ while it is doubled with $OBOR_2^+$ and multiplied by 5 for $OBOR_2^{++}$. Assuming that every node is connected with L links and the short term LSI values variation frequency is 50 ms, during the signaling of these LSI_{short} values, forwarding cost (in packet) is $(L - 1) * L$ for each node every 50 ms. Consequently, the global system overhead cost is $((L - 1) * L * N/50.10^{-3})$ packet/s for $OBOR_2$, $((L - 1) * L * N/25.10^{-3})$ packet/s for $OBOR_2^+$, $((L - 1) * L * N/10.10^{-3})$ packet/s for $OBOR_2^{++}$.
 - The short term LSI measurements forwarding to all links connecting a node are synchronized (called synchronized forwarding and which makes the lower bound). The data will be collected with a minor delay but $L - 1$ short term LSI measurements can be included in a same signaling packet meaning that we can divide the overhead by $L - 1$. Consequently, the total system overhead cost can be decreased to $(L * N/50.10^{-3})$ packet/s for $OBOR_2$, $(L * N/25.10^{-3})$ packet/s for $OBOR_2^+$, $(L * N/10.10^{-3})$ packet/s for $OBOR_2^{++}$.
 - $OBOR_3 \implies$ The measured short term LSI and BO values must be forwarded to each neighbor at two hops. Respecting the same assumptions, we have to deal with two cases:
 - Every single detection of variation, we forward the updated LSI values (unsynchronized forwarding: upper bound). As $OBOR_2$, this packets will be sent to its $L - 1$ direct neighbors, L times⁴⁷. Each connected node have to forward this data to its own neighbors once. Consequently, the total system overhead cost is increased and respectively for $OBOR_3$, $OBOR_3^+$ and $OBOR_3^{++}$ equal to: $(L - 1)^2 * L * N/50.10^{-3}$, $(L - 1)^2 * L * N/25.10^{-3}$, $((L - 1)^2 * L * N/10.10^{-3})$ packet/s.

47. It is the worst case where the diffusing algorithm floods all nodes. It is the most robust but also the worse in term of overhead

- The short term LSI measurements forwarding to every links connecting a node are synchronized to be combined in only one single signaling packet (synchronized forwarding: lower bound). Total system overhead cost can be widely decreased for $OBOR_3$, $OBOR_3^+$ and $OBOR_3^{++}$ to: $(L - 1) * L * N/50.10^{-3}$, $(L - 1) * L * N/25.10^{-3}$ and $((L - 1) * L * N/10.10^{-3})$ packet/s.
- $OBOR_k \implies$ The measured short term LSI and BO values must be forwarded to each neighbors at $k - 1$ hops. Respecting the same assumptions, we still have to deal with two cases:
 - At every single variation detection, we forward updated LSI values (unsynchronized forwarding: upper bound). Total system overhead cost is heightened as: $((L - 1)^{(k-1)} * L * N/T)$ packet/s.
 - The short term LSI measurements forwarding to every link connecting a node are synchronized to be combined in only one single signaling packet (synchronized forwarding: lower bound). Total system overhead cost can be largely estimated to $((L - 1)^{(k-2)} * L * N/T)$ packet/s.

As aforementioned, $OBOR_1$, $OBOR_2$ and $OBOR_1^+$ do not require much overhead and can easily be considered. $OBOR_2^+$ and $OBOR_3$ can also be considered, though they generate more overhead. Higher values of k can only be considered in topologies with lower connectivity (low L values). This is why we focus on $OBOR_2$ (having the same overhead cost than $OBOR_1$ when forwarding is synchronized) and $OBOR_2^+$ in this study. $OBOR_2$ and $OBOR_2^+$ still widely outperform OLSR, AODV, WCETT-LBA and $LSOR_6$.

Overhead comparison on one example

To finely discern cost C in terms of overhead for every solution, we studied their amount of overhead in a grid topology with $n = 3$ and with a constant 50 Mbps traffic load T . In this context, $L = 2.66$ on average and $N = 9$. To create and keep updated routing tables with average LSI values, OLSR needs a precise signaling number (defined as S_{OLSR} in the following). The cost is $C = S_{OLSR}$ for WCETT-LBA, OLSR, and $LSOR_1$ solutions. The cost is $C = S_{OLSR} + (L * N/50.10^{-3})\text{packet/s}$ for $OBOR_1$ protocol. Signaling packets of OBOR incorporate router ID, BO and LSI values. Their length can be evaluated as almost equal to 60 bytes. Overhead signaling for $OBOR_1$ consequently cost $C = S_{OLSR} + 230$ kbps. As previously explained in the last subsection, additional overhead cost to collect short term BO and LSI values is bounded by $L * N/50.10^{-3}$ and

$(L - 1) * L * N / 50.10^{-3}$ signaling packets per second for $OBOR_2$. So, the $OBOR_2$ total network overhead cost is bounded by $S_{OLSR} + 230$ kbps and $S_{OLSR} + 382$ kbps. For $OBOR_{k_{complete}}$ (meaning total knowledge, thus $OBOR_4$), these values are bounded by 633 kbps and $1.051 * 10^3$ kbps (S_{OLSR} is not required since mean LSI values are useless in this case). For AODV, a few signaling packet are sent during route selection by broadcasting and can be neglected if the connection lasts long enough.

Table 4.1 – Global overhead cost estimation for the grid topology with $n = 3$.

Solution	Cost from mean values	Cost from short term values (kbps)	Total overhead cost C (kbps)
OLSR	S_{OLSR}	0	S_{OLSR}
AODV	0	0	ε
WCETT-LBA	S_{OLSR}	0	S_{OLSR}
$LSOR_4$	0	[633, 1051]	$S_{OLSR} + [633, 1051]$
$OBOR_1$	S_{OLSR}	230	$S_{OLSR} + 230$
$OBOR_2$	S_{OLSR}	[230, 382]	$S_{OLSR} + [230, 382]$
$OBOR_4$	0	[633, 1051]	[633, 1051]

4.4 Conclusion

Routing has always been a critical issue in multihop wireless networks. Multipath fading effect has been left behind in state-of-the-art solutions though it is a relevant wireless network particularity affecting every link capacity. It makes it way harder for these protocols to permanently choose the best route with an optimal throughput value. OLSR exploits its routing table knowledge and keeps link value updated considering average LSI values. AODV selects a route by broadcasting packets (RREQ) and selects the path travelled by the first packet reaching destination. But in both cases, the chosen path does not change on the short term variation of the link quality and node buffer load which condemns them not to be optimal in terms of delay and throughput. We are persuaded that links' short term Signal to Noise Ratio values must be widely considered by the use of multipath fading variation knowledge. Previous works on LSOR protocol pointed out that taking into consideration this information is truly beneficial. Collecting these inputs have been proven as realistic in the access point radio resource management research domain (opportunistic scheduling) and grants a massive network performance improvement. The

LSOR protocol can profit from decreasing delay and rise throughput values by more than several tens of percent pushing back the system congestion. However, we demonstrated in this paper the limits of the LSOR algorithm that does not consider router buffer occupancy level in its management. This sometimes conducts to some links over-exploitation while several other links are under-exploited and this phenomenon particularly appears when network traffic load is high. This chapter proposes to add the previous parameter in the path selection process. The new OBOR approach uses less signaling information than LSOR and succeeds in outperforming it. Packet delay is widely reduced thanks to an efficient load balancing while system capacity limit and packet delivery ratio are improved.

INTERFERENCE MANAGEMENT IN A CELL-LESS CONTEXT

One of the main objective of this thesis is to investigate the benefits of the combination of the opportunistic resource allocation and the cell-less context. Unlike classical cellular network, the cell-less approach considers the network as a "hyper-cell" rather than a sum of quite independent cells (cf. Sections 1.1 and 1.3). Consequently, the network is no more cut into cells but into clusters of gNBs that are dynamically constituted according to the users needs. This allows to the gNBs to share more information such as scheduling information, CSI, user data and so on. Thanks to the results obtained and the knowledge acquired from the works presented in previous chapters, we can establish that these kinds of resource allocation (i.e. scheduling, meta-scheduling and routing) do not entirely benefit from such emerging paradigm. Indeed, the scheduler is in charge of allocating PRBs to UEs for only one gNB. In the end, designing a scheduler that deals with more gNBs results in managing interference between gNBs. Routing is already a field where information is shared among nodes and where the decision making can be centralized. Therefore, the benefits -if they exist- of designing routing solutions in a cell-less context are hardly palpable.

On the contrary, the ICI management is a multi-cellular issue where the resource allocation has to be performed according to neighboring gNBs. Thus, new and more dynamical ICI management solutions can be designed thanks to the cell-context. This chapter introduces two new solutions for managing ICI. The first, is a solution that dynamically performs frequency reuse for the PRBs considered. The second is an extension that uses the CoMP and investigates the benefits of such approach.

5.1 Introduction

One of the main objective of network operators is to provide to their clients a coverage area anywhere on the territory. This often involves to have a sufficient number of gNBs and adapted to users' needs. The increase in the number of base stations and the variety of technology lead to the so-called ultra-dense and heterogeneous networks. The side effect of such approach is that there can be overlapping areas between cells or some cells may be included in others. This rises the probability that ICI occurs. Consequently, network operators cannot only rely on cell densification to ensure broadband services and to face the exponential increase of the number of devices in coming years. Therefore, ICI management is a major concern to increase the spectral efficiency as well as QoS of UEs. Moreover, it is important to have a particular attention to UEs located at cell edges, as they are far from their gNB and close to neighboring cells. Due to a high path loss and the magnitude of interference potentially received, they have on average, a poor Signal-to-Interference-and-Noise Ratio (SINR). This makes difficult the meeting of their application requirements. Most of the time, the way that ICI management solutions deals with these UEs is highly correlated to the performance provided by these solutions (especially in terms of spectral efficiency, system capacity and QoS of users).

Classically in wireless networks, ICI are managed thanks to frequency reuse schemes¹. This kind of solutions decides beforehand if a frequency already used by a neighboring cell will be reused by the considered cell. On the one hand, if the frequency is reused, all the bandwidth can be shared between the cell and its neighbors but ICI occurs. On the other hand, if the frequency is not reused, this prevents ICI from occurring at the expense of a high bandwidth waste. Therefore, these solutions attempt to solve a multi-cellular issue (i.e. ICI) in the intra-cellular context. Indeed, as there is no information related to the resource allocation shared between gNB, they have to roughly manage ICI by preventing gNB from transmitting on parts of the bandwidth that are already used. This comes down to either sacrifice cell edge UEs or either reducing the usage of the spectrum. This kind of solutions and their varieties are further detailed and explained in section 5.2.

Lately, the emergence of new network architectures such as Cloud Radio Access Network (C-RAN) or Open Radio Access Network (O-RAN) enables to open new prospects and to consider innovative approaches to replace the classical cellular view of wireless

1. Elayoubi, S., Haddada, O. Ben, and Fourestie, B., « Performance evaluation of frequency planning schemes in OFDMA-based networks », *in: IEEE Transactions on Wireless Communications* (2008), pp. 1623–1633.

networks. Traditionally in C-RAN architecture^{2,3}, base stations are split into three parts: the Remote Radio Head (RRH), the Baseband Unit (BBU) and the optical communication link with high-performance, low delay and high bandwidth front-haul which connects the RRH to the cloud pool where BBUs are gathered. In C-RAN, most of the computational tasks can be considered as moved to BBU pool such as the CSI estimation, the centralized signal processing management of the RAN etc⁴. The RRHs include radio antennas with their associated amplifier and are dispatched among several remote sites⁵. Thus, it is possible to distinguish two C-RAN oriented approaches.

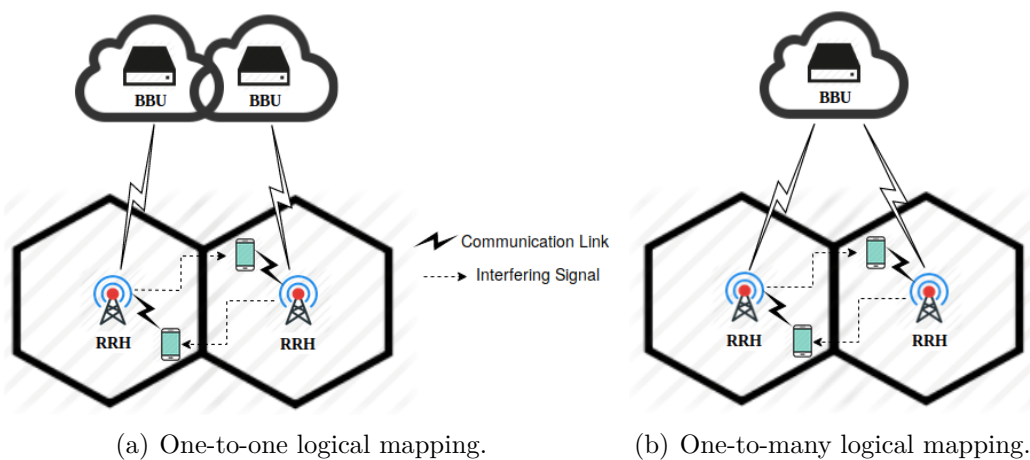


Figure 5.1 – Example of C-RAN approaches

The first is the one-to-one logical mapping (Fig. 5.1(a)) where one BBU is assigned to one RRH. Although BBUs are gathered in the same BBU pool which can reduce the maintenance costs, this mapping only allows to manage one RRH at the same time. This makes ICI management solutions to have a static behavior. Consequently, it is not suitable to easily react to the channel state variations and to the UE mobility between two cells.

Another approach is the one-to-many mapping (Fig. 5.1(b)), where one BBU is assigned to several RRHs. It is in this context that the cell-less approach has emerged. The network is now seen as an "hyper-cell" rather than a sum of independent cells. The cells become a set of gNBs and the decision making for this cluster is logically centralized. In

2. Checko, A., *Cloud Radio Access Network Architecture. Towards 5G Mobile Networks*, 2016.

3. Park, S. et al., « Robust and Efficient Distributed Compression for Cloud Radio Access Networks », *in: IEEE Transactions on Vehicular Technology* (2013), pp. 692–703.

4. Wang, K., Yang, K., and Magurawalage, C.S., « Joint Energy Minimization and Resource Allocation in C-RAN with Mobile Cloud », *in: IEEE Transactions on Cloud Computing* (2018), pp. 760–770.

5. Ezzaouia, M. et al., « A Dynamic Transmission Strategy Based on Network Slicing for Cloud Radio Access Networks », *in: Wireless Day*, 2018, pp. 40–45.

this way, the cell-less approach and the one-to-many mapping allow to manage several cells with only one entity. Thus, ICI can be fully controlled in a multi-cellular oriented approach allowing to share frames between RRHs. This provides a better usage of the bandwidth while efficiently reducing the magnitude of interference received. This is done at the expense of higher signaling rate and computation complexity. However, this thesis only focuses on the RAN and although being aware of these drawbacks, they won't be dealt with in what follows.

In the first part of this chapter, a new ICI management solution is presented. The Dynamic Cell-less Radio Access Network Meta- Scheduler (DC-RAN-MS) is based on the one-to many mapping and takes benefits from the centralized decision making provided by this architecture. It relies on the following analysis:

- ICI management performed in intra-cellular domain often induces a waste of bandwidth due to frequencies allocated beforehand. For instance, classical solution does not take into account that a UE could not be interfered at a given time which lead to needlessly bandwidth waste.
- Static solutions cannot adapt to the distribution of UEs and to their channel state. In this way, they neglect that UEs at cell edges could have a sufficient SINR to perform a Reuse 1 strategy.

