

Dissémination de contenus populaires et tolérants au délai dans les réseaux cellulaires

Salah-Eddine Belouanas

▶ **To cite this version:**

Salah-Eddine Belouanas. Dissémination de contenus populaires et tolérants au délai dans les réseaux cellulaires. Réseaux et télécommunications [cs.NI]. Université Pierre et Marie Curie - Paris VI, 2017. Français. NNT : 2017PA066230 . tel-01767009

HAL Id: tel-01767009
https://theses.hal.science/tel-01767009

Submitted on 15 Apr 2018

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.

UPMC Sorbonne Universités
13 October 2017

Dissémination de contenus populaires et tolérants au délai dans les réseaux cellulaires

Thèse présentée
par
Salah-Eddine BELOUANAS

pour obtenir
le grade de docteur
spécialité
Informatique, Télécommunications et Électronique

Jury :

Véronique Vèque	Rapporteur	Professeur, Université Paris-Sud
Yacine Ghamri-Doudane	Rapporteur	Professeur, Université de la Rochelle
Aline Carneiro Viana	Examinateur	Chargée de recherche, Inria Saclay-Île-de-France
Vania Conan	Examinateur	Directeur des Études Amont, Thales C&S
Kim-Loan Thai	Encadrante	Maître de conférence, UPMC Sorbonne Universités
Prométhée Spathis	Encadrant	Maître de conférence, UPMC Sorbonne Universités
Marcelo Dias de Amorim	Directeur	Directeur de Recherche CNRS

À mes parents

Ma femme et toute ma famille

Remerciement

J'affirme honnêtement que le chemin parcouru pour obtenir ma thèse n'aurait pas été possible sans le soutien de nombreuses personnes.

Tout d'abord, je veux remercier Marcelo, mon directeur de thèse et mon guru, pour ses excellentes orientations, son soutien et son dévouement pendant la durée de ma thèse. Ça était un grand bonheur d'effectuer mes travaux de recherche sous sa direction. Je lui suis reconnaissant pour tout, en particulier, pour m'accorder la liberté de développer mes idées ainsi que sa façon de simplifier les choses. Ses suggestions, ses recommandations et sa capacité à gérer plusieurs choses diverses étaient très remarquables pour moi. J'en profite de cette occasion pour lui exprimer ma plus grande gratitude d'avoir été un excellent conseiller.

Je remercie tout particulièrement mes deux honorables encadrants *Kim* et *Prométhée*. Leurs motivations et idées ont été un véritable moteur pour achever ce travail. J'ai appris beaucoup de choses en bénéficiant de leurs expériences dans le domaine de la recherche scientifique ainsi que leurs excellentes compétences.

Je remercie l'Agence National de Recherche (ANR) de m'avoir accordé la chance de développer cette thèse dans le cadre du projet DataTweet (ANR-13-INFR-0008). J'aimerai exprimer ma plus grande appréciation pour leur soutien financier qui m'a permit de présenter mes travaux dans des évènements international.

Je tiens à exprimer mes très chaleureux remerciements aux Professeurs Véronique Vèque et Yacine Ghamri-Doudane pour leur temps précieux pour évaluer et rapporter cette thèse. Je suis très conscient que les suggestions qu'ils m'ont adressées ont une grande contribution à l'amélioration de la qualité de mon travail. Je tiens aussi à remercier Aline Carneiro Viana et Vania Conan pour avoir accepté de faire partie de mon jury et évaluer les travaux que j'ai réalisés pendant ces trois dernières années.

J'ai passé de très merveilleux moments au LIP6. Je dois remercier tous mes collègues de l'équipe NPA, et en particulier : Mustapha, Amr, Walid, Wafa, Fadwa, Nesrine, Bu, Minh, Mehdi,

Clément, Benjamin, Alexandre, Narcisse, Quentin, Giulio, Davide, Antonella, Florian, Filippo, Thiago, Alexendru, Tiph et Matteo. Sans oublier nos présents stagiaires et futurs doctorants : Bahiya, Mustapha et Wahib. Je m'excuse à l'avance si j'oublie quelqu'un.

Cette thèse n'aurait jamais été réalisable sans l'aide de ma famille. J'exprime mon extrême reconnaissance à : ma mère (Mebarka), mon père (Lakhdar), ma soeur (Djamila), mes deux frères (Nourdine et Abdelwahab), ma cousine (Amel) et mes neveux (Adel, Mouhamed, Abdallah, Ayoub et Youcef) qui m'ont toujours supporté et soutenu. Votre amour et encouragement restent toujours ma source d'inspiration dans la vie.

Asma, ma très chère femme, tu es toujours ma force. Un grand merci pour ta patience, ton amour et ton soutien qui m'ont accompagné pendant toutes ces années. Je sais que ce n'était pas simple pour toi. "*A big thanks to you my dear love.*"

Je tiens aussi à remercier mes chers amis depuis longtemps et très particulièrement : Fayçal (mon ami d'enfance), Ramzi (le médecin qui connait tout), Chouiab (monsieur l'escroc), Bassem (chubby boy), Lamine & Raouf (les jumeaux), Ahmed, Younes, Hakim (the tall guy), Lokman (loka), Ismail (sma), Fayçal (jari), Djaafar, Rachid, Hatem (el3ifa), Abdallah, Lotfi (chrix), Assem (VW expert), Reda et Salim.

Résumé

L'explosion du marché des dispositifs mobiles intelligents (*smartphones*, tablettes et ordinateurs portables) a généré un déferlement de données utilisateurs sur les réseaux depuis quelques années. Cela explique l'augmentation du trafic mobile mondial de 7 exaoctets par mois en 2016 à 49 exaoctets par mois d'ici 2021. En conséquence, les réseaux cellulaires souffrent d'une forte pression afin de faire face à cette surcharge de trafic. Hélas, les approches existantes ne pourront pas suivre indéfiniment cette avalanche des données sans le déploiement d'une nouvelle infrastructure coûteuse. Certains opérateurs adoptent des stratégies pour limiter l'utilisation des ressources cellulaires, telles que la mise en œuvre d'une stratégie d'affectation de ressources en deux phases, où les utilisateurs voient leur bande passante diminuée après un certain seuil de consommation de données.

Toutefois, des stratégies plus récentes, regroupées sous le nom de *délestage de données*, s'articulent autour de l'exploitation de technologies alternatives. En particulier, les communications du type dispositif-à-dispositif profitent de la coopération des utilisateurs mobiles dans les phases de transmission pour transférer une partie des contenus à la place de l'infrastructure. Ces dernières stratégies profitent de la tolérance au délai des contenus ainsi que de la mobilité des utilisateurs pour apporter un grand avantage à l'opérateur. En adoptant cette nouvelle tendance technologique, nous sommes fortement convaincus que toute économie de ressources cellulaires est la bienvenue.

Dans cette thèse, nous étudions le problème de dissémination de contenus dans un réseau cellulaire surchargé. Cela signifie que l'infrastructure est saturée par le trafic qui passe, mais qu'elle doit répondre aux demandes des utilisateurs mobiles. Dans un premier temps, nous étudions les différentes stratégies proposées dans la littérature selon deux critères de classification. Le premier critère concerne le support de transmission (cellulaire ou D2D), alors que le deuxième critère est temporel, sachant que nous traitons des contenus tolérants aux délais. Notre première contribution est SCoD (*Scheduled Content Delivery*), une stratégie de dissémi-

nation purement cellulaire qui exploite la mobilité des utilisateurs ainsi que leur tolérance aux délais afin d'effectuer des transmissions différées. L'idée de base de SCoD est d'attendre que les utilisateurs se rassemblent autour d'un minimum de points d'accès afin que le nombre total de transmissions soit réduit. Afin de déterminer les meilleurs instants pour déclencher des transmissions, SCoD repose sur différentes fonctions de décision qui lancent, le cas échéant, une transmission en multicast.

Les communications opportunistes D2D font référence aux transmissions directes entre dispositifs mobiles proches, sans faire appel à l'infrastructure cellulaire. Ces transmissions directes offrent un large éventail de bénéfices, incluant l'efficacité du spectre cellulaire, l'amélioration de la capacité du système et la réduction de latence. La deuxième partie de cette thèse traite encore le problème d'économie de ressources cellulaires, mais sous un autre angle. Nous proposons d'associer les communications D2D au multicast cellulaire afin de réduire la charge sur l'infrastructure. En d'autres termes, nous exploitons les communications D2D pour soulager le goulot d'étranglement que constitue le canal d'accès cellulaire. Nous étudions les avantages des communications opportunistes dans le contexte d'un service de diffusion de contenus basé sur les abonnements des utilisateurs. Au début, le contenu est livré selon un processus en deux phases. Dans une première phase, le contenu est transmis à un nombre restreint d'utilisateurs mobiles chargés de transmettre le contenu aux abonnés qu'ils rencontreront (ce sont des *mules de données*). La seconde phase a lieu pendant le temps restant avant la fin de la période de validité du contenu, et au terme de laquelle l'infrastructure devra intervenir si nécessaire. Au cours de cette phase, le contenu est transmis par les mules de données, dans les cellules qui hébergent des utilisateurs n'ayant pas encore reçu le contenu. Nous présentons nos résultats selon deux cas de figure. Tout d'abord, nous illustrons le bénéfice des communications D2D dans le cadre d'un processus de diffusion, où le coût d'une transmission opportuniste est négligeable par rapport à celui d'une transmission cellulaire. Nous considérons ensuite le cas où les mules de données doivent être indemnisées pour leur participation au processus de diffusion. Cela rend forcément la transmission D2D plus coûteuse à un moment donné que la simple transmission cellulaire. Il y a donc un compromis à trouver : nous fournissons ainsi des lignes directrices en ce sens et nous montrons la variation d'un tel compromis en fonction de plusieurs paramètres de réseau, tels que la portée de diffusion en D2D et le nombre de mules de données sélectionnées.