Consequently, the DC-RAN-MS merges the ICI management with the scheduling in order to provide a more accurate resource allocation. The proposed solution allows or prevents schedulers to allocate resources according to: the CSI (i.e. the SINR/SNR), the magnitude of interference experienced and the number of UEs available to transmit in each RRH. The entire bandwidth is shared between the cluster of RRHs considered and the resources that are not allocated by a RRH can be used by another. This provides a more accurate and dynamic ICI management which reduces the bandwidth waste and the magnitude of interference experienced.

However, this solution can be improved by the usage of the Coordinated MultiPoint (CoMP). The CoMP has been introduced for LTE networks by 3GPP⁶ and is now considered as a key feature for 5G wireless networks to mitigate ICI⁷. It fits appropriately with a cell-less approach and allows to manage ICI more efficiently. The use of the Joint Transmission Coordinated MultiPoint (JT-CoMP) provides better performance than other type

6. 3GPP, « Coordinated Multi-Point Operation for LTE Physical Layer Aspects », *in: 3rd Generation Partnership Project (3GPP), TR 36.19 R11 v11.2.0* (2013).

7. Li, Q.C. et al., « 5G Network Capacity: Key Elements and Technologies », *in: IEEE Vehicular Technology Magazine* (2014), pp. 71–78.

of CoMP^{8,9}. It allows to any UE to receive signal from many gNBs at the same time. Consequently, serving gNBs will need to jointly transmit the same user data and to reserve the same radio resource. This leads to mitigate ICI and to slightly increase throughput at the edges at the expense of higher radio resource consumption, signaling and computing among access points.

The second part of this chapter is therefore, the presentation of a new ICI management solution named Hybrid Joint-Transmission Coordinated MultiPoint (H-JT-CoMP). It is an extension of DC-RAN-MS that uses the JT-CoMP. For each PRB, H-JT-CoMP dynamically performs either frequency reuse or JT-CoMP.

5.2 Related work

The main objective of this section is to provide the reader an overview of the literature regarding ICI management. Therefore, six solutions have been selected as they are representative of each specialization: frequency reuse (Reuse 1), interference avoidance (Sliced Bandwidth), combination of the two previous specialization (Fractional Frequency Reuse), combination of power and resource allocation (Soft Frequency Reuse), coordination of scheduling and ICI management (Hybrid Static) and Coordinated Multipoint (Power Level Difference CoMP).

Note that the solutions using beamforming are not presented since it induces flat fading which is in contradiction to the opportunistic approach investigated in this thesis. However, a new approach known as the User Centric Cell-Free Massive MIMO that combines both the benefits of the JT-CoMP and the massive MIMO seems really promising¹⁰. Indeed, this takes even further the concept of cell-less as there is no more gNB but only distributed antenna. This context should be addressed in future works.

5.2.1 Reuse 1 (R1)

The Reuse 1 (R1) solution is the classical frequency reuse scheme¹¹. Figure 5.2 illustrates a simple representation of the mode of operation of R1. Each cell uses the total

8. Qamar, F. et al., « A Comprehensive Review on Coordinated Multi-Point Operation for LTE-A », in: *Computer Networks* (2017).

9. Bassoy, S. et al., « Coordinated Multi-Point Clustering Schemes: A Survey », in: *IEEE Communications Surveys Tutorials* (2017), pp. 743–7648.

10. Demir, Ö., Björnson, E., and Sanguinetti, L., *Foundations of User-Centric Cell-Free Massive MIMO*, vol. 14, 3-4, Now Publishers, 2021, pp. 164–472.

11. Elayoubi, Haddada, and Fourestie, *op. cit.*

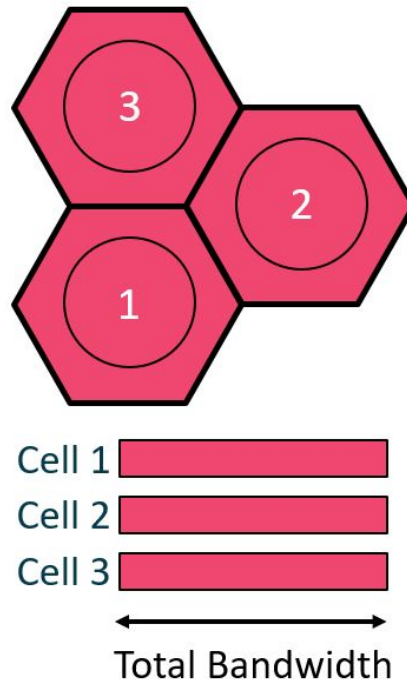


Figure 5.2 – Reuse 1 strategy.

available bandwidth (represented by the same color), meaning all the radio resources defined by the frequency and time (i.e. PRBs) are reused. Interference occurs when at least two cells allocate the same PRB to their respective UEs. In this case, depending on the UE proximity to neighboring cell, the impact of the interference may be very significant on the transmission efficiency. Inner cell UEs receive weak magnitude of interference and their application requirements are easily met while cell edge mobiles are not protected from interference leading to a poor QoS. Although this solution reuse all the available bandwidth, it is highly unfair between mobiles at the inner and edges of the cell

5.2.2 Sliced Bandwidth (SB)

The classical ICI avoidance Sliced Bandwidth (SB) has been designed in order to protect UEs at the edges¹². Figure 5.3 illustrates a simple representation of the mode of operation of R2. Each color represents a frequency range allocated to only one cell. For instance, the first frequency range (in red) is allocated to the cell 1. Inside each cell, PRBs allocated to each UE are different (because not on the same frequency range),

¹². *Ibid.*

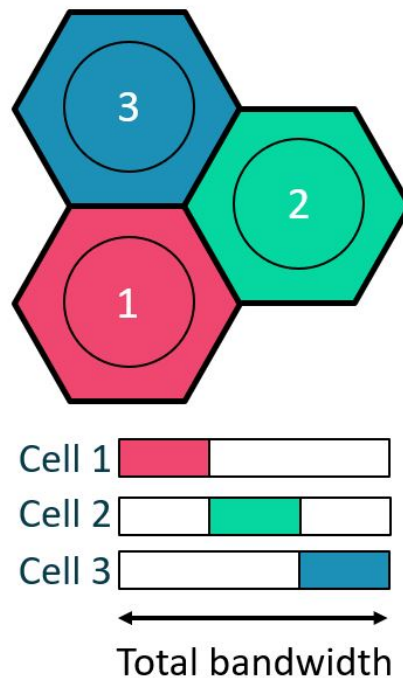


Figure 5.3 – Sliced Bandwidth strategy.

preventing interference from occurring. It allows UE to have a good throughput but the system capacity is very limited due to the high bandwidth waste.

5.2.3 Fractional Frequency Reuse

Fractional Frequency Reuse (FFR) is a more evolved approach that combines the benefits of R1 and SB^{13 14}. As shown on Figure 5.4, R1 is performed in the inner of the cell while SB is for the edges. It does not require any cooperation between cells as the number of PRB allocated to each area is fixed beforehand. This induces that FFR is not able to adapt to the PRB allocation to the UE distribution in the cell.

5.2.4 Soft Frequency Reuse

Soft Frequency Reuse (SFR) is similar to FFR but the main difference lies in it combines both resource and power allocation^{15 16}. At cell edges, the maximum power is used

13. Qian, M. et al., « Adaptive Soft Frequency Reuse Scheme for Wireless Cellular Networks », *in: IEEE Transactions on Vehicular Technology* (2015), pp. 118–131.

14. Kumar, S., Kalyani, S., and Giridhar, K., « Impact of Sub-Band Correlation on SFR and Comparison of FFR and SFR », *in: IEEE Transactions on Wireless Communications* (2016), pp. 5156–5166.

15. Qian et al., *op. cit.*

16. Kumar, Kalyani, and Giridhar, *op. cit.*

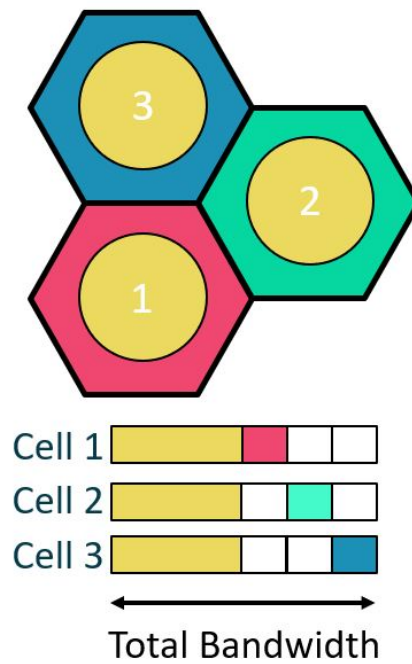


Figure 5.4 – Fractional Frequency Reuse strategy.

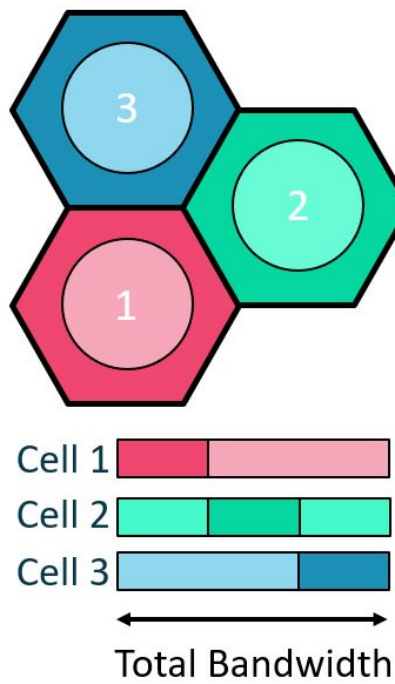


Figure 5.5 – Soft Frequency Reuse strategy

to transmit while in the inner the power transmission is reduced. This leads to a better use of the spectrum than FFR (Fig. 5.5) but it has the same drawbacks (i.e. it is a static solution).

5.2.5 Hybrid static (HS)

In¹⁷, the Hybrid Static (HS) strategy is presented. It is based on the analysis that giving PRBs to UEs from different cells, at the same time and on the same frequency is better than protected them when the magnitude of interference received is not significant. Thus, the HS splits its bandwidth into two slices. In the first slice, a R1 strategy is used and it is allocated to UEs in inner cell. The second slice is allocated to UEs at cell edges where a SB strategy is performed. In this slice, the HS compares the channel state of mobiles chosen by the intra-cell scheduler and allocates the resource to the one with the best SNR. Thus, this slice is fully shared between the RRHs which allows to optimize the resource allocation regardless of the number of mobiles in the edges of cells considered. Consequently, the HS solution makes a better usage of the spectrum than SFR and FFR approaches. However, due to a static boundary between the two slices, the same number of PRBs is allocated to cell edges and inner UEs. This makes this solution to have a static behavior which can hardly adapt to the context. Moreover, depending on the mobile distribution, one slice reaches the congestion before the other while PRBs are still available in the other slice.

5.2.6 Power Level Difference CoMP (PLDCoMP)

The last solution from the literature uses the JT-CoMP. One of the main issue of this field is to perform a CoMP clustering. It means that clusters of serving gNBs for a given UE have to be dynamically created. Moreover, as the use of JT-CoMP leads to a high bandwidth usage, it is necessary to limit the use of JT-CoMP to the UEs highly interfered.

Shami proposes¹⁸ a JT-CoMP clustering. In the following of the chapter, this solution is referenced as Power Level Difference CoMP (PLDCoMP). The PLDCoMP selects the two strongest signals received by an UE to define serving gNBs to perform CoMP. Then,

17. Ezzaouia et al., *op. cit.*

18. Shami et al., *op. cit.*

it relies on a similar method presented in ^{19 20}, based on the Power Level Difference (PLD) value, to determine whether the signal powers are comparable. If they are, the UE is in CoMP mode, else the UE is in non-CoMP mode. Unlike previously described solutions, the PLDCoMP dynamically performs its ICI management by allocating resources to UE in non-CoMP and CoMP modes such as:

$$BW_{i,non-ComP} = \frac{RUs}{(NC + (b \times CO))} \times NC \quad (5.1)$$

$$BW_{i,ComP} = BW_{i,total} - BW_{i,non-ComP} \quad (5.2)$$

where $BW_{i,non-ComP}$ and $BW_{i,ComP}$ are the bandwidth allocated to the number of NC mobiles in non-CoMP mode and to the number of CO mobiles in CoMP mode, respectively. RUs is the number of available resources while b is a constant set to 0.25 to limit the usage of the bandwidth for mobiles in CoMP mode. Consequently, this solution makes a better bandwidth usage than previously described solutions and allocates resources depending on the number of mobiles in CoMP and non-CoMP modes. However, the b constant induces a restriction in the bandwidth usage for edge mobiles which provides them a poor Quality of Service. Although the PLDCoMP relies on a CoMP implementation with a cell-less approach, this solution does not entirely take benefits from the flexibility provided by the CoMP.

5.3 System description

This study focuses on the ICI management issue for the set of UEs located in the coverage zone of several gNBs. A logically centralized approach is assumed since this allows efficient opportunistic approaches.

The physical layer is considered to operate using a Time Division Duplex mode (TDD) which allows a good compatibility with the OFDM based transmission mode²¹. The global available bandwidth is split into sub-frequency bands called sub-carriers. Radio resources are distributed in time domain in frames themselves split in Time Slots (TS)

19. Sakr, A.H. and Hossain, E., « Location-Aware Cross-Tier Coordinated MultiPoint Transmission in Two-Tier Cellular Networks », *in: IEEE Transactions on Wireless Communications* (2014), pp. 6311–6325.

20. Bassoy, S. et al., « Load Aware Self-Organising User-Centric Dynamic CoMP Clustering for 5G Networks », *in: IEEE Access* (2016), pp. 2895–2906.

21. Kela, Turkka, and Costa, *op. cit.*

of constant duration (which is an integer multiple of the OFDM symbol duration). Each frame duration is assumed equal to a value inferior to the coherence time of the channel, allowing transmission on each sub-frequency to experience flat fading during each frame. Full knowledge (through CSI, for instance) of radio conditions is supposed to be available at the receiver²². Thanks to SNR and SINR measurements of the signal sent by each UE (for instance during the uplink contention subframe), the gNB is thus, able to estimate their channel state attenuation at a given TS on each sub-carrier. According to²³, channel state can be assumed stable on a scale of 50 ms. This requires that UEs have to transmit their control information alternatively on each sub-carrier once for several frames (in order for the gNB to successfully refresh the channel state information). A PRB, defined as a (15 sub-carriers, time slot) pair, can be allocated to any UEs with a specific modulation order. Transmissions performed on different PRBs by different UEs have independent channel state variations. On each PRB, a modulation order adapted to the channel state (between the gNB and the selected mobile) is assumed. This provides the flexible resource allocation allowing opportunistic approaches.