Table des matières

1 Introduction **2**

1.1 Contexte et motivations 2

1.2 Défis et positionnement 3

1.3 Contributions de la thèse 5

1.4 Les applications ciblées 6

2 Économiser les ressources cellulaires : Un état de l'art **8**

2.1 Classification des solutions 8

2.2 Solutions avec infrastructure cellulaire 9

2.3 Solutions sans infrastructure cellulaire 22

2.4 Positionnement de la thèse vis-à-vis de l'état de l'art 33

3 SCoD : Livraison de contenu planifiée en environnement purement cellulaire **35**

3.1 Introduction 35

3.2 Modèle de système, hypothèses et illustration du problème 38

3.3 Jusqu'où la tolérance au délai peut-elle aider à économiser les ressources radio ? 40

3.4 SCoD : Livraison de contenu planifiée 43

3.5 Trace de mobilité et les paramètres d'évaluation 47

3.6 Résultats d'évaluation 49

3.7 Conclusion 55

4 Résister au déferlement de données cellulaires au moyen de communications opportunistes **57**

4.1 Introduction 57

4.2 Modèle de système et hypothèses 59
4.3 Les stratégies de diffusion 62
4.4 La compagne d'évaluation 64
4.5 Les résultats numériques 68
4.6 Conclusion 72

5 Conclusion et Perspectives **74**
5.1 Conclusion 74
5.2 Perspectives 76

References **88**

Table des figures

1.1 Illustration du principes de la dissémination différée d'un contenu. 4
1.2 Illustration d'un scénario cellulaire avec les communications D2D. 6

2.1 Classification des stratégies d'économie des ressources cellulaires. 9
2.2 Architecture à 3 niveaux d'un réseau d'infostations. 12
2.3 Scénarios de délestage de données (*mobile data offloading*). 14
2.4 Les Femtocellules déployées dans des environnements résidentiels et des entreprises. 17
2.5 Évolution vers un système LTE-U (intra-réseau). 20
2.6 Communications D2D : *underlay* vs *overlay*. 23
2.7 Communications D2D : utilisation d'un spectre sans licence. 27

3.1 Principe de fonctionnement d'une dissémination de contenu en environnement purement cellulaire. 36
3.2 Les stratégies Hot Potato et Oracle. 41
3.3 Hot Potato et Oracle : Le nombre de transmissions réalisées en fonction de la tolérance au délai et de la population (a) $|\mathbf{N}^c| = 500$ (b) $|\mathbf{N}^c| = 1000$. . 42
3.4 Fonctionnement de SCoD. 45
3.5 Variation du taux de regroupement pour des intervalles consécutifs de durée τ = 300 s. 46
3.6 Exemple d'un diagramme de Voronoi avec 100 cellules déployées dans la ville de Bologne. 48
3.7 Le gain moyen de transmission pour un deadline $\tau = 180\ s$ à Luxembourg : SCoD vs les stratégies de base. 49

3.8 Le gain moyen de transmission pour un deadline $\tau = 300\ s$ à Luxembourg : SCoD vs les stratégies de base. 50
3.9 Nombre moyen de transmissions des variantes real run et ideal run pour Luxembourg. 51
3.10 Les graphiques montrent les trois fonctions de retardement $\beta(t)$ utilisées par l'ensemble de données real run et ideal run pour la trace de Luxembourg. . . 51
3.11 Le gain moyen de transmission pour un deadline $\tau = 180\ s$ à Bologne : SCoD vs. les stratégies de base. 53
3.12 Le gain moyen de transmission pour un deadline $\tau = 300\ s$ à Bologne : SCoD vs les stratégies de base. 53
3.13 Nombre moyen de transmissions des variantes real run et ideal run pour Bologne. 54
3.14 Comparaison des fonctions de retardement $\beta(t)$ pour la trace de Bologne. . 55

4.1 (a) La première transmission satisfait quatre utilisateurs, trois d'entre eux sont choisis comme mules de données (4, 5 et 6). (b) La mobilité des utilisateurs, y compris les mules de données, permet la satisfaction d'un certain nombre d'utilisateurs intéressés de manière opportuniste. (c) Une transmission cellulaire supplémentaire est nécessaire lorsque le deadline du contenu est atteinte afin de satisfaire les utilisateurs n'ayant pas encore reçu le contenu. 61
4.2 Gain moyen de transmission pour les stratégies étudiées : Wait-and-See, Hot Potato et Oracle. Le deadline considéré est τ = 180 s. 66
4.3 Gain moyen de transmission pour les stratégies étudiées : Wait-and-See, Hot Potato et Oracle. Le deadline considéré est τ = 300 s. 67
4.4 Le pourcentage de couverture des utilisateurs intéressés avec les mules de données et les points d'accès cellulaires. Plusieurs paramètres sont illustrés. . 69
4.5 La relation entre le coût cellulaire et le coût des transmissions opportunistes. 70

Si j'avais une heure pour résoudre un problème, je passerais 55 minutes à réfléchir au problème et 5 minutes à penser à des solutions.

Albert Einstein

1 Introduction

1.1 Contexte et motivations

De nos jours, les objets intelligents deviennent omniprésents. En connectant le monde physique au monde numérique, ils tirent parti de nouveaux scénarios vers des villes intelligentes, des maisons intelligentes, des infrastructures intelligentes et des systèmes de transport intelligents. Ces objets sont des sources de mesure et de contenu ainsi qu'un outil d'amélioration de la vie quotidienne. De plus, la montée de l'industrie des objets intelligents a conduit à une demande explosive en termes d'applications et de services mobiles répondant à des besoins d'utilisateurs comme par exemple l'organisation, les voyages, la communication et les informations.

Chaque année, plusieurs nouvelles gammes de dispositifs communicants se différenciant par leurs formes, dimensions et capacités ainsi que par leur intelligence, sont introduites sur le marché. En 2016, le nombre de dispositifs qui ont été connectés sur le réseau mondial est passé à 8 milliards, contre 7,6 milliards en 2015. Ce chiffre va atteindre les 11,6 milliards d'ici 2021 [29, 40]. Par conséquent, le trafic mobile en 2016 a atteint le volume considérable de 7 exaoctets par mois, et il pourrait atteindre les 49 exaoctets par mois d'ici 2021 (7 fois plus). L'in-

frastructure actuelle ne sera tout simplement pas capable de s'adapter à cette demande massive. Une conséquence directe est que les opérateurs cellulaires mènent une lutte permanente pour supporter la croissance fulgurante du trafic mobile.

Dans un proche avenir, les opérateurs n'auront d'autre choix que de combiner un large éventail de techniques pour économiser les ressources cellulaires. De multiples solutions ont été introduites ces dernières années afin de parer à cette croissance de trafic. Jusqu'à présent, les principaux efforts ont porté sur l'évolution de la technologie cellulaire en passant de la 3G vers la 4G et bientôt la 5G et l'accroissement du plan de déploiement à l'aide de nouvelles macro-stations de base ou de micro-stations de base telles que Femtocellules [23]. D'autres opérateurs se focalisent sur la mise en œuvre d'une stratégie d'affectation de ressources en deux phases, où les utilisateurs voient leur bande passante significativement limitée après le dépassement d'un seuil de consommation de données (généralement de l'ordre de quelques Gigabits) [66, 92]. Hélas, les tentatives précédentes restent coûteuses ou ne répondent pas aux attentes des utilisateurs.

À ce stade, une alternative a été introduite, appelée le *délestage de données* (*data offloading*) [80] qui consiste à rediriger une partie du trafic vers le spectre sans licence (c.-à-d. des points d'accès Wi-Fi ou vers des communications du type dispositif-à-dispositif). Cette dernière solution a attiré l'attention de la communauté au cours des dernières années car elle exploite des technologies moins coûteuses que celles utilisées par les transmissions cellulaires. Bien que le mécanisme de délestage de données marche très bien, il y a des limites qui peuvent jouer sur la satisfaction des utilisateurs ou même de l'opérateur. Un tel réseau nécessite la gestion d'un système supplémentaire pour assurer la commutation sûre entre les différentes technologies (cellulaire et Wi-Fi par exemple). Cela peut réduire considérablement la qualité de service (QoS) fournie aux utilisateurs mobiles. La compréhension de l'attente des utilisateurs ainsi que du type d'applications dont ils ont besoin devient fondamentale pour concevoir un mécanisme d'échange de contenus efficace.

1.2 Défis et positionnement

Dans cette thèse, nous abordons le problème de la diffusion de contenus en présence de lourdes contraintes cellulaires en termes de disponibilité de ressources. Cela signifie que le réseau cellulaire est surchargé par le trafic et qu'il doit répondre de la façon la plus efficace possible aux demandes des utilisateurs. Tout au long de ce travail, nous nous attacherons à garantir une satisfaction totale des utilisateurs (taux de livraison de 100%). Nous évaluons plusieurs stra-

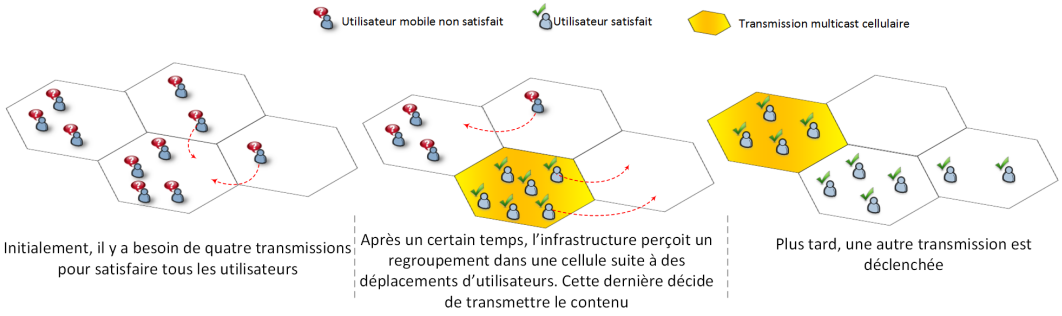


Figure 1.1 – Illustration du principes de la dissémination différée d'un contenu.

tégies de dissémination sous des angles différents. Les résultats de ce travail de thèse ont été développés dans le cadre du projet ANR DataTweet [10]. Les stratégies de transmissions qui ont été développées au cours de cette thèse sont évaluées via une implantation en utilisant de multiples outils. Cette thèse s'attaque aux questions suivantes :

- Comment transmettre un contenu satisfaisant les demandes des utilisateurs en réduisant le gaspillage de ressources cellulaires ?
- La mobilité des utilisateurs et la tolérance au délai de certains types de contenus peuvent-elles contribuer à améliorer le processus de dissémination ?
- Les communications opportunistes sont-elles vraiment bénéfiques pour réduire le trafic cellulaire ?

Tout au début de cette thèse, nous présentons un chapitre de l'état de l'art avec une taxonomie des solutions et techniques de dissémination de contenus. Ces dernières ont pour objectif commun de faire diminuer le trafic cellulaire sur l'opérateur et donc d'économiser des ressources qui pourront être exploitées pour d'autres usages. Nous fournissons une classification des techniques selon deux dimensions. La première concerne le support de transmission proprement dit : est-ce l'infrastructure cellulaire qui déclenche les transmissions ou les terminaux mobiles ? Quant à la seconde dimension, elle se rapporte à un critère temporel, puisque nous nous inscrivons dans un contexte de contenus tolérants aux délais. Ce chapitre nous permet également de nous positionner par rapport aux différentes techniques proposées dans la littérature.

1.3 Contributions de la thèse

1.3.1 Contribution 1 : Dissémination efficace en environnement purement cellulaire

Dans le chapitre 2, nous proposons d'exploiter, pour la première partie de ce travail, une architecture de réseau mobile purement cellulaire. Le chapitre 3 présente dans la foulée SCoD (*Scheduled Content Delivery*), notre stratégie de dissémination permettant une économie de ressources. SCoD s'appuie sur la mobilité des utilisateurs pour réduire le nombre de transmissions multicast cellulaires (Figure 1.1). L'idée est de reporter ultérieurement une transmission afin de maximiser les chances d'avoir un regroupement important d'utilisateurs dans la même cellule. Nous exploitons la tolérance au délai de certains types de contenus, combinée avec la tendance naturelle des utilisateurs à se réunir dans les mêmes endroits. En nous appuyant sur les communications multicast (p. ex. le multicast LTE) [34], nous pouvons réduire le nombre de transmissions nécessaires pour satisfaire tous les utilisateurs intéressés par un même contenu. SCoD repose sur les observations précédentes pour déterminer, de manière adaptative, le nombre d'utilisateurs qui devraient déclencher une transmission multicast dans une cellule.

1.3.2 Contribution 2 : Dissémination hybride cellulaire/opportuniste

Le multicast représente un moyen très simple et efficace pour la distribution des contenus à de multiple utilisateurs. Dans la deuxième partie de cette thèse, nous étudions le gain obtenu en couplant des communications opportunistes (*device-to-device*) avec les transmissions multicast cellulaires [37]. Dans cette optique, nous proposons une stratégie de diffusion d'un contenu qui exploite les communications D2D pour rendre plus efficiente la transmission de l'information en diminuant la quantité de trafic passée par le canal cellulaire. L'idée est de tirer parti de la tolérance aux délais des contenus ainsi que de la mobilité des utilisateurs afin de rediriger une partie du trafic (initialement prévu sur des liens cellulaires) vers des liens opportunistes. Cela passe par la sélection d'un ensemble de nœuds, appelés *mules de données*, qui sont les seuls à transmettre en D2D (figure 1.2). L'infrastructure cellulaire retarde ses transmissions et suit les utilisateurs mobiles pendant leurs mouvements afin de détecter de possibles regroupements. Juste avant la fin de validité du contenu, l'infrastructure cellulaire intervient pour satisfaire le reste des utilisateurs qui n'ont pas encore reçu leur contenu.

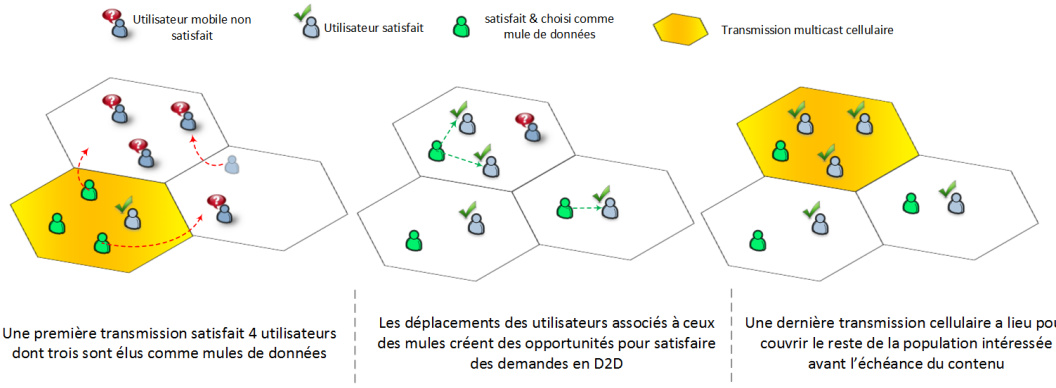


Figure 1.2 – Illustration d'un scénario cellulaire avec les communications D2D.

Notre première étude, dans cette partie, s'attache au cas où les mules de données sont altruistes. Cela veut dire qu'elles n'exigent pas de contrepartie pour l'utilisation de leur appareil comme moyen d'échanger des données avec d'autres utilisateurs. En revanche, pour intégrer le D2D dans un système cellulaire, il faut soigneusement étudier cette coopération selon différentes dimensions, y compris le coût.

À ce stade, nous introduisons un second cas, dans lequel l'opérateur va se montrer plus généreux et prendre en compte le coût des échanges D2D. Dans ce cas, les mules de données ne sont pas traitées comme les autres nœuds puisqu'elles doivent être indemnisées pour permettre l'utilisation des capacités de leur appareil dans le processus de diffusion. L'objectif dans notre analyse est de trouver le point à partir duquel le D2D n'est plus utile ni bénéfique pour l'opérateur.

1.4 Les applications ciblées

Nous ciblons dans cette thèse les applications de diffusion de données dans les réseaux cellulaires, et en particulier, les applications qui tolèrent un certain délai avant l'expiration des contenus. À titre d'exemple, nous pouvons citer : les éditions numériques des journaux, les prévisions météorologiques, des événements liés à une ville, les contenus populaires sur les réseaux sociaux, des informations sur l'état des routes, la localisation de foules, ou même les informations sportives. Ces applications répondent à une réelle tendance à la production et la consommation de contenus générés par les utilisateurs.

Actuellement, les services mobiles livrent leurs contenus en mode point-à-point. En conséquence, l'infrastructure qui fournit du contenu à plusieurs utilisateurs à la fois doit établir et

maintenir une connexion séparée avec chacun d'eux. Cette approche fonctionne bien pour un nombre faible ou modéré d'abonnés, mais tend à être très inefficace au fur et à mesure que le nombre d'abonnés augmente. Ceci explique un regain d'intérêt pour les communications suivant le modèle point-à-multipoint. De ce fait, l'organisme de standardisation 3GPP [1] développe des activités dans le cadre des *Multimedia Broadcast and Multicast Services* (MBMS)[4]. MBMS prend en charge deux modes de transmission de base pour la livraison de paquets IP : diffusion et multidiffusion. Il se concentre donc sur les aspects transport des services de diffusion en se basant sur le protocole IP multicast [70]. Les premiers essais ont été faits par l'opérateur téléphonique Verizon [3] aux États-Unis en 2014 lors un événement sportif. L'expérience a été concluante. Elle a été élargie à d'autres services tels que des informations d'intérêt général, des notifications, etc.

Étudie le Passé si tu veux connaître l'Avenir.

Confucius

2

Économiser les ressources cellulaires : Un état de l'art

Ce travail de thèse se situe dans le domaine de la transmission de contenus sur réseaux cellulaires, et porte tout particulièrement sur le cas des réseaux assujettis à un trafic mobile important. Dans ce chapitre, nous examinons les travaux récents dont le but est de diminuer la charge sur l'infrastructure cellulaire et à augmenter l'efficacité du spectre radio. Pour chacun de ces travaux, nous présentons les caractéristiques des stratégies proposées. Cela nous permettra par la suite de positionner nos propres travaux par rapport à l'existant ; nous verrons en particulier que notre proposition peut venir en complément de la plupart de ces solutions, sans les concurrencer.

2.1 Classification des solutions

La diffusion de contenus sur réseaux mobiles à forte charge a été largement étudiée dans la littérature. De nombreuses stratégies pour limiter la surcharge en débit ont été ainsi proposées, et plusieurs classifications sont possibles en fonction du critère retenu. Nous avons choisi de

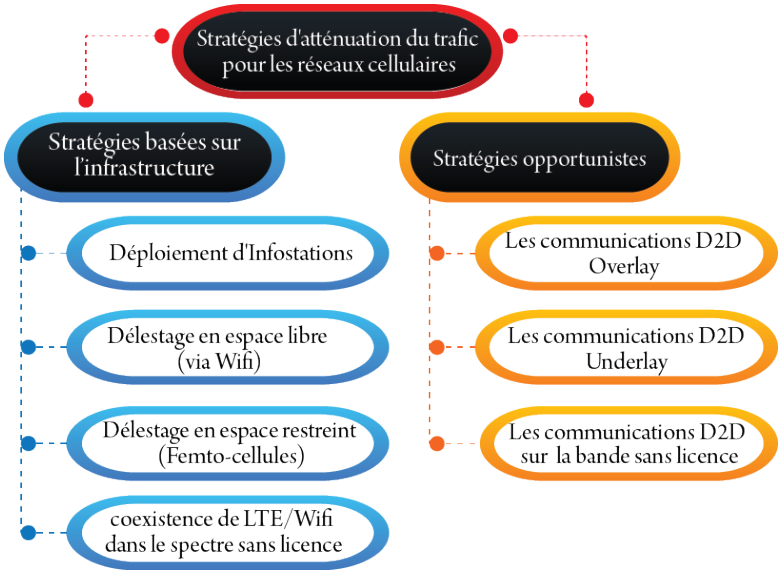


Figure 2.1 – Classification des stratégies d'économie des ressources cellulaires.

présenter l'état de l'art avec comme premier critère de classification les entités impliquées dans la transmission de contenus lors du processus de dissémination : *l'infrastructure cellulaire* ou les *terminaux mobiles* eux-mêmes (voir figure 2.1). Par ailleurs, nous avons considéré, dans le contexte de cette thèse, des contenus tolérants au délai, c'est-à-dire pouvant accepter un délai supplémentaire avant d'être délivrés à leurs destinataires finaux. Ainsi, pour une raison d'homogénéité avec les chapitres suivants, nous avons retenu comme second critère une dimension temporelle, sachant que les applications peuvent avoir des contraintes temporelles de façon à garantir la satisfaction de leurs utilisateurs. Nous mettrons ainsi en exergue, lorsque c'est le cas, la notion de tolérance au délai pour chacun des travaux présentés.

2.2 Solutions avec infrastructure cellulaire

En parcourant la littérature, nous avons pu constater que les systèmes de communication actuels ont été conçus pour servir les utilisateurs "n'importe quand et n'importe où". Cependant, les exigences de plus en plus fortes sur un tel service en augmentent nécessairement le coût, alors que le réseau cellulaire se doit d'assurer un service de qualité à ses utilisateurs mobiles. À ce niveau, plusieurs méthodes ont été envisagées pour accroître la connectivité du réseau, mais qui induisent un coût supplémentaire et une consommation d'énergie très élevée. Ces propositions ne peuvent donc pas constituer des solutions à long terme, tant d'un point de vue économique que d'un point de vue opérationnel. De ce fait et toujours dans l'optique

de permettre aux opérateurs d'améliorer l'expérience des utilisateurs, des alternatives ont été proposées pour une distribution des contenus efficace, en d'autres termes, tenant compte de la rareté du spectre cellulaire. Nous nous focalisons dans ce paragraphe sur ce type de solutions et qui exploitent essentiellement l'infrastructure de communication. Toutes ces solutions ont pour point commun de répondre à la question suivante : *Comment un opérateur cellulaire peut-il satisfaire efficacement les demandes de diffusion de contenus de ses utilisateurs, alors que son infrastructure est le seul moyen de desservir ces derniers ?*

2.2.1 Déploiement d'infostations

Au tout début des années 2000, des chercheurs du laboratoire WINLAB (Rutgers University) ont proposé un nouveau type de réseau sans fil [44, 45] comme solution architecturale. Cette solution prend en compte les problèmes de coût et de limitation énergétique des terminaux, ces derniers étant considérés comme des facteurs de ralentissement à la mise en place de nouveaux services sans fil. Son originalité réside dans le fait que la zone de couverture n'est pas globale ; a contrario, leur architecture se compose de zones géographiques discontinues, appelées *infostations*. À l'intérieur des infostations, des bornes sans fil agissent comme des passerelles entre l'infrastructure et les utilisateurs finaux. Cibler des contenus tolérant des délais supplémentaires avant leur remise aux utilisateurs permet alors d'offrir un service de communication en environnement mobile à haut débit et à faible coût. Par ailleurs, pour servir au mieux les utilisateurs, les infostations peuvent être placées dans des endroits stratégiques, tels que les aéroports, les entrées de bâtiments d'entreprises, les routes très empruntées, etc. Ce concept d'infostations a été ensuite largement repris et étudié sous de multiples facettes.