A C-RAN architecture is considered leading to split gNB into RRHs and BBUs. The channel gain between the RRH i and the UE k on the sub-carrier n is given by :

$$G_{k,n}^i = h \times 10^{\frac{X\sigma}{10}} \times \left(\frac{1}{d_{k,i}}\right)^\alpha \quad (5.3)$$

where h represents the Rayleigh multi-path fading, which is modeled by an exponential distribution, X is a standard Gaussian random variable, σ is the standard deviation of shadowing in dB, $d_{k,i}$ is the distance between the mobile k and the RRH i and α is the path loss exponent. We denote C , the set of L neighboring RRHs of RRH i considered mapped to the BBU of the DC-RAN. Thus, the Signal-to-Interference-plus-Noise Ratio of user k on sub-carrier n associated to RRH i is given by :

$$\gamma_{k,n}^i = \frac{P_n^i G_{k,n}^i}{B_{sub} N_0 + \sum_{(j=1, j \in C)}^L P_n^j G_{k,n}^j} \quad (5.4)$$

where P_n^i and P_n^j are respectively the transmitted power on sub-carrier n of RRH i and the transmitted power on sub-carrier n of the interfering RRH j that belongs to the set C . $G_{k,n}^j$ is the channel gain between the mobile k and RRH j , the parameter N_0 is the

22. Li, Seshadri, and Ariyavisitakul, *op. cit.*

23. Truman and Brodersen, *op. cit.*

thermal noise power density, and B_{sub} is the sub-carrier spacing²⁴.

To compute the spectral efficiency $\eta_{k,n}$ of UE k on sub-carrier n associated to RRH i , the Shannon's formula is used such as:

$$\eta_{k,n}^i = \log_2\left(1 + \frac{\gamma_{k,n}^i}{\Gamma}\right) \quad (5.5)$$

with Γ , a SNR correction factor that takes into account the difference between the information-theoretic performances and the practical implementation of the MCS²⁵ defined as follows:

$$\Gamma = -\frac{\ln(5.E)}{1.5} \quad (5.6)$$

where E is a BER Target.

For the second part of this chapter, the computation of spectral efficiency of a UE in CoMP mode is defined such as:

$$\eta_{k,n}^{CoMP} = \log_2\left(1 + \sum_{(i=1, i \in C')}^N \frac{\gamma_{k,n}^i}{\Gamma}\right) \quad (5.7)$$

5.4 Interference management with logically centralized decision making

To cope with the ever increasing needs of UEs in terms of throughput and to manage their mobility, a wise interference management that adapts to the context is required. The proposed solution leverage the C-RAN architecture and fully take benefits from the one-to-many mapping. This leads to improve solutions from the literature by dynamically performing ICI management for each PRB according to the channel state of UEs for the cluster of RRH considered.

5.4.1 Contribution: Dynamic cell-less Radio Access Network Meta Scheduler (DC-RAN-MS)

The proposed solution

The DC-RAN-MS relies on a one-to-many mapping within the scope of a "cell-less" approach. In this way, it manages several RRHs at the same time and aims to coordinate

24. Ezzaouia, *op. cit.*

25. Seo and Lee, *op. cit.*

with the intra-cell schedulers. The proposed solution is compatible with the most acknowledged schedulers while emphasizing their specific features (throughput and system capacity for the MaxSNR, system capacity and fairness for the PF, etc.).

The DC-RAN-MS avoids ICI when their magnitude is high by restricting the usage of a sub-carrier n for a given TS. Otherwise, the proposed solution uses the entirety of the spectrum. This decision is taken according to the channel state of the mobile m in each RRH. $\eta_{m,n_{interf}}^i$ is defined as the $\eta_{m,n}^i$ while the UE m of the RRH i is interfered on the sub-carrier n . In this way, the DC-RAN-MS performs its resource allocation following this inequality:

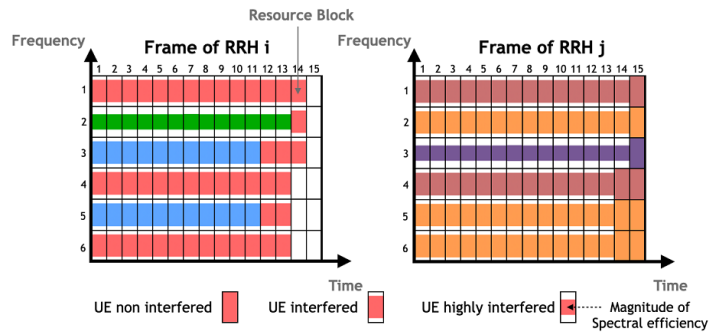
$$\sum_{(i=1, i \in C)}^L \eta_{m,n_{interf}}^i \geq \max \{ \eta_{m,n}^i \}, \forall i \in C \quad (5.8)$$

Thanks to this inequality, DC-RAN-MS can deduce the number of UEs available to transmit (i.e. if they are interfered). Depending on their CSI, the proposed solution performs different frequency reuses. When the magnitude of interference is too high DC-RAN-MS uses a Sliced Bandwidth strategy. Otherwise, if the interference is not significant, it performs a Reuse 1 scheme. Consequently, the DC-RAN-MS does not rely on a criteria in correlation with interference (i.e the distance like state of the art solutions) but directly on the interference themselves.

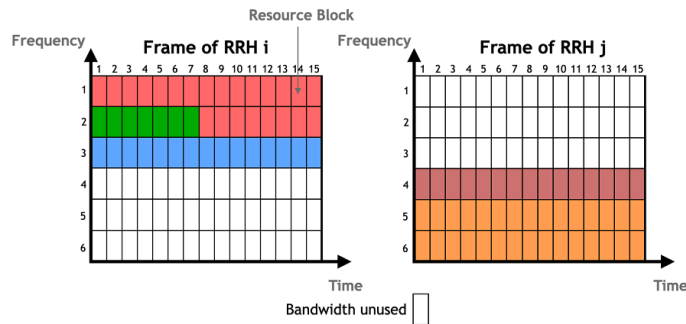
Operation of the DC-RAN-MS

The Figure 5.6 illustrates the four solutions previously described. For each solution, we consider two frames belonging respectively to the RRH i and the RRH j . Within the RRHs, the scheduling algorithm performed is a MaxSNR. The PRBs are allocated TS per TS and denoted by the pair (set of sub-carriers, TS). A UE is embodied by a specified color. The rate of color filling illustrates the magnitude of spectral efficiency regarding the interference received. To provide a realistic scenario, the UEs have different needs in term of application requirements which leads to different PRB allocations between the UEs. The green and purple UEs are considered in cell edges while the others are in the inner. In addition, the red UE has the most largest number of packets to transmit.

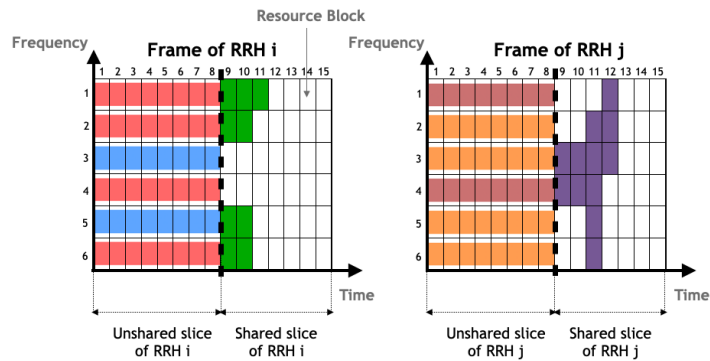
The Reuse 1 strategy (Fig. 5.6(a)) uses the entirety of the bandwidth without attempting to avoid ICI. Consequently, the green and purple UEs are highly interfered and their spectral efficiency is significantly downgraded. Consequently they need much more PRBs than usual to end their transmission. As the resource allocation is performed TS per TS in this example, the red UE takes advantage that others UEs have ended their



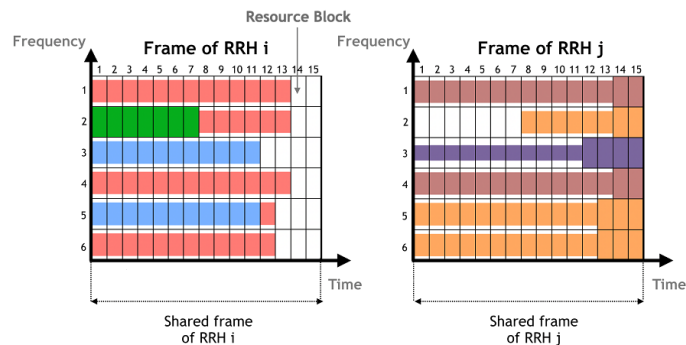
(a) Reuse 1 strategy



(b) Sliced Bandwidth strategy



(c) Hybrid Static strategy



(d) DC-RAN-MS strategy

Figure 5.6 – Interference management solutions with a MaxSNR allocation

transmission to transmit on the sub-carriers they previously used (PRBs : (3,12), (5,12), (3,13), (5,13), (2,14) and ends on the PRB (3,14)). However, considering that the red UE is likely to do not have a good SNR on these sub-carriers and interferes UEs of the RRH j , it is not always efficient that the red uses these PRBs. Depending on the channel state of this UE and the number of PRBs available, it could be more profitable for the system that the red waits for transmitting on sub-carriers 1 and 4. Notice that users from the RRH j are not interfered on the PRBs (4,14), (5,14) and (6,14) as well as on the TS 15 since users from the RRH i have ended their transmission.

The classical ICI avoidance Sliced Bandwidth strategy (Fig. 5.6(b)), splits its bandwidth into two parts. The Sliced Bandwidth strategy protects the green UE which is highly interfered. However, due to the MaxSNR allocation, the purple UE does not have any resource as the channel state of brown and orange UEs is better. In addition, the half of the bandwidth is unused on each RRH which leads to a poor system capacity.

The HS solution (Fig. 5.6(c)) divides the bandwidth into two slices. The left slice is allocated to UEs in inner cell where a Reuse 1 strategy is performed. In the right slice, a Sliced Bandwidth scheme is used to protect cell edge UEs. On this side, the HS solution compares the $\eta_{green,n}^i$ with the $\eta_{purple,n}^j$ and allocates the resource on the sub-carrier n to the UE with the best value. Although this solution attempts to provide a better fairness among UEs by protecting the most affected from interference, its static behavior is not well suitable for wireless network specificities to be efficient. In this example, the Reuse 1 slice is overloaded and the PRBs unused in the Sliced bandwidth part can not be allocated to the Reuse 1 one. Indeed, this solution relies on an ideal distribution between UEs in inner cell and cell edge ones in order to avoid a load asymmetry between the two slices. In addition, even with a moving boundary, as the radio conditions may highly vary, it is not enough flexible to react to the channel state of UEs and to their mobility.

In order to dynamically react to the context, the DC-RAN-MS performs an ICI management that does not rely on the locations of UEs. The Figure 5.6(d) illustrates the resource management of the proposed solution. For instance, in accordance with the inequality (5.8), the result of the sum of $\eta_{green,2interf}^i$ and $\eta_{orange,2interf}^j$ is not enough significant to perform a double allocation, in comparison with a single allocation. In this way, as the green UE has a better channel state than the orange, it starts to transmit on the PRB (2,1) and ends on the PRB (2,7). The purple UE also highly interfered, can transmit on the same TS and sub-carrier than the blue as $(\eta_{blue,3interf}^i, \eta_{purple,3interf}^j)$ is greater than $\max(\eta_{blue,3}^i, \eta_{purple,3}^j)$. This allocation is not possible with state of the art solutions as they

does not take into account the channel state but only the locations of UEs. In addition, the red UE takes advantage that blue and green UEs have ended their transmission to use the following PRBs: (2,8) to (2,13) and (5,12). As the sum of $\eta_{red,n_{interf}}^i$ and $\eta_{orange,n_{interf}}^j$ on sub-carriers 2, 5 and 6 is greater than their respective $\eta_{m,n}$, red and orange UEs are allowed to transmit at the same time. However, red UE can not transmit on the same TS than the purple (i.e. PRB (3,13)) due to the high interference experienced by this last and to the poor channel state.

To conclude, The DC-RAN-MS may perform a wise resource allocation by providing a more accurate interference management according to: the possibility that a UE could not be interfered, the magnitude of interference experienced and their $\eta_{m,n}$, $\eta_{m,n_{interf}}$. This leads to an optimized decision making: either allocating the resources to all UEs when the interference are not significant or only allowing to one UE to transmit when the channel state is poor. Thereby, the DC-RAN-MS may optimize the usage of the bandwidth, reduces the interference experienced and protects the most affected UEs. This may lead to increase the system capacity and the QoS.

5.4.2 Performance evaluation

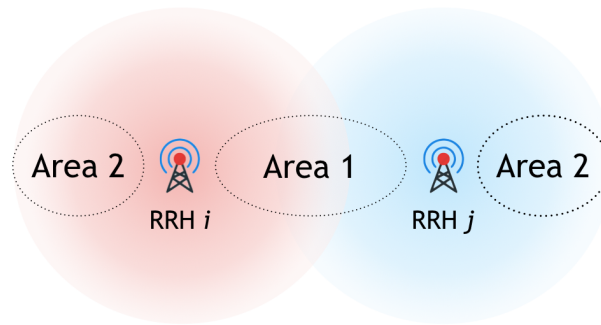


Figure 5.7 – Conceptual representation of the simulation context

Scenarios and Key Performance Indicators

In this section, the DC-RAN-MS is compared to the classical C-RAN strategy, the HS solution, as well as to the classical ICI management strategies: the Reuse 1 and Sliced Bandwidth schemes.

In the simulations, a small network of 2 adjacent cells is considered. This allows to minimize the simulation time while facilitating the analysis of the results. The area 1 is a

zone where the UEs are likely to be highly interfered as they are closer to the neighboring cell than UEs in area 2 (where the interference are less significant). The Figure 5.7 illustrates an example of the position of these areas. The scheduling algorithms of each cell are the same which allows to only study the influence of the interference management solutions. The scheduler chosen is the MaxSNR for its ability to increase the system capacity and as it is one of the most acknowledged scheduler. In addition, the UEs are at the same distance from their respective RRH to neglect the unfair behavior of this scheduler. The traffic generated by sources is considered realistic and follows a Youtube streaming profile of traffic²⁶. This significantly complicates the task of schedulers because they must face high traffic variations at a given time between UEs while ensuring their QoS. Simulations parameters are described in the table 5.1.

The performances evaluation is composed of 2 scenarios. For each scenario, only one parameter varies. The first examines when UEs are equally distributed between area 1 and 2. Thus, the traffic load increases until the congestion with this particular distribution. The second scenario analyzes the performances of the different strategies when the repartition of UEs between the area 1 and 2 varies.

This study focuses on two KPI to evaluate each solution:

- **The spectral efficiency** is the mean number of bits received on each PRB used (bits/PRB).
- **The mean packet delay** is the mean delay to transmit one packet (in ms).

Parameters	Value
Cell Radius	500 m
Number of sub-carriers	32
Number of Time Slots	10
RRH transmit power	20 W (43 dBm)
Standard deviation of shadowing	$\sigma = 8$ dB
Path-loss exponent	3.5 (urban context)
Target BER	5×10^{-5}
Sub-carrier spacing	15 kHz
Thermal noise power density (N_0)	- 174 dBm/Hz

Table 5.1 – Simulations parameters.

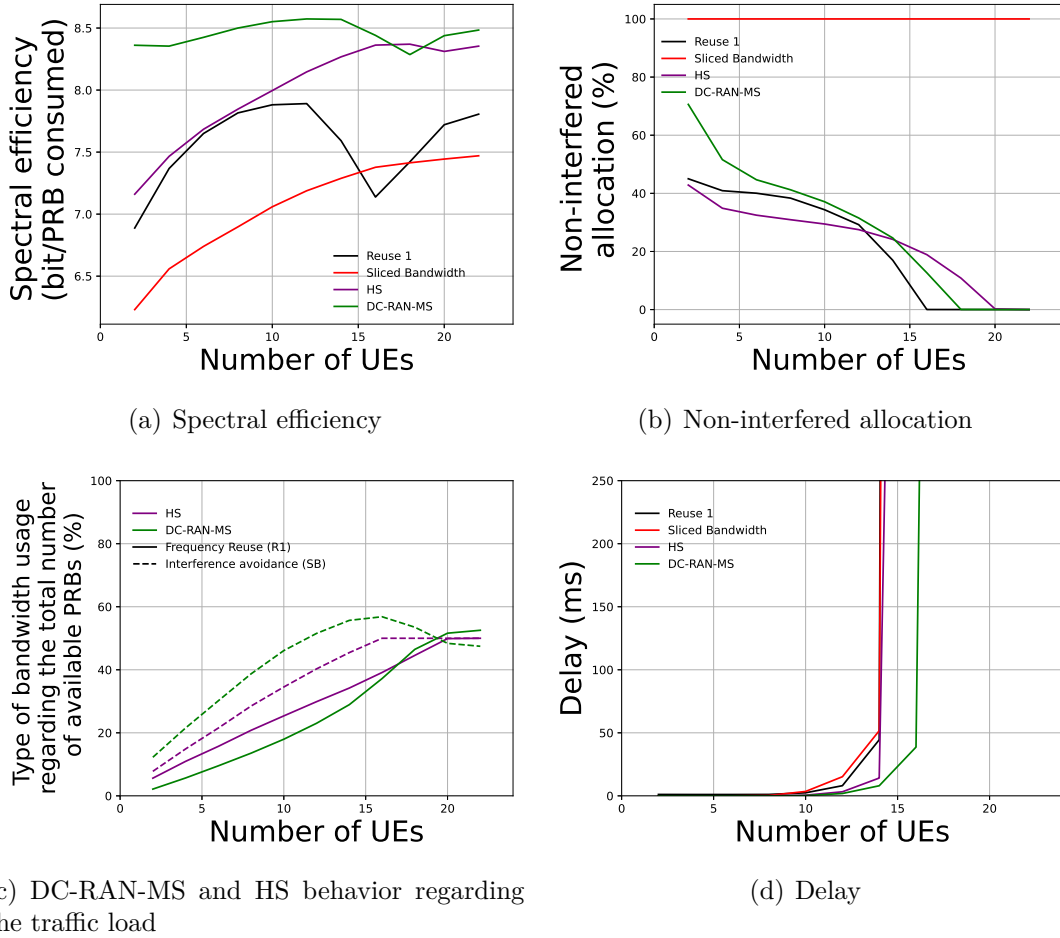


Figure 5.8 – Scenario 1

Scenario 1

In this scenario, considering the two cells, the UEs are equally distributed between the area 1 and 2 such as: 25% are located in area 2 of the RRH i , 50% in the area 1 and 25% in area 2 of the RRH j . In this way, UEs are added two per two in each cell (one per area).

Spectral efficiency: figure 5.8(a) shows the spectral efficiency obtained with each solution for different traffic load in the system. The spectral efficiency presented takes into account the two cells considered. It means that it is the average of bit per PRB consumed. A PRB is considered as consumed either when it is used by the both RRHs or when an ICI management prevents a RRH from allocating the PRB in order to protect a

26. Horvath and Fazekas, *op. cit.*

UE in the other cell. Computing the spectral efficiency only allocated PRBs is misleading because it does not take into account PRBs unused to avoid interference. For instance, given a RRH i and a RRH j which have only one UE, respectively the UE m_i and the UE m_j . In this example it is considered that they both get the same $\eta_{m,n}$ and $\eta_{m,n_{interf}}$ equal to 14 and 6.5. With a Reuse 1 solution their classical spectral efficiency (bits/PRB) is equal to 6.5 while it is equal to 14 with a Sliced Bandwidth scheme as only the UE m_i is allowed to transmit (according to the frequency reuse). However, this last value does not take into account the PRB of the RHH j unused. To provide a more accurate indicator of the bandwidth waste, the average of bits per PRB consumed seems to be more appropriate to compute the spectral efficiency. Thus, with the previous example, the spectral efficiency (bits/PRB consumed) of m_i and m_j is equal to 6.5 with a Reuse 1 solution while for m_i it is equal to 7 with a Sliced Bandwidth strategy ($14 \div 2$, where 2 is the number of PRBs consumed).

The Figure 5.8(b) shows the percentage of non-interfered allocation for different traffic loads. An interference appears when a UE of the RRH i and a UE of the RRH j are chosen to transmit on the same frequency and time. In underloaded context, non-interfered allocations occur more frequently compared to a loaded system. Thus, this Figure shows that the UEs become more and more interfered as the traffic load increases. Indeed, it is rare that only one UE is chosen to transmit on a given PRB among the RRHs since their bandwidth become more and more filled. Notice that when the percentage of non-interfered allocation reaches zero, it means that there is no more PRB available without being interfered.

The Figure 5.8(c) shows the type of bandwidth usage regarding the total number of available PRBs for different traffic load. It illustrates the usage ratio of the different ICI management (either Reuse 1 or Sliced Bandwidth) performed by the HS and DC-RAN-MS solutions.

On the Figure 5.8(a), the solutions adopt at low traffic load, a typical behavior of a MaxSNR allocation which benefits from the multi-user and frequency diversities. In this way, their spectral efficiency increases with the traffic load. However, depending on the ICI management, interference effects on this metric highly vary.

The Sliced Bandwidth avoids interference and keeps the opportunistic MaxSNR behavior. Nevertheless, as the part of unused bandwidth is high, it provides a poor spectral efficiency. Notice that with a classical spectral efficiency computation (Bits/PRB used) the result is significantly different.

From 12 UEs to 16, the Reuse 1 solution experiences a high degradation of its spectral efficiency due to the decrease of the number of the allocations non-interfered (Fig. 5.8(b)). At 16 UEs, the number of possible interference has reached its peak: all the PRBs are interfered and the system is overloaded. Since there are two phenomenons that face each other (the MaxSNR opportunistic behavior and the magnitude of interference), from 16 UEs the MaxSNR allocation can counterbalance interference effects as it favors the less interfered UEs. This leads to increase the spectral efficiency of the Reuse 1 solution.

HS and DC-RAN-MS solutions are less affected by interference as they provide a better ICI management.

The HS solution puts the cell edges UEs (i.e. area 2) in its right part (Fig. 5.6(c)) where a Sliced Bandwidth is performed. Consequently, only UEs in inner cells are affected by interference. As they are less interfered, the magnitude of ICI experienced is not significant making the spectral efficiency less affected by interference (Fig. 5.8(a)).

The DC-RAN-MS dynamically performs its ICI management thanks to appropriate allocations according to the channel state (Inequality (5.8)). In addition, unlike the HS solution, the DC-RAN-MS does not restrict the usage of the ICI management schemes (Reuse 1 or Sliced Bandwidth) to 50% of its bandwidth (Fig. 5.8(c)). This provides a better usage of the bandwidth which leads to a high spectral efficiency (Fig 5.8(a)). **and DC-RAN-MS behaviors and system capacities:** the system capacity provided by a solution does not only rely on its spectral efficiency. For instance, the HS solution has a better spectral efficiency at a traffic load of 18 UEs (Fig 5.8(a)). As only UEs in inner cell are affected by interference, the percentage of mono allocation reaches 0% at a traffic load of 20 UEs (Fig. 5.8(b)). However, this is not relevant from the system capacity provided by the HS solution. The Figure 5.8(c) shows that the interference avoidance (i.e. Sliced Bandwidth) part of the HS solution is overloaded since a while (it reaches 50% of the available bandwidth at a traffic load of 16 UEs). Thus, as the boundary of the HS solution is static, it can not take advantage of the unused PRBs in the Reuse 1 slice to allocate them to the Sliced Bandwidth part. This leads to an asymmetry load between UEs in inner and edges of the cell.

Unlike the HS solution, the DC-RAN-MS can adapt its behavior to the context. When the system is underloaded, the DC-RAN-MS performs more Sliced Bandwidth allocation than Reuse 1 ones as only one UE is allowed to transmit with this kind of frequency reuse (Fig. 5.8(c)). As the DC-RAN-MS does not rely on a boundary, it can allocate the PRBs unused by the Reuse 1 strategy to perform its Sliced Bandwidth allocations. This

provides a better system capacity than the HS solution for UEs in inner cell and edges without an asymmetry.

Mean packet delay: a crucial indicator of the QoS experienced by UEs is the latency. Figure 5.8(d) represents the mean packet delay in the system in milliseconds.

The Sliced Bandwidth strategy provides a poor system capacity as a half of the bandwidth is unused. Thus, even if when UEs are not interfered, they will quickly experience an increase of their mean packet delay due to the poor available PRBs number as the traffic load rises.

The Reuse 1 solution uses the total available bandwidth. However, as there is no interference management, the most interfered mobiles are highly penalized which leads to provide a poor global QoS.

The HS strategy attempts to protect the most interfered mobiles by dividing the bandwidth. Nevertheless, as only one mobile is allowed to transmit on the Sliced Bandwidth part, this side is quickly overloaded (Fig. 5.8(c)). As the boundary is static, the Reuse 1 part can not be used to provide a better system capacity and the PRBs unused stays wasted until this part is overloaded too. Consequently it can better handle the QoS of mobiles than the two previously described solutions, but overall, the system collapses at the same traffic load than others solutions (i.e. at 14 UEs).

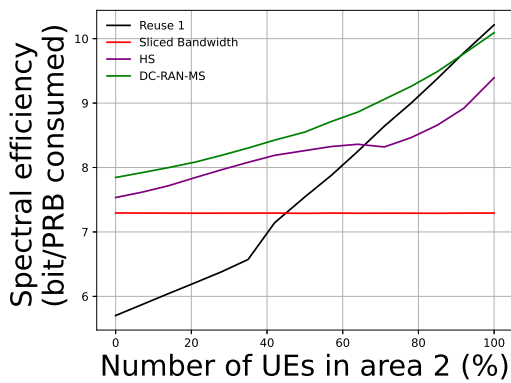
The DC-RAN-MS optimizes the spectrum usage (Fig. 5.8(a)) and allows to reduce the magnitude of interference. The proposed solution performs a Sliced Bandwidth strategy only when its necessary (Fig. 5.8(c)). This reduces the waste of bandwidth and provides a more accurate and fairer resource allocation. Consequently a better QoS is experienced and the system can handle the traffic load longer (i.e. at 16 UEs) than others solutions (Fig. 5.8(d)).

Scenario 2

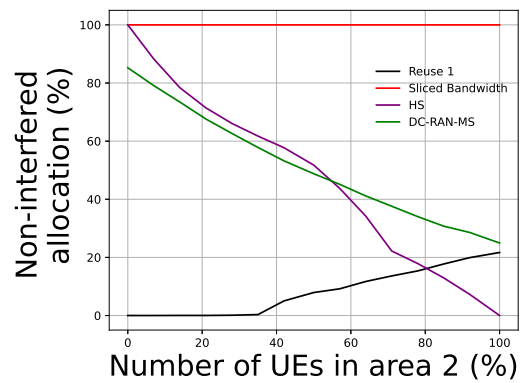
This scenario analyzes the behavior of the solutions when UEs distribution varies between the areas 1 and 2 for a traffic load of 14 UEs. First, UEs are located in the area 1. Then they move one by one to the area 2.

Spectral efficiency and system capacity: Sliced Bandwidth strategy has a constant spectral efficiency (Fig. 5.9(a)) as this solution manages the interference by dividing its bandwidth. In this way, its spectral efficiency depends only on the traffic load and not on UEs distribution (i.e magnitude of interference experienced)

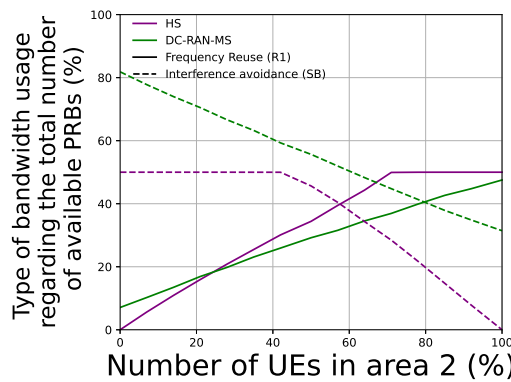
The Reuse 1 solution provides a poor spectral efficiency (Fig. 5.9(a)) when UEs



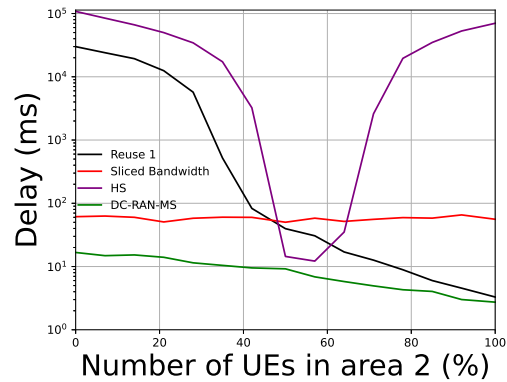
(a) Spectral efficiency



(b) Non-interfered allocation



(c) DC-RAN-MS and HS behavior regarding the mobile distribution



(d) Delay

Figure 5.9 – Scenario 2

are highly interfered (i.e located in area 1). Thus, UEs need much more PRBs than usual to end their transmission which leads to consume all the PRBs available and an overloaded system as the ratio of non-interfered allocation reaches 0% (Fig. 5.9(b)). Like a virtuous circle, the rise of the spectral efficiency increases the percentage of non-interfered allocation. As the number of PRBs interfered decreases, it provides a better spectral efficiency and so on.

The HS and DC-RAN-MS schemes provides a better ICI management leading to a better spectral efficiency.

However, as the HS solution is static, there is an asymmetry between the Sliced Bandwidth part and the Reuse 1 one (Fig. 5.9(c)). Until 42% of UEs in the area 2, the Sliced Bandwidth part is overloaded and the spectral efficiency increases thanks to the opportunistic MaxSNR behavior. Then, the number of UEs interfered rises which decreases the percentage of non interfered allocation. At 71% of UEs in area 2, the Reuse 1 slice is overloaded (Fig. 5.9(c)) which means that the influence of interference has reached its peak. Thus, the MaxSNR prioritizes less interfered UEs and counterbalances the ICI effects. In this way, the spectral efficiency increases from 71% until 100% of UEs in area 2. Notice that the percentage of non-interfered allocations reaches 0% only when UEs are all located in area 2 as before, PRBs are still allocated in the Sliced Bandwidth part.

Unlike the HS solution, the DC-RAN-MS has been designed to dynamically react to the context. The DC-RAN-MS does not take into account the location of UEs but only their channel state. This allows to be even more accurate than a solution with a dynamic boundary. When the UEs are mostly in the area 1 (i.e highly interfered) it uses mainly the Sliced Bandwidth strategy (Fig. 5.8(c)). Nevertheless, it can performs a Reuse 1 strategy provided the radio conditions of UEs are enough significant. The same analysis can be apply when UEs are mostly located in area 2 (i.e. less interfered) where the percentage of Sliced Bandwidth strategy used is still important. Consequently, the DC-RAN-MS provides a better spectral efficiency than other solutions (Fig. 5.9(a)) and a better system capacity.

Mean packet delay: Sliced Bandwidth provides a constant delay (Fig. 5.9(d)) as its spectral efficiency is the same whatever the distribution of UEs.