Le problème du placement optimal des infostations dans une zone géographique donnée a été ainsi traité dans [93]. Xiong et al. y étudient la diffusion d'un contenu avec contrainte temporelle en cherchant à minimiser le coût d'urbanisation en infostations. Pour cela, les auteurs construisent un graphe de mobilité modélisant les déplacements des différents utilisateurs, et calculent la couverture minimum par sommets [36]. Les résultats obtenus par heuristique et par simulation avec des traces de mobilité réelles montrent que la méthode proposée atteint son objectif avec seulement un nombre réduit d'infostations.

Dans le même contexte, Wang et al. proposent dans [88] le concept d'*Integrated Infostation – Cellular Network* (IICN), qui est un modèle de système dans lequel plusieurs infostations sont déployées dans une zone cellulaire à faible bande passante. L'idée de base est que les infostations transmettent de façon continue vers les clients et que l'utilisation du réseau cellulaire

n'a lieu que sur demande explicite du client, de façon à éviter le gaspillage des ressources du réseau. Un plan de gestion des caches est nécessaire pour gérer les données reçues à la fois via les deux canaux. Les résultats de simulation montrent une amélioration significative en termes d'épargne de ressources cellulaires, obtenue grâce aux caches, en diminuant le nombre de transmissions cellulaires réalisées en réponse aux demandes explicites des clients.

Iacono et al. étudient le problème de la livraison d'un fichier de type arbitraire via un système d'infostations [52]. Les auteurs supposent que différentes infostations sont réparties le long du trajet d'un utilisateur mobile et que le fichier désiré est segmenté et réparti sur ces infostations. De plus, la population devant être servie par le système est constituée de ce seul et unique utilisateur se déplaçant à vitesse constante, le modèle de mobilité considéré étant de type *random walk*. Les auteurs développent un modèle analytique pour calculer les bornes du délai de livraison du fichier. Les résultats montrent de façon formelle que le délai de livraison est réduit lorsque la vitesse de l'utilisateur augmente. Intuitivement, on comprend aisément que l'utilisateur est servi d'autant plus vite qu'il traverse rapidement les infostations détenant les segments du fichier.

Dans le but de proposer une architecture complémentaire au réseau cellulaire, Eriksson et al. proposent un réseau véhiculaire pour la distribution de contenus, *Cabernet* [41], où les véhicules en mouvement reçoivent les contenus via des points d'accès (*Access Point*, AP) Wi-Fi approchés lors de leurs trajets. Dans leur étude, la connectivité éphémère (de l'ordre de quelques secondes seulement) et la couverture discontinue fournie par les points d'accès sont prises en compte. Afin de traiter ces deux problèmes et dans le but d'améliorer l'expérience des utilisateurs, Cabernet intègre deux composants essentiels : le premier est le *QuickWiFi* qui permet d'établir rapidement une connexion entre les AP opportunistes et un véhicule ; le second est le protocole CTP (*Cabernet Transport Protocol*) qui permet de gérer la congestion sur les liaisons sans fil. La solution proposée a été intégralement implémentée et déployée, notamment à Boston. L'évaluation, menée avec 10 taxis en mouvement dans la ville, montre des résultats très encourageants. Néanmoins, le problème du passage à l'échelle reste posé et demande à être étudié en profondeur.

Pour un même type de contexte, le système *TACO-DTN* a été proposé pour la dissémination de contenus dans des réseaux tolérants aux délais [86]. Ce système repose sur l'existence d'infostations dans le réseau. TACO-DTN est implémenté comme un système *publish/subscribe* [42] permettant de diffuser des informations possédant une durée de validité à des utilisateurs mobiles ayant préalablement exprimé des intérêts pour des types de contenus. Le système s'attaque à plusieurs défis, et en particulier à celui de la gestion des tampons, pour lequel il faut sé-

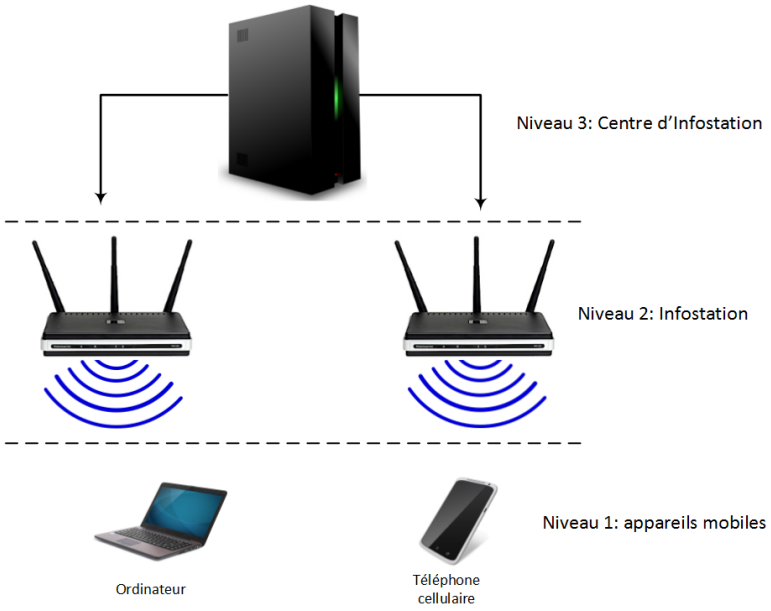


Figure 2.2 – Architecture à 3 niveaux d'un réseau d'infostations.

lectionner les données à stocker en cas de capacité limitée des tampons. Les auteurs proposent également une méthode pour sélectionner les infostations chargées de publier les contenus et de servir les utilisateurs intéressés.

Galluccio et al. vont encore plus loin en introduisant une fonctionnalité supplémentaire au niveau des services des infostations [46]. En effet, les auteurs proposent un cadre détaillant deux processus essentiels de téléchargement et de collecte de contenus par les nœuds mobiles. En d'autres termes, afin d'offrir une connectivité supplémentaire aux utilisateurs en dehors de la couverture des infostations, des nœuds mobiles agissent comme des mules en transportant les données de ces utilisateurs isolés. Les données, ensuite, sont téléchargées lorsque les utilisateurs mobiles se trouvent à proximité d'une infostation. Cela veut dire que les données téléchargées sont recueillies entre deux passages consécutifs sur les infostations. Les auteurs présentent une étude analytique pour évaluer la stabilité du système à infostations. La condition pour obtenir la stabilité est de trouver quand la moyenne du processus de collecte est inférieure à la moyenne du processus de téléchargement, ce qui est prouvé par les résultats obtenus. En revanche, tout comme pour l'étude précédente, le problème du passage à l'échelle n'est pas pris en compte, sachant qu'il peut avoir un impact significatif sur les résultats de simulation. Il serait donc nécessaire de poursuivre l'étude avec un réseau plus complexe, comprenant davantage de nœuds.

Ganchev et al. proposent une architecture de type infostations quelque peu différente, destinée à des services eLearning [47, 48]. Leur architecture permet à un ensemble d'utilisateurs mobiles d'accéder à des services électroniques (eServices) via plusieurs infostations déployées sur un campus universitaire. La figure 2.2 illustre l'architecture de réseau proposée, qui comporte trois niveaux. Le premier niveau est celui qui correspond à l'utilisateur mobile, supposé être équipé d'une technologie sans fil (telle que 4G, Wi-Fi ou Bluetooth). Le deuxième niveau correspond aux infostations proprement dites : elles fournissent un accès intermittent aux différents eServices et sont utilisées comme caches pour les contenus déjà servis. Le troisième et dernier niveau est le centre d'infostations : sa fonction principale est de gérer les mises à jour des informations, ainsi que les profils des utilisateurs. L'évaluation de l'architecture a été réalisée au moyen d'une implémentation et du déploiement d'un système réel. Les services eLearning et eTeaching ont été développés de façon conjointe, afin d'offrir un moyen d'échanger des contenus sans pour autant stresser le réseau cellulaire réel [71, 72]. Les auteurs notent toutefois les lacunes de leur architecture restant à combler avant de parvenir à un système parfaitement opérationnel, telles que le coût et la facturation des utilisateurs, ou encore le nombre de ces derniers.

2.2.2 Délestage des données en espace libre (via Wi-Fi)

Le délestage des données (*mobile data offloading*) est considéré actuellement comme étant l'une des approches les plus prometteuses pour diminuer la charge des réseaux à infrastructure. Il suscite de ce fait un intérêt croissant dans la communauté scientifique. L'idée de base consiste à exploiter le spectre non employé par les autres technologies (tel que celui du Wi-Fi) pour y basculer une partie du trafic initialement destiné à être convoyé par le réseau cellulaire (voir figure 2.3). Cela permet naturellement d'alléger la charge du réseau cellulaire et de diminuer le coût de transmission. Le délestage de trafic a été largement étudié dans la littérature et des états de l'art détaillés sont présentés dans [8, 80]. Dans un souci de concision, nous nous limiterons ici aux propositions qui nous semblent les plus proches de ce travail de thèse.

D'après Lee et al., il existe deux types principaux de délestage de données [62]. Le type le plus courant est appelé *on-the-spot offloading* : il consiste à rediriger tout le trafic mobile vers les AP, dès lors qu'un réseau Wi-Fi est disponible ; dans le cas contraire, le trafic continue à transiter par le réseau cellulaire. Le second type, le *delayed offloading*, ne traite que les informations pouvant tolérer un délai supplémentaire avant d'être délivrées à leurs utilisateurs intéressés : à chaque contenu est associée une durée de validité (*deadline*) qui a un sens tant du point de

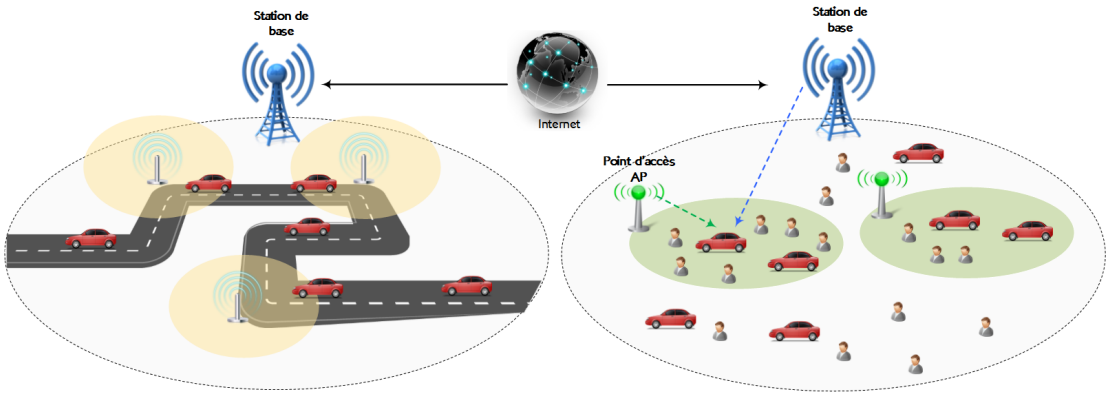


Figure 2.3 – Scénarios de délestage de données (*mobile data offloading*).

vue de l'utilisateur que de celui de l'application finale ; les transmissions ont lieu lorsqu'un utilisateur destinataire entre en contact avec un point d'accès Wi-Fi, mais à l'approche imminente de l'échéance, le transfert de bout-en-bout est achevé au moyen de transmissions cellulaires. Les deux types de délestage peuvent ainsi réduire la charge sur l'infrastructure cellulaire et en épargner les ressources.

Le positionnement des points d'accès Wi-Fi a été largement étudié dans la littérature [17, 35, 82]. Dimatteo et al. proposent un algorithme pour le positionnement des AP dans un réseau cellulaire [35]. Les données considérées correspondent à de grosses quantités d'information (données issues d'expériences scientifiques) et des fichiers vidéo (c.-à-d. fichier AVI), qui vont occuper 82% du trafic mondial d'ici 2020 [29]. Ces données ont comme caractéristique commune d'être tolérantes au délai. L'algorithme a été évalué avec une trace de mobilité réelle, et les auteurs reportent des résultats très encourageants : près de 50% du trafic est délesté via le Wi-Fi, même avec un nombre relativement faible d'AP. Un point négatif est cependant à signaler : nous notons que l'algorithme proposé repose sur un déploiement séquentiel ; de ce fait, il ne peut en aucun cas garantir de fournir le plan de déploiement qui optimiserait le processus de délestage.

Ristanovic et al. proposent l'algorithme *HotZones* qui permet aux AP de couvrir l'ensemble de la zone géographique occupée par le réseau cellulaire [82]. HotZones positionne les AP en fonction de la densité de population des macro-cellules, avec l'idée que plus la cellule est dense, mieux cela devrait être. Cependant, une sous-région d'une macro-cellule dense peut très bien ne pas générer autant de données qu'une sous-région d'une autre macro-cellule moins dense, d'autant que les régions denses ne sont pas nécessairement celles qui génèrent les quantités les plus importantes de données. Avec la même problématique que les deux articles susmention-

nés, Bulut et Szymanski développent un algorithme de positionnement qui cherche à maximiser la capacité de délestage en fonction de la densité des requêtes des utilisateurs [17]. Pour ce faire, les auteurs réalisent une modélisation par un programme linéaire, qui s'avère être NP-complet. L'heuristique utilisée pour la résolution montre que le trafic cellulaire pourrait être réduit de l'ordre de 20% à 65%.

Outre le problème du positionnement des AP, le délestage de données mobiles a été étudié sous d'autres angles. En particulier, les solutions à base de réseaux de véhicules (VANETs) occupent une part de lion. Wu et al. présentent le système iCAR (*integrated Cellular and Ad hoc Relaying* -– relayage cellulaire et ad hoc intégré) qui répartit de façon efficace le trafic entre l'infrastructure cellulaire et les stations ad hoc [91]. Se basant sur l'idée développée dans iCAR, et en utilisant des modèles de mobilité d'utilisateurs, Yoon et al. ont conçu leur propre système qui suppose que chaque véhicule est doté d'interfaces cellulaire et Wi-Fi : la première est dédiée au plan de contrôle, à des fins de signalisation, alors que la seconde prend en charge les transmissions de données proprement dites [97]. Les points d'accès sont situés aux intersections les plus importantes du réseau routier et sont reliés entre eux par des liens sans fil ; en plus de ces liens, un ou plusieurs de ces AP sont inter-connectés en filaire. L'objectif principal est donc de maximiser la quantité de données transmises par les AP déployés. Les simulations effectuées avec des cartes routières de la Suisse montrent des résultats de performance encourageants, avec des économies en termes de coût. Nous remarquons toutefois que les auteurs ne spécifient rien quant à la sélection des AP à connecter en filaire.

Ahn et al. étudient le problème de réduction du coût pour la diffusion d'une information aux VT (*Vehicle Terminal*) via des communications V2I (*Vehicle-to-Infrastructure*) [7]. En d'autres termes, l'objectif est de maximiser le nombre d'utilisateurs obtenant un contenu considéré, tout en minimisant l'utilisation de l'infrastructure cellulaire. Les informations échangées sont considérées comme tolérantes au délai. Les auteurs analysent mathématiquement le processus de diffusion en présentant le problème sous forme d'un problème d'optimisation. Les résultats d'analyse sont obtenus par simulation avec des traces de mobilité réelle (trace véhiculaire de Pékin) : ils mettent en évidence qu'un contenu peut très bien être diffusé à l'ensemble des véhicules intéressés à moindre coût. Il est nécessaire aussi de noter que le modèle proposé considère l'effet de la popularité des intérêts des utilisateurs, et qu'il a un impact très important sur les résultats obtenus. Par contre, pour des raisons de simplicité, les auteurs ont supposé que les rencontres entre deux voitures quelconques sont identiques. Cependant, la probabilité que deux véhicules se rencontrent est liée à plusieurs paramètres comme par exemple : leur proximité, leurs vitesses et leurs habitudes. Ces derniers peuvent influer sur le modèle proposé.

Malandrino et al. proposent quant à eux l'utilisation de RSU (*Road Side Unit*) afin de réaliser leur système de délestage. Malandrino et al. détaillent leur modèle analytique de performances [68]. Le principe est le suivant : dans un premier temps, on calcule la fraction maximale du contenu que les utilisateurs pourraient obtenir directement de la part de l'infrastructure et sans utilisation supplémentaire des ressources de cette dernière. Dans un second temps, une prédiction de mobilité des utilisateurs est réalisée afin de planifier les transmissions du contenu restant à faire, des RSU vers les véhicules. Nous notons toutefois que l'article ne considère pas les transferts entre RSU alors qu'ils pourraient affecter l'efficacité de la solution.

Plus récemment, une définition de la satisfaction des utilisateurs a été proposée dans [89]. Les auteurs partent du fait qu'a priori les utilisateurs cherchent à obtenir un contenu non seulement dans le laps de temps le plus bref possible, mais également au coût le faible possible. Il s'agit alors d'obtenir un compromis entre tolérance au délai et coût de transmission. Dans cette proposition, les utilisateurs conservent la possibilité de télécharger un contenu à tout moment, à partir du lien cellulaire ; en plus, ils ont la possibilité d'obtenir ce même contenu à partir des *RSU* (Road-Side Unit) installées le long des routes, et ce à un coût réduit. La formulation du problème revient alors à une maximisation de la satisfaction des utilisateurs et à la sélection de l'interface de réception adéquate. Les résultats, obtenus avec une grande trace de mobilité, montrent un bon taux de délestage et de satisfaction des utilisateurs.

Wiffler est un cadre proposé par Balasubramanian et al. qui cible des applications tolérantes au délai [15]. Wiffler exploite à la fois la tolérance au délai que possède certains types de contenus et une commutation rapide entre deux technologies de transmission (Wi-Fi et 3G) pour faire face à la disponibilité fluctuante du Wi-Fi. Des mesures détaillées rapportent que la moitié du trafic cellulaire peut être pratiquement délestée, avec une tolérance de l'ordre de la minute. Malheureusement, le nombre maximum d'utilisateurs lors des mesures effectuées est de 20, et nous n'avons aucune idée du comportement du système lorsque la population d'utilisateurs augmente (c.-à-d. passage à l'échelle).

Lee et al. présentent une étude quantitative des performances du délestage de donnés mobiles via le Wi-Fi [62]. Ils ont suivi à la trace une centaine d'utilisateurs piétons, et ont ainsi pu collecter des statistiques sur leur connectivité Wi-Fi à l'intérieur d'une zone métropolitaine. Leurs simulations, à partir de cette trace, montrent que les AP Wi-Fi pourraient prendre en charge 65% du trafic total de données mobiles, tout en économisant 55% de la batterie des terminaux, et cela sans introduire de délai supplémentaire pour les utilisateurs. Un autre résultat intéressant est également présenté : considérer des données tolérantes au délai et accepter un délai de 100 secondes ne permet d'obtenir qu'un gain inférieur à 3%, ce qui pourrait paraître

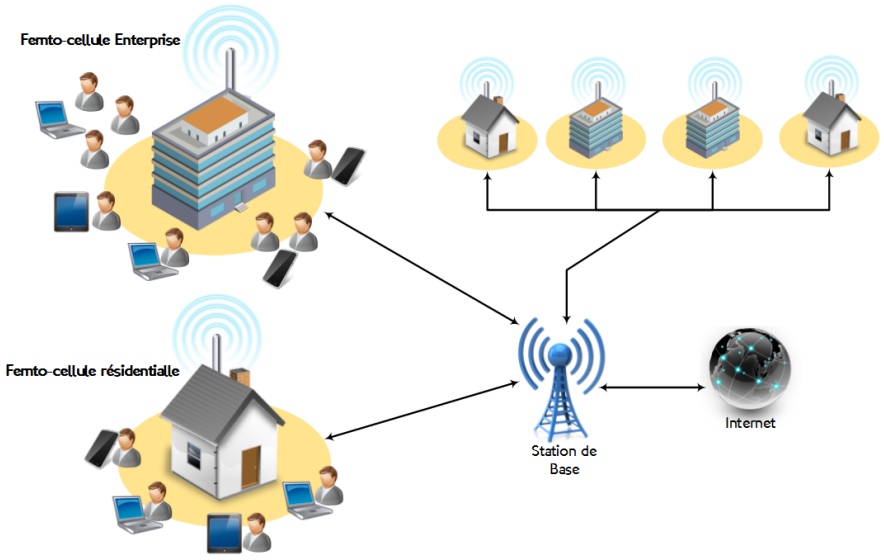


Figure 2.4 – Les Femtocellules déployées dans des environnements résidentiels et des entreprises.

surprenant, mais qui s'explique par le fait de la faible vitesse de mouvement des utilisateurs -- piétons -- et par conséquent des durées de connexion plus importantes que dans le cas de véhicules.

Xu et al. proposent une application innovante appelée OAD (*On-roads Ads Delivery*), qui permet la diffusion d'annonces publicitaires ciblées afin d'inciter les utilisateurs à bord de véhicules à s'arrêter sur leurs trajets quotidiens (domicile -- lieu de travail) et à se transformer en clients consommateurs [64]. Le système repose sur la mise en œuvre de points d'accès routiers (*RSAP – Road Side Access Point*) afin de préserver le spectre cellulaire. L'étude est formulée comme un problème d'allocation de bande passante, dans lequel l'emplacement des RSAP est prédéterminé. Les auteurs fournissent plusieurs heuristiques pour échelonner les messages publicitaires ciblés et pour allouer la bande passante des AP. Les simulations réalisées mettent en évidence la diminution de charge sur le réseau cellulaire.

2.2.3 Délestage de données en espace restreint (via Femtocellules)

La technologie des *Femtocells* offre des solutions possibles au problème de la surcharge des réseaux cellulaires en utilisant le spectre exploitable sous licence, contrairement au Wi-Fi [9, 23]. Elle nécessite le déploiement de stations de base *plug-and-play* à courte portée dans les environnements résidentiels ou les entreprises : la figure 2.4 met en évidence le fait que cela évite à des utilisateurs d'être servis directement par les macro-cellules et donc le gaspillage de res-

sources cellulaires. Néanmoins, étant donné que les Femtocellules utilisent la même partie du spectre que les macro-cellules, il peut y avoir des interférences. Elles peuvent être directement liées au déploiement (surtout lorsqu'il est non planifié) et à la configuration des Femtocellules. L'optimisation de la localisation Femtocellulaire est un enjeu fondamental, qui a pour objectif de fournir une qualité de service (QoS) à l'utilisateur avec un coût minimum du système. Un rapport du 3GPP (*3rd Generation Partnership Project*) identifie trois critères pour déployer des Femtocellules [5] :

— *Canal dédié vs co-canal* : soit un canal dédié est attribué à la Femtocellule, soit un même canal est partagé entre la Femtocellule et la macro-cellule [22, 67].