The mean packet delay provided by the Reuse 1 solution highly depends on the UE locations (i.e. magnitude of interference received). In this way, the UEs have a better QoS when they are located in the area 2 than in the area 1.

The HS solution is designed with a static boundary. This leads to a poor QoS expe-

rienced by UEs except when they have an ideal distribution (between 42% and 61%).

The DC-RAN-MS dynamically adapt its behavior to the channel state of UEs regardless of their location. Thanks to a better spectral efficiency and system capacity provided by the inequality (5.8), the proposed solution outperforms state of the art schemes on this metric.

5.4.3 Conclusion

The continuously growing needs in term of throughput, system capacity and delay requirements lead to search new manner to optimize transmission efficiency particularly in ultra dense networks. In this way, the contribution of this work named DC-RAN-MS, aims to merge the ICI management and the scheduling processes to optimize the resource allocation. Thanks to a cell approach based on the C-RAN architecture, the proposed solution allocates resources for a cluster of RRHs. The proposed solution does not rely on a criteria in correlation with ICI (like the distance) but directly on the interference themselves. According to the channel state of UEs, the proposed solution performs dynamically its ICI management either by: allowing the usage of the entire bandwidth or by preventing schedulers from allocating resources when the magnitude of interference experienced are too high. This leads to optimize the usage of the spectrum which decreases the global ratio of unused bandwidth while efficiently reducing the magnitude of interference. This results in higher spectral efficiency, higher system capacity and a QoS increased.

5.5 Interference management with Coordinated Multi-Point

In this section a new interference management is introduced. The Hybrid Joint-Transmission Coordinated MultiPoint (H-JT-CoMP) is an extension of the previous solution that uses CoMP. For each PRB, the proposed solution dynamically performs either R1 either JT-CoMP. We assume that this could enhance the spectral efficiency especially at cell edges.

5.5.1 Contribution: Hybrid Joint-Transmission CoMP (H-JT-CoMP)

The H-JT-CoMP clustering is based on a Cell-less approach. Consequently, it can manage several RRHs at the same time and aims to coordinate with the intra-cellular scheduling. The proposed solution operates its ICI management after the scheduling. When interference occurs, the H-JT-CoMP clustering performs either a JT-CoMP or a Reuse 1 strategy based on CSI according to this inequality :

$$\sum_{(i=1, i \in C)}^L \eta_{m_i, n_{interf}}^i \geq \max \{ \eta_{m_i, n}^{CoMP}, \forall i \in C \} \quad (5.9)$$

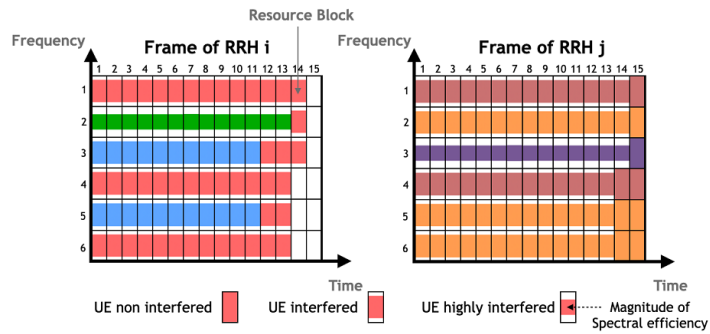
The H-JT-CoMP clustering compares the sum of interfered radio conditions of all mobiles selected in the set of cells C with the maximal value of their radio conditions in CoMP mode. Based on this inequality, the proposed solution will perform a different ICI management strategy, defined such as:

$$\text{ICI Strategy} = \begin{cases} \text{Reuse 1} & \text{if Equation 5.9 is true} \\ \text{CoMP mode} & \text{else} \end{cases} \quad (5.10)$$

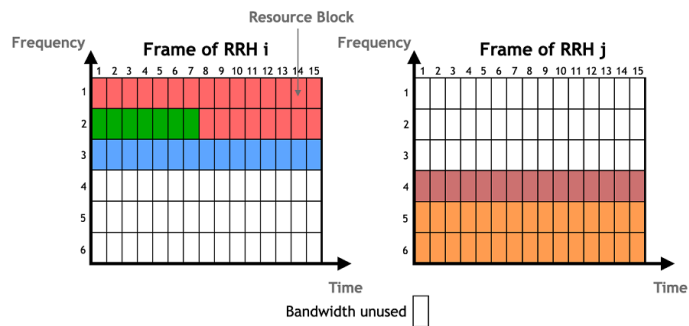
When H-JT-CoMP clustering uses a Reuse 1 strategy it means that radio conditions of UEs selected by the intra-cellular scheduler are enough significant from the system point of view to let interference to occur while CoMP mode is favored when UEs received important magnitude of interference. Thanks to this dynamic approach based on the CSI, the proposed solution optimizes the radio resource usage of the system while efficiently mitigating the ICI received. Moreover, no additional overhead is required as CSI is already computed by opportunistic intra-cellular schedulers (such as the MaxSNR) and cell coordination is performed by optical fibers. This provides an increase in system capacity and spectral efficiency, a better fairness between UEs at inner and edges of the cell leading to an overall increase in the QoS.

5.5.2 Operation mode

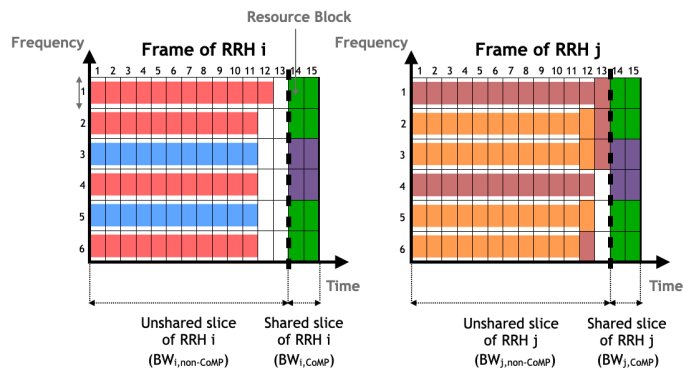
The Figure 5.10 illustrates some previously described solutions. For this example, a cluster of two cells is considered with their associated frames, respectively i and j . The user selection algorithm is a MaxSNR allocation. PRBs are allocated TS per TS and



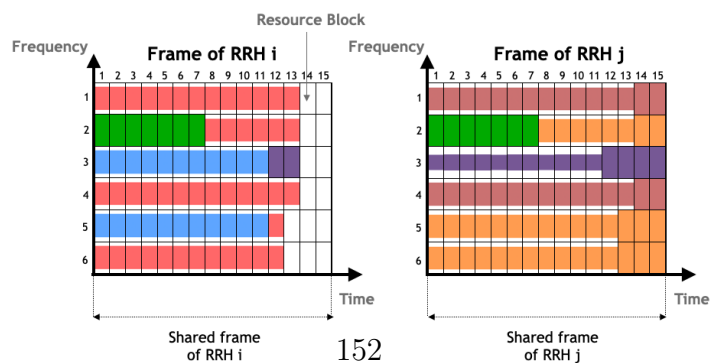
(a) Reuse 1 strategy



(b) Sliced Bandwidth strategy



(c) PLDCoMP strategy



(d) H-JT-CoMP strategy

Figure 5.10 – Interference management solutions with a MaxSNR allocation.

denoted by (set of sub-carriers, time slot). Three UEs are associated to each cell and are embodied by a specific color. In the RRH i , blue and red UEs are at inner of the cell, while green UE is at edge. In the RRH j , brown and orange UEs are at inner of the cell, while purple UE is at edge. The rate of color filling in each frame, illustrates spectral efficiency variations regarding interference received. Mobiles have different needs in term of application requirements which leads to different number of PRBs required to end the transmission between the UEs. For instance, the red UE has the most packets to transmit.

In the Figure 5.10(a), the Reuse 1 operation mode is presented. As this solution allows interference to occur, UEs at inner of cells have a decent spectral efficiency while edge UEs (green and purple) are highly impacted by interference. Note that the purple UE is not interfered on the PRB (3,15) as nobody is transmitting on this PRB in the RRH i .

The Figure 5.10(b) introduces the Sliced Bandwidth strategy. As a cluster of 2 cells is considered, this solution splits its bandwidth into two parts. This allows to avoid interference at the expense of high bandwidth waste. In addition, the purple UE in frame j , has not been selected by the MaxSNR allocation and cannot transmit during this frame.

In the Figure 5.10(c), the PLDCoMP strategy is presented. Green and purple UEs are at edges, they are considered in CoMP mode, while others are in non-CoMP mode. In this example (Fig. 5.10), in each cell 90 PRBs are available (15×6) and 1 UE is in CoMP mode while 2 UEs are in non-CoMP mode. According to the equation 5.1, the number of PRBs allocated to UEs in non-CoMP mode is 80 while only 10 PRBs are allocated to CoMP mode UE²⁷. Consequently, the PLDCoMP hardly restricts the usage of the bandwidth to edge UEs (green and purple). Note that with a JT-CoMP, a same PRB in both RRH is reserved for a given UE. For instance, on PRB (1,14) in frame i and j the PRB is allocated to the green UE.

The Figure 5.10(d) illustrates the H-JT-CoMP operation mode. The ICI management is performed for each PRB. As the green UE is highly interfered on sub-carrier 2 by the RRH j which transmits to orange UE, the H-JT-CoMP puts the green UE in CoMP mode. On sub-carrier 3, purple and blue UEs have enough significant radio conditions to be let in Reuse 1. When the blue UE end its transmission (after TS number 11), the red UE is able to transmit on sub-carrier 3. This time, the SINR of purple and red UEs are poor which induces a CoMP mode for purple UE on PRBs (3,12) and (3,13). Thus,

²⁷. For presentation matters, the bandwidth is split by time slot. In the Figure 5.10(c), the PLDCoMP allocates 12 PRBs to edge user instead of 10

the HT-CoMP performs a more accurate ICI management, allowing to efficiently mitigate interference while optimizing the spectrum usage. This leads to protect cell edge users when the magnitude of interference received is too high, increasing their QoS.

5.5.3 Performance evaluation

Parameters	Value
Cell radius	500 m
Number of sub-carriers	32
Number of time Slots	10
RRH transmit power	20 W (43 dBm)
Standard deviation of shadowing	$\sigma = 8$ dB
Path-loss exponent (α)	3.5 (urban context)
Target BER	5×10^{-5}
Sub-carrier spacing	15 kHz
Thermal noise power density (N_0)	- 174 dBm/Hz
Simulation duration	500 000 frames

Table 5.2 – Simulations parameters.

In the simulation, a network of two adjacent cells is considered. According to section 5.3, $L = 1$ and $N = 2$. The proposed solution is compared to the classical ICI schemes, respectively Reuse 1 (R1) and Sliced Bandwidth (SB) solutions, as well as to the PLDCoMP clustering. The intra-cellular scheduling solution used in both cells is a MaxSNR allocation for its ability to increase the system capacity and as it is one of the most acknowledged scheduler²⁸. The traffic generated by sources is considered realistic and variable which produces high volume of data with important sporadic and tight delay requirements²⁹. This significantly complicates the task of resource allocation schemes. Simulations parameters are described in the table 5.1.

Scenarios and Key Performance Indicators

Two deployments scenarii are provided. The first scenario is a proof of concept. Considering the MaxSNR unfair behavior regarding the distance of mobiles from their access point, UEs are at the same distance from their access point. UEs are split into two groups : the first is far from the neighboring cell and is likely not interfered. The second group

28. Bechir Dadi and Belgacem Chibani, *op. cit.*

29. Horvath and Fazekas, *op. cit.*

is close to the neighboring cell and receive high magnitude of interference. This scenario allows to study the behavior of each solution in a simple context.

In the last scenario, users are uniformly distributed inside cells and solutions are studied for a given traffic load, when all solutions are close to congestion to let appear some packet delay.

This work focuses on four Key Performance Indicators (KPIs) to evaluate the performance of each solution. These KPIs have already been used in previous chapters:

The bandwidth consumption ratio is the ratio between the number of RUs used by a solution and the total number of PRBs available. When this ratio is equal to 100% it means that the system is congested.

The spectral efficiency is the mean number of bits received on each PRB used. In this work, the spectral efficiency takes into account both cells and is computed as the mean number of bits on each PRB consumed. Computing the spectral efficiency only on PRBs allocated is misleading because it does not take into account PRBs unused to avoid interference (like SB). For instance, given an UE with a $\eta_{m,n}$ equals to 14 with a SB strategy (avoiding ICI), its classical spectral efficiency is 14 (as only 1 PRB is used) while its spectral efficiency (bits/PRBs consumed) used in this work is 7 (as 2 PRBs are consumed : one for transmitting and one for avoiding ICI). This provides a more accurate indicator on the bandwidth waste.

The mean packet delay is the mean delay to transmit one packet (ms).

Jain's fairness index corresponds to the Jain's fairness index on delay. In this work, this KPI is only provided for scenario 1.

Scenario 1

In this scenario, UEs are split into two groups. One group is far from neighboring cell and is likely not interfered. This position is referenced as area 1. The second group is close to neighboring cell and receives high magnitude of interference. This position is referenced as area 2. To neglect the unfair behavior of MaxSNR regarding the distance, UEs are at the same distance (500 m) from their access points regardless their group. Mobiles are added in each cell two by two (one per area). This scenario studies the performance of each solutions according to the traffic load increase.

The Figure 5.11(a) shows the spectral efficiency of each solution. The Sliced Bandwidth strategy avoids interference and has a typical MaxSNR allocation behavior, taking benefits from the multi-user diversity which increases spectral efficiency as the traffic load

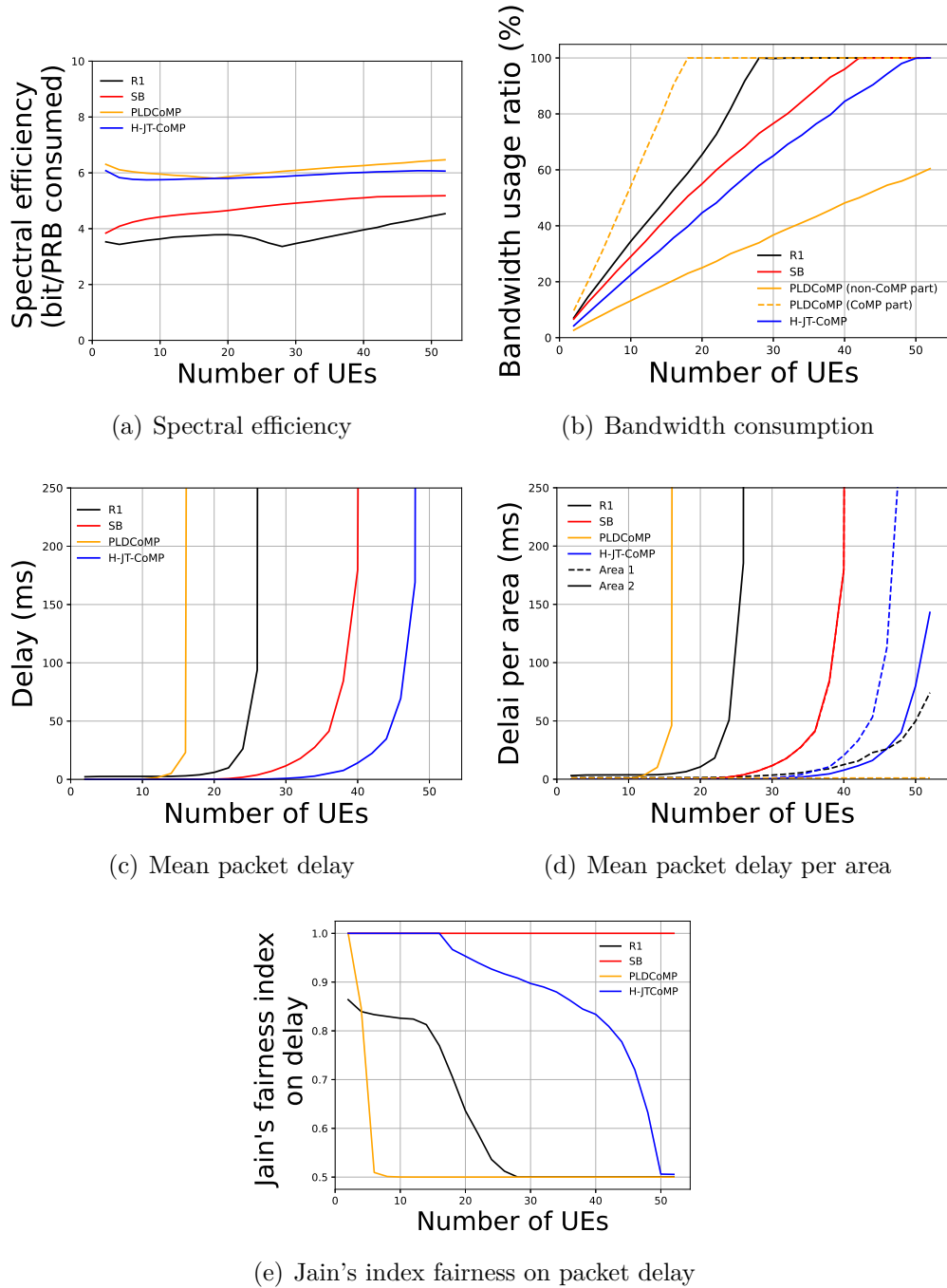


Figure 5.11 – Scenario 1

rises. Unlike this solution, the Reuse 1 lets interference to occur. From 2 users to 20, it takes benefits from the multi-user diversity. At a traffic load of 22 users, the number of PRBs without interference is decreasing as well as the spectral efficiency, highlighting that the system is close to be congested. At 28 users, all PRBs are interfered and the magnitude of interference has reached its peak. The system is congested since the bandwidth consumption ratio of the R1 solution reached 100% (Fig. 5.11(b)). Then, the MaxSNR behavior counterbalances ICI effect by taking benefits from the multi-user diversity. This leads to increase the spectral efficiency after the system congestion. Concerning solutions performing CoMP, PLDCoMP and H-JT-CoMP, their spectral efficiencies decrease at the beginning, as the number of PRBs interfered rises. Indeed, this spectral efficiency takes into account the PRBs consumed. As these solutions use CoMP to mitigate ICI, they consume more PRBs (2 PRBs reserved for the same user due to the JT-CoMP). However, the PLDCoMP splits its bandwidth into 2 parts : one for UEs in CoMP mode and the other for UEs in non-CoMP mode. This induces there is never a UE interfered (area 2) and a UE non-interfered (area 1) scheduled at the same time and on the same frequency. This avoids the situation where the UE interfered has a poor SINR while the UE non-interfered has a great SNR. Consequently, this provides a great spectral efficiency. Note that at a traffic load of 18 UEs, the bandwidth allocated to UEs in CoMP mode is congested (Fig. 5.11(b), Equation. 5.1). As the PLDCoMP cannot serve more UEs in CoMP mode while there are PRBs available for UEs in non-CoMP mode, the spectral efficiency increases. Indeed, non-CoMP mode UEs which are not interfered (i.e. have better radio conditions), will be more important in the spectral efficiency computation than CoMP mode UEs. The H-JT-CoMP clustering provides a spectral efficiency close to PLDCoMP solution results. Considering that the proposed solution does not segregate UE depending on their position neither limiting their bandwidth usage, the H-JT-CoMP provides also a great spectral efficiency.

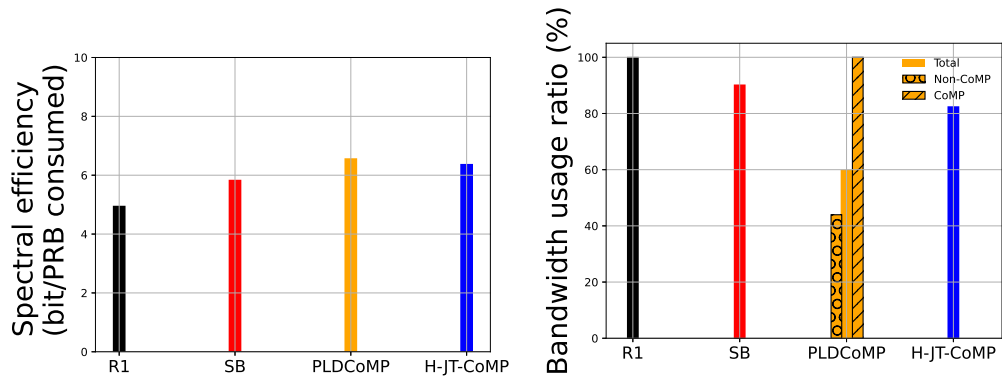
The Figure 5.11(b) shows the bandwidth consumption ratio of each solution. According to the spectral efficiency (Fig. 5.11(a)), the R1 strategy is the first to provide a congested system, then the SB scheme and the H-JT-CoMP solution. Concerning the PLDCoMP, after 50 users in the system, this solution has still PRBs available but only for UE in non-CoMP mode since its bandwidth part allocated to CoMP mode UE is overloaded since a while (18 UEs).

Figures 5.11(c), 5.11(d) and 5.11(e) illustrate the mean packet delay, the mean packet delay per area and the Jain's index fairness on delay, respectively. This last metric is

computed with the mean packet delay of each group. This induces the limit value to be equal to 0,5. Due to its poor spectral efficiency and system capacity, the Reuse 1 solution provides a poor QoS to UEs. In addition, using Reuse 1 strategy to manage interference is highly unfair (Fig. 5.11(e)) regarding the UE position. In this way, even if UEs less interfered (area 1) have a decent QoS, UEs in area 2 receive high magnitude of interference which leads to a poor QoS. The Sliced bandwidth, thanks to a better spectral efficiency, provides a better QoS than R1 solution. Its ICI avoidance scheme allows, in this context, to have a perfect fairness among UEs (Fig. 5.11(e)). Nevertheless, it induces a high bandwidth waste. The system capacity is limited and the mean packet delay of UEs rises quickly. The PLDCoMP solution provides a poor QoS to users due to the inequality of resources available between UEs in non-COMP mode and CoMP mode. Thus, the CoMP part is quickly overloaded (Fig. 5.11(b)) leading to a poor QoS for UE in CoMP mode (Figs.5.11(d),5.11(e)). The H-JT-CoMP solution outperforms other solutions on this KPI. Thanks to a wise ICI management, interference received at edges are mitigated while the spectrum is efficiently used (Fig. 5.11(b)). This leads to increase the QoS of users (Fig. 5.11(c)), even for the most interfered. As UEs in area 2 are closer to neighboring cell than UEs in area 1, the CoMP is much more efficient for them. According to the equation 5.9, if a UE in area 1 (less interfered) and a UE in area 2 (more interfered) are selected respectively on RRHs i and j to transmit with CoMP, the H-JT-CoMP will favor the UE in area 2. This explains that UEs in area 2 have a better QoS than UEs in area 1 (Fig. 5.11(d)). However, the H-JT-CoMP provides a decent fairness between the two groups of UEs (Fig. 5.11(e)).

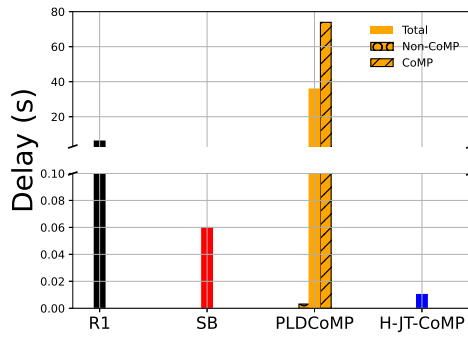
Scenario 2

In this scenario, users are uniformly distributed in cells. Solution performances are studied for the same fixed traffic load where all solutions are experiencing difficulties to ensure UE application requirements. This allows to let appear delay on packets in order to compare solutions. In this context, UEs are at different positions. Consequently, UEs at edges are impacted both by a high magnitude of interference received and the unfair behavior of MaxSNR allocation regarding the distance from the access point. Note that since Reuse 1 and the CoMP part of PLDCoMP solution are congested, the mean packet delay depends on the simulation duration. In addition, for the PLDCoMP solution, mean packet delay and bandwidth usage ratio are also given depending on the UE modes in order to explain the behavior of this solution.



(a) Spectral efficiency

(b) Bandwidth consumption



(c) Mean packet delay

Figure 5.12 – Scenario 2

According to results in the scenario 1, the R1 strategy provides a weak spectral efficiency to the system (Fig. 5.12(a)) leading to a poor system capacity (Fig. 5.12(b)). Thus, the QoS of UEs is highly degraded (Fig. 5.12(c)). The SB solution thanks to its ICI avoidance provides a decent spectral efficiency (Fig. 5.12(a)) leading to increase the QoS of UEs compared to Reuse 1 solution (Fig. 5.12(c)). The PLDCoMP, is unfair regarding the mode of UEs. This induces a poor number of resource available for UEs in CoMP mode (Fig. 5.12(b)) and a poor QoS, especially for CoMP-mode UEs (Fig. 5.12(c)). The proposed solution provides a spectral efficiency close to PLDCoMP performance (Fig. 5.12(a)) thanks to its ICI management based on CSI. This leads to increase the system capacity (Fig. 5.12(b)) and to outperform other solutions on delay metric by highly increasing QoS of UEs (Fig. 5.12(c)).

5.5.4 Conclusion

The contribution of this work, Hybrid Joint-Transmission Coordinated MultiPoint clustering (H-JT-CoMP) performs its ICI management to mitigate interference at edges while efficiently optimizing spectrum usage. In order to reduce interference at edges, it relies on the Joint-Transmission CoMP. For each Resource Unit, the proposed solution dynamically performs either a Reuse 1 or either a JT-CoMP strategy according to the Channel State Information and the magnitude of interference received. Performance evaluations emphasize this solution increases the spectral efficiency and system capacity while efficiently reducing interference. This leads to a QoS increased and a better fairness between users at inner and edges of cell.

5.6 Conclusion

The first new solution introduced in this chapter, DC-RAN-MS, relies on a C-RAN architecture and particularly on a one-to-many mapping. This allows to leverage the centralized decision making to gather user data and scheduling information from neighboring cells. Thanks to this information, the proposed solution dynamically manages the bandwidth by preventing or allowing interference to occur according to the CSI of UEs. It means that, based on the SINR of UEs, the DC-RAN-MS performs either R1 or SB. Consequently, the ICI management is carried out in an opportunistic way and not on fixed criterion (e.g. distance from the gNB). This leads to outperform solution from the

literature in terms of spectral efficiency, system capacity and delay.

The second new solution introduced in this chapter, H-JT-CoMP is an extension of DC-RAN-MS. It relies on the use of the JT-CoMP which is a more efficient technique to manage interference at cell edges than SB solution. Indeed, gNB can simultaneously transmit the same user data to an UE on the same radio resource, avoiding interference and increasing the throughput. This is done at the expense of a higher bandwidth consumption and higher signalling between gNBs.

To conclude, this chapter emphasizes that the combination of a dynamic PRB management in a cell context with an opportunistic approach allows to significantly improve the system performance in terms of spectral efficiency, system capacity and delay. The results obtained with H-JT-CoMP show that the use of JT-CoMP provides a slightly better spectral efficiency than DC-RAN-MS solution. However, these results should be seen in the context of the CoMP implementation that requires information exchange between gNBs as well as coordination to transmit data. If the ratio between the efficiency and complexity of a solution is not done, then the performance provided by the JT-CoMP is disappointing. Nevertheless, a new emerging context named Cell-Free Massive MIMO combines the benefits of JT-CoMP and mMIMO. In such context, antennas are distributed around UEs and clusters are created according to user needs. This allows to improve the benefits of JT-CoMP as it is possible to dynamically manage the whole network and for instance turn off parts of the network that are unused.

CONCLUSION

The main objective of this chapter is to give the reader an overview of this thesis. The context and the main issues of current 5G wireless networks are reminded in the first section. The second section summarizes the works and draws up the assessment of the main results. This chapter ends with the third section where improvement and perspectives of this thesis are presented.

6.1 Context

The ever-increasing needs of users in terms of throughput and latency have led mobile networks to evolve. 5G brings noteworthy improvements whether in the core network (e.g. new architecture allowing to perform network slicing) or in the RAN (e.g. new radio interface). Regarding the RAN, 5G enables the use of higher frequencies and the use of different numerologies. This allows to reach higher throughput and to reduce the latency. These improvements will probably create new user needs that will lead to the emergence of new type of services or applications. This will require to increase spectral and energy efficiencies and system capacity to ensure an adequate level of QoS. Network operators cannot only rely on these enhancements nor on cell densification to face these challenges in the coming years.

6.2 Summary of the main results

In this context, this thesis focuses on resource allocation to provide solutions that aim to improve the aforementioned metrics such as: spectral and energy efficiencies, QoS and latency. In particular, our works take benefits from the multi-user diversity to provide scheduling and routing solutions. Moreover, in combination with the Joint-Transmission Coordinated MultiPoint, this also allows the design of efficient interference management solutions.

6.2.1 Multi-user diversity

Chapter 2 introduces the multi-user diversity principle. The multi-user diversity can be defined as the pool of UEs that can transmit or receive data at a given time. This principle is well known but in the literature, it is often considered and mentioned as a consequence of the radio resource allocation. The novelty of this contribution lies in the in-depth analysis of the multi-user diversity, where this concern is addressed as a fully fledged issue. We first show that voluntarily increasing or decreasing the multi-user diversity allows to vary the system performance in terms of spectral and energy efficiencies. Then, this work emphasizes that wisely designing resource allocation solutions that take into account the multi-user diversity allows to increase the overall performance of such solutions. For instance, designing scheduling solutions which segregate UEs according to their class of traffic or their QoS requirements may lead to a poor spectral efficiency and system capacity. In addition, this contribution highlights that the way or the values of some simulation parameters that are chosen when the resource allocation solutions are evaluated are not harmless. For example, the service profile used to simulate traffic of UEs has a major impact on multi-user diversity and can lead to misleading analysis of simulation results.

6.2.2 Scheduling

Chapter 3 introduces the scheduling principle. A scheduler allocates PRB to UEs according to their needs and the available bandwidth. Based on the analysis of the multi-user diversity in Chapter 2, we designed a new scheduler. This solution known as Fair Energy efficient scheduler for high system Capacity (FEC), directly influences the multi-user diversity. Thanks to a QoS metric, this scheduler can dynamically adapt and adjust its priority according to the needs of UEs. When the QoS of UEs is high, the focus is put on reducing their energy consumption. When the network starts to experience difficulties to ensure an adequate QoS to UEs, FEC swaps of objective and the focus is put on increasing the spectral efficiency in order to consume less bandwidth so to increase the system capacity. Performance evaluation shows that FEC reaches performance close to the specialized schedulers in terms of energy efficiency and QoS. FEC takes advantages of this two worlds by providing energy efficiency when QoS of UEs is high while increasing system capacity in order to ensure QoS requirements of UEs when the bandwidth is limited.