— *Accès ouvert vs groupe fermé d'abonnés* : soit tous les utilisateurs des macro-cellules peuvent accéder à la Femtocellule, soit l'accès à la Femtocellule est restreint à un groupe d'utilisateurs identifiés [31, 95].

— *Puissance d'émission descendante fixe vs puissance d'émission descendante adaptative* : soit la puissance d'émission de la Femtocellule est prédéfinie et fixe, soit elle varie dans le temps pour s'adapter en tenant compte de la station de base de la macro-cellule et des Femtocellules voisines [24].

Kolding et al. s'intéressent à la question du déploiement des Femtocellules et déterminent combien de macro-cellules peuvent déplacer le trafic vers les Femtocellules [58]. Pour cela, différentes stratégies de déploiement de points d'accès Femtocellule ont été envisagées avec leurs configurations ainsi que la façon dont les utilisateurs sont affectés à chaque point d'accès. Pour augmenter l'effet du délestage, ils proposent d'utiliser une combinaison d'étalonnage de puissance adaptatif et de plan d'allocation de ressources prenant en compte le rapport signal sur interférence plus bruit (SINR). Dans les résultats présentés, le trafic délesté vers les Femtocellules peut atteindre jusqu'à 20% du trafic. Yavus et al., quant à eux, s'attaquent au problème des interférences entre le réseau cellulaire et les Femtocellules installées [94]. Dans cet article, ils décrivent un modèle de gestion à la fois de la liaison montante et de la liaison descendante qui vise à réduire les interférences et par conséquent à améliorer la qualité d'expérience des utilisateurs. Leurs résultats de simulation montrent que l'utilisation de Femtocellules amène une nette amélioration de la capacité et de la couverture réseau.

Hu et al. comparent le gain obtenu avec l'utilisation des Femtocellules par rapport aux techniques basées sur les AP (en Wi-Fi), en termes de QoS et notamment d'amélioration du débit [51]. Les auteurs développent un modèle analytique permettant de calculer les bornes d'une

stratégie optimale de délestage. Les simulations menées sur un scénario urbain et réaliste à grande échelle mettent en évidence une augmentation linéaire du débit lorsque le nombre de points d'accès Femtocellules augmente. L'illustration des résultats montrent que l'opérateur a besoin de 1200 Femtocellules/km^2 pour atteindre les performances requises. À notre sens, et pour notre problématique, la faiblesse principale du modèle se situe dans le fait qu'il ne concerne que les utilisateurs en intérieur ; en d'autres termes, la mobilité des utilisateurs (c.-à-d. les utilisateurs en extérieur, en plein air) n'est pas prise en compte, ce qui peut affecter les résultats.

Chaussen et al. étudient les avantages qu'aurait un opérateur cellulaire à utiliser des Femtocellule pour délester le trafic des utilisateurs en intérieur [32]. Ils indiquent que ces utilisateurs souffrent de mauvaises conditions de canal, comparé aux utilisateurs qui sont en extérieur, et que la commutation de leur trafic vers les Femtocellules pourrait entraîner un gain potentiel de ressources pour les macro-cellules. En d'autres termes, servir les utilisateurs ayant de faibles conditions de canal avec des Femtocellules libère des ressources pour les macro-cellules exploitables pour servir d'autres utilisateurs en extérieur. Dans leur article, les auteurs étudient un certain nombre de scénarios afin de quantifier l'intérêt du délestage cellulaire dans un environnement réaliste de déploiement de Femtocellules. Le critère de performance considéré est le gain de capacité du processus de délestage, qui est défini comme étant le nombre d'utilisateurs en extérieur servis en plus et en même temps qu'un utilisateur est servi par une Femtocellule. Les scénarios sont simulés et comparés à une situation de base, dans laquelle seule la macrocellule peut servir les utilisateurs, qu'ils soient en extérieur ou en intérieur. Les résultats de simulation indiquent que le gain obtenu avec le délestage peut atteindre jusqu'à 30%, sachant que les valeurs obtenues dépendent à la fois de la puissance d'émission de la macro-cellule et de la densité de sa population d'utilisateurs.

Deux études plus récentes se sont penchées sur la même question, à savoir les bénéfices apportés par le délestage au moyen de Femtocellules. Chandar et al., sachant qu'un déploiement dense de Femtocellules peut provoquer des interférences de façon aléatoire et non contrôlée, surtout en mode co-canal, proposent un modèle analytique pour évaluer l'augmentation de capacité attendue pour des réseaux à macro et Femtocellules utilisant le mode co-canal avec la technique d'accès OFDMA (*Orthogonal Frequency-Division Multiple Access*) [21]. Ils étudient notamment le gain du délestage, en termes de nombre d'utilisateurs supportés dans le système et l'impact des configurations de puissance des macro-cellules sur les performances du réseau. Leur étude montre que les performances du réseau dépendent largement des puissances utilisées dans les macro-cellules et de la charge du réseau, que le gain augmente avec la densité

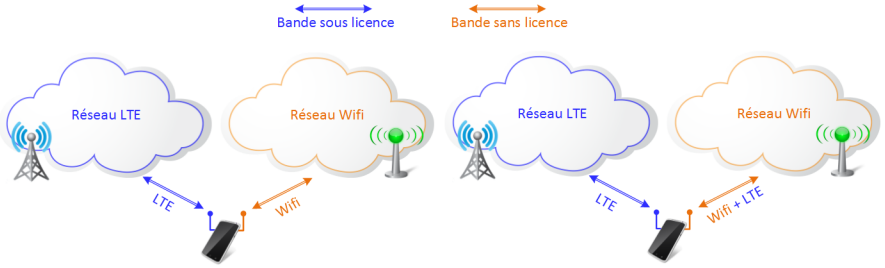


Figure 2.5 – Évolution vers un système LTE-U (intra-réseau).

des Femtocellules et également que ce gain peut être accru en choisissant de façon adéquate la puissance de transmission et la charge des Femtocellules en fonction des conditions de trafic du réseau.

Mitra et al. étudient le délestage pour les réseaux LTE-Femtocellule, leur objectif étant de maximiser le nombre d'utilisateurs dans le système ainsi que le débit moyen [73]. L'effet des différents paramètres des Femtocellules sur le processus de délestage est analysé, ce qui permet des préconisations pour le déploiement des Femtocellules. Les résultats obtenus en utilisant un délestage par Femtocellules pour les utilisateurs en intérieur montrent que le débit peut être augmenté jusqu'à 170 fois sa moyenne et que le système peut supporter jusqu'à 20% d'utilisateurs supplémentaires.

2.2.4 Équilibrage du trafic (coexistence de LTE et du Wi-Fi dans le spectre sans licence)

Ce type d'approche est justifié par la forte limitation du spectre sous licence et de son manque de disponibilité. Pour faire face à la croissance du trafic et aux demandes des utilisateurs, le groupement 3GPP a travaillé, dans le cadre de sa "Release 13", sur une solution nommée LTE-U *(LTE-Unlicensed)* [74]. La figure 2.5 illustre le principe qui consiste à agréger le spectre sous licence avec le spectre sans licence. Une description complète est donnée dans un autre document, tant du point de vue de l'opérateur que des perspectives de l'utilisateur [100]. Cet article présente également les principaux défis auxquels le LTE-U doit faire face. Entre autres, citons le fait que les deux technologies reposent sur des méthodes d'accès différentes : le Wi-Fi exploite une méthode d'accès basée sur une contention spécifique [19], alors que le LTE utilise des méthodes de type TDMA/OFDMA. Par conséquent, il est nécessaire de modifier les méthodes de gestion de ressources et les méthodes d'accès afin d'assurer une bonne coexistence. En d'autres termes, le délestage dans le cadre d'une diffusion de contenus doit être revu. En

effet, contrairement aux techniques plus anciennes qui basculent le trafic vers l'interface Wi-Fi, il s'agit ici, avec le LTE-U, de prendre en compte les différentes activités des utilisateurs pour le spectre sous licence et pour le spectre sans licence, puisqu'ils partageront les mêmes ressources. Des résultats de simulation montrent un débit meilleur dans des petites cellules que celui obtenu avec du Wi-Fi.

Dans le même genre d'approche, Chen et al. relèvent que le délestage de données et le partage de ressources sont les deux méthodes communément proposées pour véhiculer des données cellulaires sur un spectre sans licence [27]. Ils développent alors une méthode hybride qui vise à cumuler les avantages des deux précédentes : les stations de base cellulaires délestent le trafic vers les réseaux Wi-Fi et occupent en même temps des tranches de temps sur la bande passante sans licence. Une étude analytique permet de comparer le débit cellulaire obtenu avec leur méthode hybride et celui des deux autres méthodes. Les résultats montrent, d'une part, que le délestage obtient de meilleures performances que le partage de ressources lorsque le nombre d'utilisateurs Wi-Fi n'excède pas un certain seuil, et d'autre part, que la méthode hybride est plus efficace que le seul délestage et obtient les mêmes performances que le partage de ressources lorsque le nombre d'utilisateurs Wi-Fi est suffisamment important.

Chen et al. revisitent le délestage de trafic [26]. En effet, inversement aux méthodes classiques de délestage de données qui basculent du trafic LTE vers les réseaux Wi-Fi, ils proposent de basculer du trafic Wi-Fi vers le système LTE. La justification réside dans le fait que l'utilisation du spectre Wi-Fi pourrait être gâchée par un nombre trop important de collisions de trames en présence de trop nombreux utilisateurs Wi-Fi. L'idée est donc d'utiliser le LTE-U et d'allouer des ressources radio sans licence au système LTE afin de lui donner une compensation pour les utilisateurs qu'il doit gérer en plus. Trois questions se posent alors : combien d'utilisateurs Wi-Fi transférer ? Quels utilisateurs Wi-Fi transférer ? Quelle quantité de ressources LTE-U faut-il ? Les auteurs étudient plusieurs stratégies de transfert d'utilisateurs, pour lesquelles ils analysent la quantité minimum de ressources sans licence nécessaires pour un nombre donné d'utilisateurs Wi-Fi basculés. Nous notons par conséquent que leur proposition ne fonctionne qu'en présence de ressources LTE-U suffisantes.