A second contribution is also presented in this chapter. Adaptive Multi User Diversity meta Scheduler (AMUDS) is an extension of FEC principle. AMUDS is a meta scheduler which could be added to any scheduler. AMUDS adjusts the appropriate number of UEs that the scheduler can use. AMUDS computes this adequate number of UEs according to their needs and the available bandwidth. Performance evaluation shows that AMUDS allows most of schedulers to have an energy efficiency feature while keeping their intrinsic properties (in terms of fairness, spectral efficiency and delay).

6.2.3 Routing

Chapter 4 introduces the routing principle. One of the main challenge of routing is to define the best path. In our contribution, we believe that the best path to find is the path that provides the shortest delay for the considered communication. Inspired by the Little's Law, finding this path is made possible by taking into account the link and nodes states. We show that our routing solution called Opportunistic Buffer Occupancy Routing (OBOR), that takes into account the channel quality (i.e. SNR) and the buffer occupancy of nodes outperforms literature solutions while using less signaling.

6.2.4 Interference management

Chapter 5 introduces the inter-cell interference management principle. Classical solutions perform ICI management without any knowledge of neighboring cells. For instance, R1 is the well known frequency reuse solution. It does not prevent interference to occur and UEs at cell-edges may be highly interfered. On the contrary, SB splits its bandwidth by using different frequency ranges among neighboring cells. This allows to prevent interference from occurring at the expense of a high bandwidth waste. We believe that a shared knowledge of the resource allocation between clusters of cells can help to address the issue of interference management. Thanks to the emerging approach known as Cell-less, the decision making can be logically centralized allowing to have a finer management of interference. Based on this new architecture, the first contribution of this chapter known as Dynamic Cell-less Radio Access Network Meta Scheduler (DC-RAN-MS) performs either R1 or SB according to the channel state of UEs (i.e. it is an opportunistic approach). Performance evaluation emphasizes that this solution provides a better spectral efficiency, system capacity and QoS, particularly for cell-edge UEs.

The second contribution of this chapter is an extension of the first that relies on the

Joint-Transmission Coordinated MultiPoint. JT-CoMP allows gNBs to simultaneously transmit data to a same UE. It is particularly efficient to prevent interference from occurring at cell edges. However, one of the main drawbacks of the JT-CoMP is that it consumes a lot of bandwidth. Consequently, it requires to be cautiously used. Our new proposed solution called Hybrid Joint-Transmission CoMP (H-JT-CoMP), performs either R1 or JT-CoMP according to the channel state of UEs. This allows to make use of JT-CoMP only when it is required to protect cell-edge UEs. Performance evaluation shows that this solution provides a better spectral efficiency, system capacity and QoS.

6.3 Perspectives

The main axe of evolution of the work of this thesis is the Cell-free Massive MIMO context¹. Rather than having gNBs serving multiple UEs and providing wide-area coverage like in classical cellular network architecture, this innovative approach reverses this paradigm where each UE is now surrounded by multiple Access Points (APs). The Cell-free Massive MIMO takes the best aspects of three technologies: the physical layer from cellular Massive MIMO, the CoMP with joint transmission and the deployment regime of ultra dense networks. This allows to have less SNR variations, to efficiently reduce ICI while having many more APs than UEs. Figure 6.1 shows an example of the deployment of Cell-free massive MIMO with four UEs with a large number of APs.

The main challenges of such approach are to efficiently constitute the clusters of service APs for the UE considered and to move from concept to the reality by being able to provide a sufficient number of APs.

The first challenge is similar to our problematic with the constitutions of serving gNBs with the JT-CoMP (cf. Chapter 5). However efficiently constituting clusters of serving APs in the Cell-free massive MIMO may provide some significant improvements (other than those already aforementioned such as better ICI management and less SNR variations). For instance, only a part of the network could be turned on according to the number of UEs to serve, reducing the energy consumption of the network. The constitution of such cluster may also allow to provide "trusted area" for security matters.

The second challenge could be addressed by the use of radio stripes network architec-

1. Demir, Björnson, and Sanguinetti, *op. cit.*

2. Demir, Ö., Björnson, E., and Sanguinetti, L., *Foundations of User-Centric Cell-Free Massive MIMO*, vol. 14, 3-4, Now Publishers, 2021, pp. 164–472.

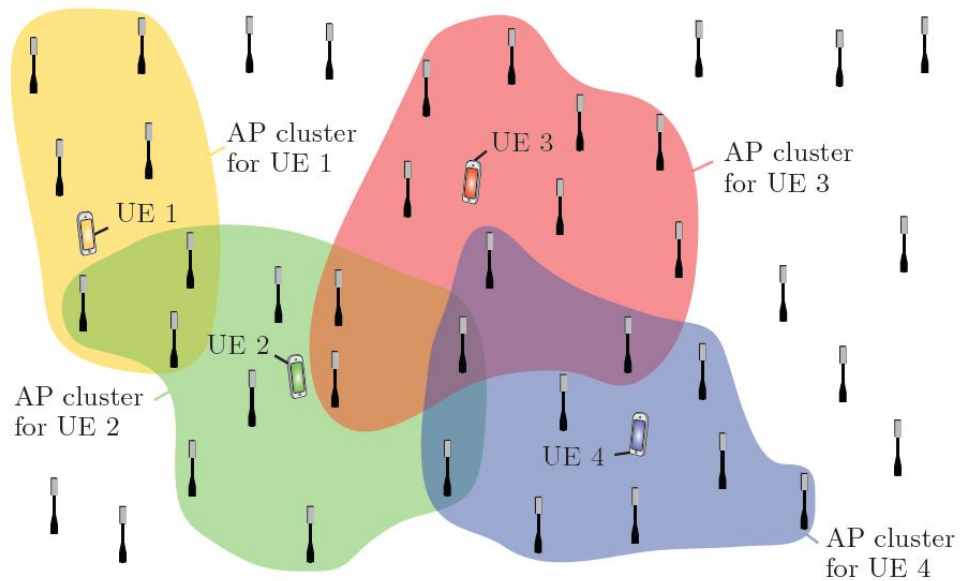


Figure 6.1 – "Example of dynamic cooperation clusters for four UEs in a Cell-free Massive MIMO network with a large number of APs" (source : Emil Björnson's book ²)
 ure ³ and Reconfigurable Intelligent Surface (RIS) ⁴.

3. Shaik, Z. H., Björnson, E., and Larsson, E. G., « Cell-free Massive MIMO with Radio Stripes and Sequential Uplink Processing », in: *IEEE International Conference and Communications Workshops*, 2020, pp. 1–6.

4. T. Van Chien, H. Q. et al., « Reconfigurable Intelligent Surface-Assisted Cell-free Massive MIMO Systems Over Spatially-Correlated Channels », in: *IEEE Transactions on Wireless Communications* (2022).

BIBLIOGRAPHY

- 3GPP, « 5G NR, Physical channels and modulation », *in: 3rd Generation Partnership Project (3GPP), TS 38.211 version 15.2.0 Release 15* (2018).
- « 5G: Study on scenarios and requirements for next generation access technologies », *in: 3rd Generation Partnership Project (3GPP), TR 38.913 version 14.3.0 Release 14* (Oct. 2017).
- « 5G: Study on scenarios and requirements for next generation access technologies », *in: 3rd Generation Partnership Project (3GPP), TR 38.913 version 14.3.0 Release 14* (2017).
- « Coordinated Multi-Point Operation for LTE Physical Layer Aspects », *in: 3rd Generation Partnership Project (3GPP), TR 36.19 R11 v11.2.0* (2013).
- « Energy efficiency of 5G », *in: 3rd Generation Partnership Project (3GPP), TS 28.310 release 17* (Mar. 2021).
- « Energy Saving Management (ESM) », *in: 3rd Generation Partnership Project (3GPP), TS 32.551 v11.0.0* (2012).
- « Evolved Universal Terrestrial Radio Access (E-UTRA); Further Enhancements for E-UTRA Physical Layer Aspect », *in: 3rd Generation Partnership Project (3GPP), Tech. Rep 68.814* (2010).
- « Max UEs/Subframe for Optimum E-UTRA DL Performance (5-20 MHz) », *in: 3rd Generation Partnership Project (3GPP), TSG-RAN WG1 R1-071354* (2007).
- « Medium Access Control (MAC) Protocol Specification », *in: 3rd Generation Partnership Project (3GPP), TS 38.321 v18.8.0* (Mar. 2022).
- « NR, Physical Layer Procedures for Data », *in: 3rd Generation Partnership Project (3GPP), TS 38.214* (2021).
- « Policy and Charging Control Architecture », *in: 3rd Generation Partnership Project (3GPP), TS 23.203 R15 v15.5.0* (2019).
- Anchun, W. et al., « Dynamic resource management in the fourth generation wireless systems », *in: Proc. IEEE Int. Conference on Communication Technology (ICCT)*, vol. 2, April 2003, pp. 1095–1098.

-
- Athanasopoulou, E. et al., « Back-Pressure-Based Packet-by-Packet Adaptive Routing in Communication Networks », *in: IEEE/ACM Transactions on Networking* 21 (Feb. 2013), pp. 244–257.
- Bassoy, S. et al., « Coordinated Multi-Point Clustering Schemes: A Survey », *in: IEEE Communications Surveys Tutorials* (2017), pp. 743–7648.
- Bassoy, S. et al., « Load Aware Self-Organising User-Centric Dynamic CoMP Clustering for 5G Networks », *in: IEEE Access* (2016), pp. 2895–2906.
- Bechir Dadi, M. and R. Belgacem Chibani, « Scheduling Performance’s Study for LTE Downlink System », *in: International Conference on Green Energy Conversion Systems (GECS)*, 2017, pp. 1–4.
- Biswas, S. and R. Morris, « ExOR: Opportunistic Multi-hop Routing for Wireless Networks », *in: SIGCOMM Comput. Commun. Rev.* 35 (2015), pp. 133–144.
- Brassard, G. and P. Bratley, *Fundamentals of Algorithmics*, Englewood Cliffs: Prentice Hall, 1996.
- Bullington, K., « Frequency Economy in Mobile Radio Bands », *in: Transactions of the IRE Professional Group on Vehicular Communications PGVC-3* (1953), pp. 4–27.
- Checko, A., *Cloud Radio Access Network Architecture. Towards 5G Mobile Networks*, 2016.
- Chen, W. et al., « Opportunistic Routing and Scheduling for Wireless Networks », *in: IEEE Transactions on Wireless Communications* (Jan. 2017).
- Clausen, T. et al., « RFC7181: The Optimized Link State Routing Protocol Version 2 », *in: IETF-Proposed Standard RFC 7681* (2014).
- Costantino, G. et al., « LoSeRO: a Locality Sensitive Routing Protocol in Opportunistic Networks », *in: Proceedings of the 31st Annual ACM Symposium on Applied Computing*, 2016, pp. 644–650.
- Darehshoorzadeh, A., L. Cerda-Alabern, and V. Pla, « Back-Pressure-Based Packet-by-Packet Adaptive Routing in Communication Networks », *in: Elsevier Computer Networks* 55 (2011), pp. 2886–2898.
- Das, S., C. Perkins, and E. Belding-Royer, *Ad hoc On-Demand Distance Vector (AODV) Routing*, RFC 3561, July 2003, DOI: 10.17487/RFC3561, URL: <https://rfc-editor.org/rfc/rfc3561.txt>.
- Del Olmo, M.R. et al., « Energy-Efficient User Scheduling Algorithm for LTE Networks », *in: 5th Joint WIC/IEEE SP Symp. Inf. Theory Signal Process. Benelux*, 2015, pp. 1–8.

-
- Demir, Ö., E. Björnson, and L. Sanguinetti, *Foundations of User-Centric Cell-Free Massive MIMO*, vol. 14, 3-4, Now Publishers, 2021, pp. 164–472.
- Elayoubi, S., O. Ben Haddada, and B. Fourestie, « Performance evaluation of frequency planning schemes in OFDMA-based networks », *in: IEEE Transactions on Wireless Communications* (2008), pp. 1623–1633.
- Elhadad, M.I., M. El-Rabaie, and M. Abd-Elnaby, « Capacity enhanced scheduler for LTE Downlink System Based on PF Algorithm », *in: Fourth International Japan-Egypt Conference on Electronics, Communications and Computers (JEC-ECC)*, 2016, pp. 1–5.
- Eom, H., « Information-Dynamics-Conscious Development of Routing Software: A Case of Routing Software that Improves Link-State Routing Based on Future Link-Delay-Information Estimation », *in: The Computer Journal* (Mar. 2008), p. 2.
- « Performance Comparison of AODV, DSDV, OLSR and DSR Routing Protocols in Mobile Ad Hoc Networks », *in: International Journal of Information Technology and Knowledge Management* (Mar. 2010), pp. 545–548.
- Ericsson, *in: Ericsson Mobility report* (2022).
- ETSI, « Assessment of mobile network energy efficiency », *in: ETSI Standard, 203 228 v1.3.1* (2020).
- Ezzaouia, M. et al., « A Dynamic Transmission Strategy Based on Network Slicing for Cloud Radio Access Networks », *in: Wireless Day*, 2018, pp. 40–45.
- Ezzaouia, M. et al., « Autonomous and Dynamic Inter-Cell Interference Coordination Techniques for Future Wireless Networks », *in: IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, Oct. 2017.
- Ferronato, J. J. and M. A. S. Trentin, « Analysis of Routing Protocols OLSR, AODV and ZRP in Real Urban Vehicular Scenario with Density Variation », *in: IEEE Latin America Transactions* 15.9 (2017), pp. 1727–1734.
- Ge, X., H. Jin, and V.C.M. Leung, « CDF-Based Scheduling Algorithm for Proportional Throughput Fairness », *in: IEEE Communications Letters* 20.5 (2016), pp. 1034–1037.
- Gokhan, M. and T. Lang, « Random Scheduling Medium access for Wireless Ad Hoc Networks », *in: Proc. IEEE Int. Conf. on MILCOM*, vol. 2, October 2002, pp. 868–872.
- Goldsmith, A., *Wireless Communications*, Cambridge University Press, 2005.
- Gonzalez, A. et al., « Deep Learning Based Speed Profiling for Mobile Users in 5G Cellular Networks », *in: IEEE Global Communications Conference (GLOBECOM)*, 2017.

-
- Gueguen, C., « Opportunistic Energy Aware Scheduler for Wireless Networks », *in: IEEE Int. 77th Vehicular Technology Conference*, 2013, pp. 1–5.
- Gueguen, C. and S. Baey, « A Fair Opportunistic Access Scheme for Multiuser OFDM Wireless Networks », *in: Journal on Wireless Communications and Networking. European Association for Signal Processing (EURASIP). Special issue: Fairness in Radio Resource Management for Wireless Network* (Feb. 2009).
- « Weighted Fair Opportunistic Scheduling for Multimedia QoS Support in Multiuser OFDM Wireless Networks », *in: IEEE Int. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2008.
- Gueguen, C., P. Fabian, and X. Lagrange, « Link State Opportunistic Routing for Multihop Wireless Networks », *in: Wireless Networks* (Feb. 2019).
- Gueguen, C. and Malo Manini, « Dynamic Tradeoff between Energy and Throughput in Wireless 5G Networks », *in: Wireless Communications and Mobile Computing Journal* (2018), pp. 1–12.
- Gueguen, C. and C. Merlhe, « Fair Energy Efficient Scheduler Providing High System Capacity for Wireless Networks », *in: SN Applied Sciences 2.12* (Dec. 2020).
- Gueguen, C., A. Rachedi, and M. Guizani, « Incentive Scheduler Algorithm for Cooperation and Coverage Extension in Wireless Networks », *in: IEEE Transactions on Vehicular Technology 62.2* (Feb. 2013), pp. 797–808, ISSN: 0018-9545.
- Hamouda H. Kabaou, M.O. and M.S. Bouhlel, « A Cross-Layer Downlink Scheduling Scheme for Balancing QoS in IEEE 802.16 Broadband Wireless Access Systems », *in: IEEE 86th Vehicular Technology Conference (VTC-Fall)*, Sept. 2017, pp. 1–5.
- Han, T. et al., « 5G Converged Cell-Less Communications in Smart Cities », *in: IEEE Communications Magazine* (2017), pp. 44–50.
- Heidarpour, A.R., M. Ardakani, and C. Tellambura, « Multiuser Diversity in Network-Coded Cooperation: Outage and Diversity Analysis », *in: IEEE Communications Letters 23.3* (2019), pp. 550–553.
- Heyman, D.P., « The GBAR Source Model for VBR Videoconferences », *in: IEEE/ACM Transactions on Networking* (1997), pp. 554–560.
- Horvath, G. and P. Fazekas, « Modelling of YouTube Traffic in High Speed Mobile Networks », *in: 21th European Wireless Conference*, Nov. 2015, pp. 1–6.
- Huhtonen, A., « Comparing AODV and OLSR routing protocols », *in: Telecommunications Software and Multimedia* (2004), pp. 1–9.

-
- Inkyu, B., K. Su Min, and S. Dan Keun, « Effects of Multiple Antennas and Imperfect Channel Knowledge on Secrecy Multiuser Diversity », *in: IEEE Communications Letters* 19.9 (2015), pp. 1564–1567.
- Jacquet, P. et al., « Optimized link state routing protocol for ad hoc networks », *in: Proceedings. IEEE International Multi Topic Conference, 2001. IEEE INMIC 2001. Technology for the 21st Century.* 2001, pp. 62–68.
- Kela, P., J. Turkka, and M. Costa, « Borderless Mobility in 5G Outdoor Ultra-Dense Networks », *in: IEEE Access* 3.3 (2015), pp. 1462–1476.
- Kolehmainen, N. et al., « Channel Quality Indication Reporting Schemes for UTRAN Long Term Evolution Downlink », *in: IEEE Vehicular Technology Conference*, Nov. 2008, pp. 2522–2526.
- Kumar, A., A. Abdelhadi, and C. Clancy, « A Delay Optimal Multiclass Packet Scheduler for General M2M Uplink », *in: IEEE Systems Journal* 13.4 (2019), pp. 3815–3826.
- Kumar, S., S. Kalyani, and K. Giridhar, « Impact of Sub-Band Correlation on SFR and Comparison of FFR and SFR », *in: IEEE Transactions on Wireless Communications* (2016), pp. 5156–5166.
- « Optimal Design Parameters for Coverage Probability in Fractional Frequency Reuse and Soft Frequency Reuse », *in: IET Communications* (2015), pp. 1324–1331.
- Kuurne, A. and A.P Miettinen, « Weighted Round Robin Scheduling Strategies in (E)GPRS Radio Interface », *in: Proc. IEEE Int. Vehicular Technology Conference (VTC)*, vol. 5, Sept. 2004, pp. 3155–3159.
- Li, Q.C. et al., « 5G Network Capacity: Key Elements and Technologies », *in: IEEE Vehicular Technology Magazine* (2014), pp. 71–78.
- Li, Q.C. et al., « Edge Cloud and Underlay Networks: Empowering 5G Cell-Less Wireless Architecture », *in: 20th European Wireless Conference*, 2014.
- Li, Y.G., N. Seshadri, and S. Ariyavisitakul, « Channel Estimation for OFDM Systems with Transmitter Diversity in Mobile Wireless Channels », *in: IEEE J. Sel. Areas Commun.* 17.3 (1999), pp. 461–471.
- Libo, J. et al., « Toward Optimal Resource Scheduling for Internet of Things Under Imperfect CSI », *in: IEEE Internet of Things Journal* 7.3 (2020), pp. 1572–1581.
- Little, J. and S Graves, *Building Intuition*, Springer, 2008.
- Liu, F., J. Riihijärvi, and M. Petrova, « Analysis of Proportional Fair Scheduling Under Bursty On-Off Traffic », *in: IEEE Communications Letters* 21 (May 2017), pp. 1175–1178.

-
- Liwei, Y., B. Bo, and C. Wei, « On Energy Efficiency Maximization in Downlink MIMO Systems Exploiting Multiuser Diversity », *in: IEEE Communications Letters* 18.12 (2014), pp. 2161–2164.
- Love, D.J. et al., « An overview of limited feedback in wireless communication systems », *in: IEEE Journal on Selected Areas in Communications* 26.8 (2008), pp. 1341–1365.
- Ma, L. and M. K. Denko, « A Routing Metric for Load-Balancing in Wireless Mesh Networks », *in: 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07)*, vol. 2, May 2007, pp. 409–414.
- Mao, X. et al., « Energy-Efficient Opportunistic Routing in Wireless Sensor Networks », *in: IEEE Transactions on Parallel and Distributed Systems* 22 (Nov. 2011), pp. 1934–1942.
- Masson, M., Z. Altman, and E. Altman, « Multi-User Collaborative Scheduling in 5G Massive MIMO Heterogeneous Networks », *in: IFIP Networking Conference*, 2020, pp. 584–588.
- Minelli, M. et al., « Scheduling Impact on the Performance of Relay-Enhanced LTE-A Networks », *in: IEEE Transactions on Vehicular Technology* 65.4 (2016), pp. 2496–2508.
- Nasralla, M.M., « A Hybrid Downlink Scheduling Approach for Multi-Traffic Classes in LTE Wireless Systems », *in: IEEE Access* 8 (2020), pp. 82173–82186.
- Nguyen, A.H. and B.D. Rao, « CDF Scheduling Methods for Finite Rate Multiuser Systems With Limited Feedback », *in: IEEE Transactions on Wireless Communications* 14.6 (2015), pp. 3086–3096.
- Ostovari, P., J. Wu, and A. Khreishah, « Cooperative Internet Access Using Helper Nodes and Opportunistic Scheduling », *in: IEEE Transactions on Vehicular Technology* 66 (July 2017), pp. 6439–6448.
- Park, S. et al., « Robust and Efficient Distributed Compression for Cloud Radio Access Networks », *in: IEEE Transactions on Vehicular Technology* (2013), pp. 692–703.
- Patriciello, N. et al., « An E2E Simulator for 5G NR Networks », *in: Elsevier Simulation Modelling Practice and Theory (SIMPAT)*, vol. 96 (2019).
- Prakash, P. and B. Chaitali, « Fair Resource Allocation to MIMO Wireless System Using Opportunistic Round Robin Scheduling Algorithm », *in: International Conference on Pervasive Computing (ICPC)*, 2015, pp. 1–3.
- Proakis, J.G. and M. Salehi, *Digital Communications*, New York: McGraw-Hill, 2007.

-
- Qamar, F. et al., « A Comprehensive Review on Coordinated Multi-Point Operation for LTE-A », *in: Computer Networks* (2017).
- Qian, M. et al., « Adaptive Soft Frequency Reuse Scheme for Wireless Cellular Networks », *in: IEEE Transactions on Vehicular Technology* (2015), pp. 118–131.
- Rong, Y. et al., « QoS Differential Scheduling in Cognitive-Radio-Based Smart Grid Networks: An Adaptive Dynamic Programming Approach », *in: IEEE Transactions on Neural Networks and Learning Systems* 27.2 (2016), pp. 435–443.
- Saffar, I. et al., « Deep Learning Based Speed Profiling for Mobile Users in 5G Cellular Networks », *in: IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–7.
- Sakr, A.H. and E. Hossain, « Location-Aware Cross-Tier Coordinated MultiPoint Transmission in Two-Tier Cellular Networks », *in: IEEE Transactions on Wireless Communications* (2014), pp. 6311–6325.
- Schulte, H. J. and W. A. Cornell, « Multi-area Mobile Telephone System », *in: Transactions of the IRE Professional Group on Vehicular Communications* VC-9.1 (1960), pp. 49–53.
- Seo, H. and B.G. Lee, « A proportional-fair Power Allocation Scheme for Fair and Efficient Multiuser OFDM Systems », *in: IEEE Global Telecommunications Conference*, Nov. 2004, pp. 3737–3741.
- Shaik, Z. H., E. Björnson, and E. G. Larsson, « Cell-free Massive MIMO with Radio Stripes and Sequential Uplink Processing », *in: IEEE International Conference and Communications Workshops*, 2020, pp. 1–6.
- Shami, T.M. et al., « User-centric JT-CoMP clustering in a 5G cell-less architecture », *in: IEEE Indoor and Mobile Radio Communications (PIMRC)*, Nov. 2018, pp. 177–191.
- Simon, C. and G. Leus, « Round-Robin Scheduling for Time-Varying Channels with Limited Feedback », *in: IEEE 10th Workshop on Signal Processing Advances in Wireless Communications*, 2009, pp. 231–234.
- Singh, D. and A. C. Ghosh, « Mobility-Aware Relay Selection in 5G D2D Communication Using Stochastic Model », *in: IEEE Transactions on Vehicular Technology* 68.3 (2019), pp. 2837–2849.
- Song, J., G. Veciana, and S. Shakkottai, « Meta-Scheduling for the Wireless Downlink through Learning with Bandit Feedback », *in: 18th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, 2020.

-
- T. Van Chien, H. Q. et al., « Reconfigurable Intelligent Surface-Assisted Cell-free Massive MIMO Systems Over Spatially-Correlated Channels », *in: IEEE Transactions on Wireless Communications* (2022).
- Tabatabaee, V. and L. Tassiulas, « Max-min Fair Self-randomized Scheduler for Input-buffered Switches », *in: Workshop on High Performance Switching and Routing (HPSR)*, 2004, pp. 299–303.
- Taejoon, K. and L. Jong-Tae, « Capacity Analysis and Feedback Threshold Optimization in Fair Multiuser Diversity System », *in: IEEE Transactions on Vehicular Technology* 61.9 (2012), pp. 4189–4194.
- « Queuing Analysis in a Multiuser Diversity System With Adaptive Modulation and Coding Scheme », *in: IEEE Transactions on Vehicular Technology* (2011), pp. 338–342.
- Tamimi, A.K.A., R. Jain, and C. So-In, « Modeling and Prediction of High Definition Video Traffic: A Real-World Case Study », *in: Second International Conferences on Advances in Multimedia*, 2010.
- Tanwir, S. and H. Perros, « A Survey of VBR Video Traffic Models », *in: IEEE Communications Surveys Tutorials*, 2013.
- Truman, T.E. and R.W. Brodersen, « A Measurement-Based Characterization of the Time Variation of an Indoor Wireless Channel », *in: Proc. IEEE Int. Universal Personal Communications Record (ICUPC)*, 1997.
- Van Dam, T. and K. Langendoen, « An Adaptive Energy-Efficient MAC Protocol for Wireless Sensor Networks », *in: Proceedings of the 1st international conference on Embedded networked sensor systems*, 2003, pp. 171–180.
- Viswanath, P., D. N. C. Tse, and R. Laroia, « Opportunistic Beamforming Using Dumb Antennas », *in: IEEE Transactions on Information Theory* 48 (June 2002), pp. 1277–1294.
- Wang, K., K. Yang, and C.S Magurawalage, « Joint Energy Minimization and Resource Allocation in C-RAN with Mobile Cloud », *in: IEEE Transactions on Cloud Computing* (2018), pp. 760–770.
- Wong, C. Y. and R. S. Cheng, « Multiuser OFDM with Adaptive Subcarrier, Bit, and Power Allocation », *in: IEEE J. Sel. Areas Commun.* (1999).
- Yao, M., « Proportional Fair Scheduling for Downlink OFDMA », *in: Proc. IEEE Int. Conf. on Communications (ICC)*, June 2007, pp. 4843–4848.

Yujia, H. and T. Xiaofeng, « Secrecy Outage Analysis of Multiuser Diversity With Unequal Average SNR in Transmit Antenna Selection Systems », *in: IEEE Communications Letters* 19.3 (2015), pp. 411–414.

Titre : Allocation des ressources dans un contexte "Cell-less" pour les réseaux sans fil 5G

Mot clés : Allocation des ressources, Cell-less, Réseaux sans fil, 5G, Qualité de service, Efficacité énergétique

Résumé : L'augmentation significative du nombre d'utilisateurs ayant des besoins toujours croissants, l'émergence de nouveaux services et de nouvelles applications ont amené les réseaux mobiles à évoluer. Les travaux de cette thèse ont pour objectif de répondre à ces enjeux. La première partie de cette thèse s'intéresse particulièrement à l'allocation des ressources classique et fournit quatre contributions : une analyse de la diversité multi-utilisateur, deux nouvelles solutions d'ordonnancement et une nouvelle solution de routage. Les résultats de ces travaux montrent que ces contributions apportent des solutions pour répondre aux enjeux des réseaux mobiles 5G de demain en augmentant par exemple l'efficacité spectrale et l'efficacité énergétique, la qualité de service tout en ré-

duisant le délai global des utilisateurs. Fort de ces résultats et des analyses qui en découlent, la deuxième partie de cette thèse se concentre sur l'allocation des ressources dans un contexte "Cell-less". Cette approche innovante permet notamment d'avoir une prise de décision de manière logiquement centralisée. Cela s'applique tout particulièrement à la gestion des interférences inter-cellulaires, où deux nouvelles solutions sont présentées dans cette thèse. Les résultats montrent une augmentation de l'efficacité spectrale et une réduction du délai des utilisateurs, particulièrement pour ceux en bordure de cellule. De plus, les résultats obtenus via l'approche "Cell-less" sont accrus grâce à l'utilisation du Joint-Transmission Coordinated MultiPoint.

Title: Resource Allocation in a Cell-less context for 5G Wireless Networks

Keywords: Resource allocation, Cell-less, Wireless network, 5G, QoS, Energy efficiency

Abstract: The significant increase of the number of users with ever-growing needs, the emergence of new services and new applications have led mobile networks to evolve. The main objective of this thesis is to respond to these challenges. The first part of this thesis focuses on classical resource allocation and provides four contributions: a multi-user diversity analysis, two new scheduling solutions and a new routing solution. The results of this work show that these contributions provide solutions to meet the challenges of tomorrow's 5G mobile network, by, for instance, increasing spectral and energy efficiencies, increasing the QoS while reducing the user de-

lay. Based on these results and the ensuing analysis, the second part of this thesis focuses on resource allocation in a "Cell-less" context. This innovative approach enables logically centralized decision making. This is particularly efficient for inter-cell interference management, where two new solutions are presented in this thesis. The results show an increase in spectral efficiency and a reduction in user delay, particularly for those located at cell-edges. In addition, the results obtained with the "Cell-less" approach are enhanced by the use of Joint-Transmission Coordinated MultiPoint.