Initialement, les technologies LTE et Wi-Fi ont été conçues pour opérer dans des bandes de fréquences distinctes. Or, étant donnée la rareté des ressources radio, il faut s'attendre à ce que ces deux technologies se mettent à coexister dans une même bande de fréquences sans licence. Abinader et al. s'intéressent donc aux problèmes liés au fonctionnement simultané de LTE et du Wi-Fi, du point de vue des ressources radio [6]. Leur analyse détaillée des fonctionnalités et caractéristiques des deux technologies montre en effet que le Wi-Fi peut être

sévèrement impacté par les transmissions LTE. Pour pallier ce problème, les auteurs listent des mécanismes de coexistence possibles, et montrent que l'utilisation de la fonctionnalité ABS *(Almost Blank Subframe)* définie dans la "Release 10" du 3GPP pour le LTE et le contrôle de la puissance d'émission sur les liens montants peuvent toutefois contribuer à amoindrir le problème. Cavalcante et al. se sont aussi penchés sur ce problème de coexistence [20]. L'angle d'attaque est moins fonctionnel, et résolument quantitatif. Leur évaluation des performances du réseau est faite par simulation, en considérant un contexte professionnel (c.-à-d. un certain nombre de bureaux sur un étage unique, ou sur plusieurs étages) et plusieurs configurations possibles (déploiement clairsemé ou dense des points d'accès et des stations). Le principal critère de performance retenu est le débit vu par l'utilisateur. Les résultats montrent une inéquité des performances : en cas de coexistence LTE/Wi-Fi, alors que le débit du LTE n'est pratiquement pas dégradé (de 0% à 4,5% dans le cas pire), celui du Wi-Fi peut chuter de 70% à 100%. Des pistes (telles l'utilisation de la fonctionnalité ABS de LTE précédemment citée) sont fournies pour limiter le phénomène.

Jeon et al. présentent une étude du fonctionnement des systèmes LTE dans le spectre sans licence, lorsqu'il est utilisé en même temps par d'autres technologies (p. ex. le Wi-Fi) [54]. Ils s'intéressent en particulier aux interférences inter-systèmes. Les résultats de leur modèle mathématique montrent que l'opérateur doit diminuer le rayon de sa cellule LTE pour diminuer l'impact de la coexistence sur ses performances. Néanmoins, cela provoque le blocage de davantage de stations Wi-Fi, ce qui réduit ses performances.

2.3 Solutions sans infrastructure cellulaire

Dans cette partie, nous continuons à nous intéresser aux techniques et stratégies qui permettent d'économiser les ressources dans les réseaux cellulaires, avec un focus sur les *communications device-to-device* (D2D), actuellement considérées comme partie intégrante de la future génération de réseaux. Ce paradigme est défini comme étant la communication entre deux terminaux mobiles sans nécessiter l'utilisation d'une infrastructure cellulaire. A première vue, les communications D2D présentent donc des points communs avec les réseaux ad hoc mobiles (MANET) [49]. Il existe cependant des caractéristiques qui les différencient nettement. En effet, les communications D2D peuvent utiliser le spectre sous licence (in-band) comme le spectre sans licence (out-band) ; par ailleurs, les échanges entre terminaux peuvent être contrôlés par un AP comme se dérouler de façon autonome (sans aucun contrôle d'un AP quelconque). Les communications D2D vont ainsi permettre une augmentation de l'efficacité

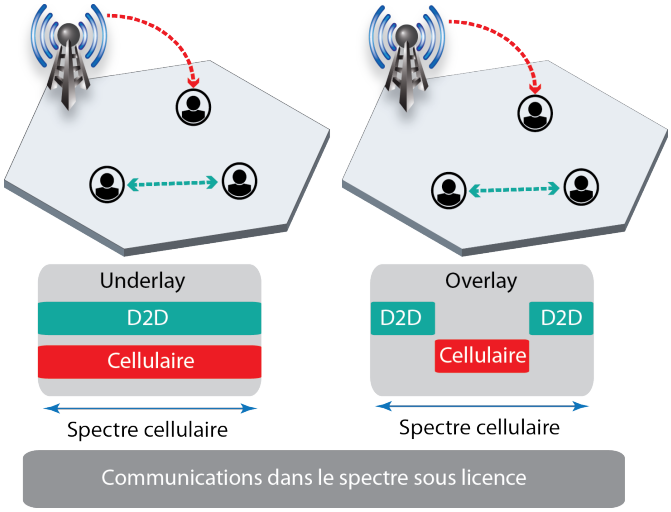


Figure 2.6 – Communications D2D : *underlay* vs *overlay*.

spectrale, une diminution de la consommation d'énergie ainsi qu'une amélioration du débit. Dans un souci de concision, nous nous limitons dans ce qui suit aux publications de recherche qui ont un lien direct avec nos travaux de thèse. La figure 2.6 illustre le partage du spectre entre les communications D2D sous licence (underlay, overlay). Les communications underlay (sous-jacentes) ne peuvent pas être traitées simultanément en raison du spectre partagé, ce qui produira des interférences à son tour. Cependant, les communications overlay (superposition) et les communications D2D peuvent avoir lieu simultanément en raison de l'absence du partage du spectre.

2.3.1 Les communications D2D *underlay*

Pendant des années, le spectre sous licence a été explicitement attribué à l'utilisation de l'infrastructure cellulaire. Malheureusement, l'utilisation qu'en fait cette dernière n'est pas efficace [56]. En vue de l'améliorer, le concept des *communications underlay* a été proposé. Le principe est de partager le spectre cellulaire entre l'infrastructure elle-même et les différents abonnés mobiles en communication D2D. Si l'idée paraît simple, elle ne soulève pas moins plusieurs questions, telles l'allocation de ressources et la gestion des interférences. De nombreux travaux de recherche ont traité ces deux problèmes afin de garantir un niveau de QoS raisonnable, à la fois pour les utilisateurs cellulaires et pour les utilisateurs opportunistes.

Kaufman et al. proposent ainsi un protocole de partage de spectre, dans lequel le réseau D2D peut utiliser les mêmes fréquences que le réseau cellulaire [56]. Pour ce faire, un utilisateur

D2D mesure la puissance du signal reçu sur sa liaison descendante pour déterminer son affaiblissement sur le trajet vers la station de base du réseau cellulaire ; il adapte ensuite sa puissance d'émission en fonction de cet affaiblissement de manière à pouvoir communiquer directement avec les autres utilisateurs D2D, tout en ne causant qu'un minimum d'interférences pour la station de base. Un modèle analytique du protocole permet de calculer la probabilité qu'une liaison D2D ne provoque pas la rupture de la liaison cellulaire par non-respect du SINR.

Peng et al. se placent dans le cas où la liaison cellulaire montante est réutilisée par les communications D2D, et présentent deux mécanismes d'évitement d'interférences [77]. Le premier vise à limiter les interférences générées par les transmissions cellulaires sur les communications D2D, et repose sur un traçage : sachant qu'un équipement D2D peut, grâce au canal de contrôle cellulaire, savoir quelles ressources radio ont été allouées et à quels utilisateurs cellulaires, il va pouvoir mesurer et tracer les interférences en provenance de ces derniers, et donc adapter ses transmissions pour éviter toute interférence nuisible de la part des communications cellulaires. Inversement, le second mécanisme a pour objectif de limiter les interférences des transmissions D2D sur les communications cellulaires : il suppose l'assistance de la station de base cellulaire, chargée de calculer le niveau d'interférences tolérable et de le diffuser à tous les utilisateurs D2D qui pourront alors choisir les meilleures ressources radio à utiliser. Les deux mécanismes peuvent être utilisés conjointement ou séparément. Les résultats de simulation montrent qu'on peut attendre une augmentation du débit sur l'ensemble du réseau de l'ordre de 41%.

Zhang et al. étudient le même problème, à savoir l'allocation des ressources radio prenant en compte les interférences entre système cellulaire et communications D2D, mais avec d'autres outils, puisqu'ils utilisent la théorie des graphes [101]. Ils représentent les interférences existant entre les différentes liaisons D2D et cellulaires dans un graphe, dans lequel un sommet représente une communication, qu'elle soit cellulaire ou D2D, et une arête une interférence mutuelle possible entre les deux sommets adjacents. Le problème de l'allocation des blocs de ressources (c.-à-d. les plages de fréquences) s'avère NP-difficile, d'autant plus que le nombre de communications est important. Les auteurs proposent alors un algorithme heuristique et montrent par simulation qu'il fournit d'excellentes performances tout en présentant une complexité de calcul réduite.

Yu et al. traitent le partage des ressources en supposant que le réseau cellulaire (plus précisément une station de base) peut contrôler la puissance de transmission et les ressources radio des communications D2D [98, 99]. Deux modes de partage de ressources sont considérés : *orthogonal* et *non orthogonal*. Les auteurs cherchent à maximiser le débit global, avec une

contrainte de puissance de transmission maximum et des contraintes de débits pour les utilisateurs. Ils considèrent pour cela deux cas. Dans le premier cas, il n'y a pas de priorité et les deux paradigmes de communication sont en concurrence pour accéder aux ressources ; la station de base exécute un algorithme de maximisation de débit de type glouton (*greedy*). Dans le second cas, la priorité est donnée à l'utilisateur cellulaire (qui a un débit minimum garanti). Leurs résultats de simulation montrent une franche augmentation du débit global et donc tout l'intérêt à faire prendre en charge du trafic local par des communications opportunistes, même en présence de priorités pour les utilisateurs cellulaires.

Doppler et al. étudient les aspects protocolaires à prendre en compte pour l'intégration de communications D2D avec un réseau LTE-Advanced [37, 38]. Leur principal objectif est d'identifier les fonctionnalités à ajouter ou les modifications à apporter dans l'architecture SAE (*System Architecture Evolution*) du réseau cellulaire, notamment pour l'établissement et la gestion des sessions D2D. Ils étudient également les mécanismes permettant de contrôler et de limiter les interférences générées par les communications D2D sur le réseau cellulaire. Leurs contributions sont validées par simulation, avec un scénario basé sur un réseau cellulaire local. Les résultats montrent que, si les utilisateurs D2D sont suffisamment proches et dans la même pièce, le débit global peut être accru de 65%.

Li et al. s'intéressent aux problèmes de partage de ressources et de coordination entre utilisateurs cellulaires et utilisateurs D2D [53]. Ils proposent l'utilisation d'un plan de communication D2D avec une technologie de relais incrémental en supposant que la station de base peut lire et comprendre toute information émise par un utilisateur D2D sur un lien opportuniste. De plus, la station de base sert de solution de repli lorsque la liaison D2D échoue à véhiculer l'information ; en cela et comme elle transmet des informations manquantes, elle contribue à améliorer le débit des communications D2D. Le problème est ensuite formulé comme un problème de maximisation de débit, avec la prise en compte de trois modes de transmission : transmission cellulaire classique, transmission *underlay* et utilisation d'un relais incrémental. Les résultats de simulation montrent que le relais incrémental proposé est plus fiable que les deux autres alternatives d'un ordre de 40% . Cependant, nous remarquons qu'en cas de défaillances multiples des liaisons D2D, la station de base doit retransmettre les informations manquantes et cela implique une consommation significative de ressources sur le lien cellulaire descendant.

2.3.2 Les communications D2D *overlay*

Contrairement au cas des communications D2D *underlay*, le principe consiste ici à consacrer une partie des ressources radio aux communications purement D2D en fonction de la charge du réseau [13]. De nombreux articles traitent du multicast dans les réseaux cellulaires, multicast réalisé grâce à un découpage en groupes (*clustering*). En particulier, Zhou et al. se concentrent sur les retransmissions de données lorsque certains utilisateurs intéressés n'ont pu recevoir le contenu souhaité [102]. Pour améliorer les performances, ils proposent une architecture dans laquelle les utilisateurs sont regroupés en communautés (clusters) dans le réseau et où une tête de groupe est élue par communauté. La raison derrière le choix de ce dernier est d'épargner les ressources à consommer par les communications D2D étant qu'elles utilisent une partie du spectre sous licence dédié. Lorsqu'une station de base effectue une transmission multicast, en raison des propriétés de la liaison cellulaire descendante, seules les têtes de groupes sont capables de décoder les informations transmises ; chaque tête de groupe est donc chargée de retransmettre le contenu au sein de sa communauté et ainsi de jouer un rôle de relais. La question primordiale est alors de déterminer les têtes de groupes et leur nombre, qui vont minimiser le temps de réponse ainsi que le nombre de ressources consommées. Les auteurs proposent un modèle analytique fournissant la fonction de densité du nombre optimal de relais, une stratégie de retransmissions optimisant l'utilisation des ressources en sélectionnant de façon adaptative le nombre de relais, ainsi qu'un algorithme de découpage en groupes qui améliore le débit des retransmissions. Leurs résultats montrent un gain important en termes d'utilisation de ressources par rapport à un nombre fixe de relais.

Seppala et al. étudient également le problème du multicast fiable dans le même contexte [84]. Les auteurs proposent de faire contrôler le processus d'allocation et la puissance d'émission par la station de base afin de garantir une QoS suffisante, et de faire prendre en charge les retransmissions par des liaisons D2D afin de garantir la fiabilité du multicast. En d'autres termes, la station de base a tout le contrôle du réseau, y compris les communications D2D. Leurs résultats de simulation montrent que le partitionnement en groupes des utilisateurs géographiquement proches et avec des besoins de communication proches permet d'accroître significativement leur débit.

Les travaux de Peng et al. se concentrent sur les stratégies de regroupement pour affiner l'utilisation du D2D dans les transmissions multicast [76]. Dans cet article, l'objectif visé est de minimiser le temps de réponse. Les contenus considérés sont des fichiers multimédias que la station de base envoie aux utilisateurs présentant les meilleures conditions de canal, à charge

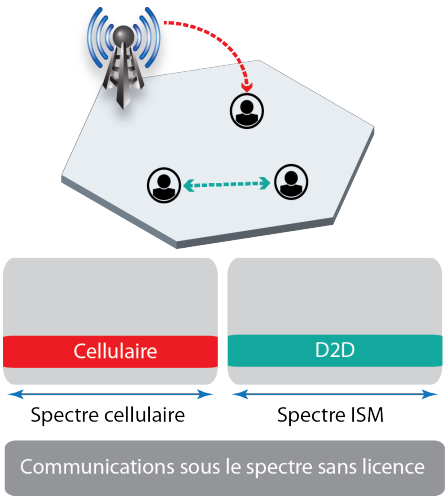


Figure 2.7 – Communications D2D : utilisation d'un spectre sans licence.

pour ces derniers de les faire suivre aux autres destinataires au sein du même groupe. Ils présentent un modèle pour la sélection des groupes, dont l'analyse repose sur la théorie des jeux. Les simulations réalisées montrent que leur stratégie diminue fortement le temps de réponse moyen sous une forte densité d'utilisateurs.

Koskela et al. étudient également l'apport des groupes dans un réseau cellulaire FDD (Frequency Division Duplex) pour le multicast, en considérant des contraintes au niveau des interférences [59, 75]. Ils présentent dans un premier temps un modèle analytique permettant de choisir, en fonction de l'ensemble des utilisateurs et de l'état du réseau, le mode optimal de communication entre deux utilisateurs. Dans un second temps, ils étendent leur système d'équations pour prendre en compte les groupes, sous l'hypothèse qu'au sein d'un même groupe il y a un utilisateur partageant un fichier avec les autres membres de son groupe et que ces derniers représentent globalement un seul destinataire. Les résultats, analytiques et de simulation, montrent que, lorsque les membres du groupe sont proches, le mode par défaut peut être le D2D avec les performances optimales. Dans le cas contraire, le gain obtenu pour l'affaiblissement de propagation en D2D par rapport aux communications cellulaires conventionnelles diminue à cause des interférences avec les autres utilisateurs. Dans cet article, les auteurs n'abordent pas les problèmes de fiabilité du multicast, de choix des têtes de groupes, et de partitionnement en groupes.

2.3.3 Les communications D2D sur la bande de fréquence sans licence

Dans cette section, nous examinons les communications D2D qui utilisent le spectre sans licence [83]. L'utilisation d'un spectre sans licence offre des avantages évidents, à commencer par l'absence d'interférences avec les communications cellulaires puisque ces dernières utilisent un spectre sous licence (cf. Fig 2.7). Cependant, elle n'en pose pas moins des problèmes à traiter avant d'exploiter complètement ces avantages. Tout d'abord, les terminaux n'ont pas un accès exclusif à la bande de fréquences (notamment dans le cas de la bande ISM) et des interférences sont donc possibles avec les autres types d'équipements. Ensuite, les terminaux étant très disparates du point de vue du débit et de la durée de transmission, il est difficile de garantir une utilisation efficace du spectre pour tous. Enfin, le spectre étant partagé de façon libre entre tous, il n'y a pas d'incitation pour un terminal à l'utiliser de façon à améliorer les performances de tous.

Zhou et al. se sont ainsi penchés sur le problème de la coexistence de communications D2D et conventionnelles dans la même zone géographique et sur la même bande de fréquences (dans leur étude, la bande ISM) [103]. Pour éviter les contentions pour l'accès au canal et les collisions de paquets, le mécanisme qu'il propose intègre une écoute de la porteuse et une pré-allocation des ressources par le point d'accès aux utilisateurs D2D ; l'idée sous-jacente est que le point d'accès forme des groupes de paires D2D ayant les mêmes contraintes de QoS ou de bande passante et qu'il ordonnance les transmissions de ces paires à l'intérieur d'un groupe. Les résultats numériques montrent une augmentation substantielle du débit global et donc de l'efficacité d'utilisation de la bande ISM, et également que le gain obtenu augmente avec le nombre de groupes. Cependant, le nombre de canaux WLAN orthogonaux dans la bande 2,4 GHz limite le nombre de liaisons simultanées.

Dans le même ordre d'idée, Asadi et al. proposent de former des groupes, les membres d'un même groupe devant être à la portée Wi-Fi les uns des autres [11, 12]. L'heuristique proposée pour la formation des groupes utilise la théorie des jeux. Au fur et à mesure que les groupes se forment, seul le membre ayant la meilleure qualité de signal cellulaire communiquera avec la station de base cellulaire ; assurant le rôle de tête de groupe, il devra faire suivre les contenus qu'il aura reçus lui-même en provenance du réseau cellulaire aux autres membres du groupe, mais cette fois en mode D2D. Les auteurs proposent également une stratégie d'ordonnancement WRR (*Weighted Round-Robin*) pour que la station de base décide quel groupe servir. Cette stratégie est analysée et comparée à des stratégies classiques. Les auteurs montrent que leurs propositions mènent à une amélioration du débit du réseau cellulaire (jusqu'à 50%), de

l'efficacité énergétique et assurent un meilleur niveau d'équité que d'autres stratégies d'ordonnancement connues. On remarque néanmoins que, dans le cas de canaux stables, le rôle de tête de groupe sera dévolu de manière durable à un même utilisateur, qui verra donc sa batterie s'épuiser rapidement.

Quant à Cai et al., ils étudient la diffusion de contenus dans un environnement de réseaux sans fil mobiles [18]. Les contenus sont supposés tolérants au délai et les transmissions peuvent se faire en mode hybride, c'est-à-dire soit directement à partir de la station de base, soit de façon opportuniste entre mobiles. Les auteurs proposent un algorithme d'ordonnancement des transmissions qui exploite à la fois la variabilité des conditions de canal (dans le temps et fonction de l'utilisateur) et les contacts aléatoires qui ont lieu entre les utilisateurs mobiles. Dans leur système, la station de base diffuse un contenu donné à plusieurs groupes dans le réseau, en respectant la date de validité associée au contenu. Les utilisateurs se déplacent selon des modèles de mobilité aléatoire et sont capables de communiquer en D2D sur un rayon donné. Lors d'un contact, les utilisateurs présents dans un même rayon de communication vont échanger leurs listes de contenus. Le problème est présenté sous la forme d'une fonction de maximisation du débit. Le modèle analytique et les résultats de simulation mettent en avant la supériorité de leur algorithme, en termes de débit et d'équité par rapport aux algorithmes ne prenant pas simultanément en compte la variabilité du canal et la mobilité des utilisateurs (de 50% à 150% de débit en plus). L'analyse est restreinte au cas où il n'y a qu'une cellule et ne prend pas en compte les interférences qui pourraient être causées par d'autres cellules.

Toujours dans le but d'alléger la charge du réseau cellulaire, d'autres types de solutions sont proposées dans la littérature et, en particulier, la classe d'alternatives à laquelle nous avons déjà fait référence sous le terme de *délestage de données mobiles* [8, 63, 80]. A l'intérieur de cette classe, nous nous concentrons uniquement sur les solutions utilisant simultanément une approche avec tolérance au délai (DTN) [43] et des communications D2D. Pour la plupart, le contenu est transféré à quelques uns des utilisateurs intéressés, ces derniers propageant ensuite l'information au reste des utilisateurs. Nous donnons ici quelques exemples de stratégies, ceux qui nous semblent relativement proches de nos contributions.

Pour mettre en avant l'intérêt d'une approche DTN, Vukadinovic et al. proposent un système de distribution de podcast [87]. Dans les systèmes conventionnels, un contenu de podcast est diffusé par la station de base sur une bande de fréquences sous licence. Là, les auteurs proposent de mettre à profit la connectivité intermittente entre les utilisateurs d'un réseau ad hoc et de réduire l'utilisation de la bande passante cellulaire en ne distribuant le contenu qu'à un nombre bien défini d'utilisateurs. Plus précisément, ils sélectionnent comme cibles de desti-

nation initiales un pourcentage p d'abonnés, lesquels présentent les meilleurs canaux de propagation. Les abonnés restants vont récupérer le contenu au travers de rencontres opportunistes avec des abonnés déjà satisfaits. Les résultats présentés montrent que la distribution opportuniste de certains services de podcast se révèle être une méthode efficace en termes d'économie de ressources et d'efficacité spectrale. Dans le même temps, elle amène une réduction des coûts, puisque l'opérateur n'aura pas à déployer des stations de base supplémentaires. Nous notons que la qualité des résultats dépend du choix des destinataires initiaux – les utilisateurs servis directement par le réseau cellulaire.

D'autres études se sont penchées sur le même type de problème, en cherchant à minimiser le nombre d'utilisateurs servis par le réseau cellulaire et, par conséquent, à maximiser les échanges opportunistes qui n'ont aucun coût pour l'opérateur cellulaire. Han et al. ont ainsi étudié les communications opportunistes dans le cadre de réseaux sociaux mobiles (*Mobile Social Network*, MSN) et ont examiné comment choisir l'ensemble initial avec seulement k utilisateurs, de sorte qu'ils peuvent minimiser la quantité de trafic de données cellulaires [50]. Les auteurs analysent trois algorithmes de sélection : *Random, Heuristic et Greedy*. Les simulations réalisées avec une trace de mobilité montrent un taux de satisfaction des utilisateurs élevé à la fin du processus (c.-à-d. au terme de la durée de validité du contenu) et une amélioration très significative de l'utilisation des ressources cellulaires, puisque le trafic est réduit jusqu'à 73%. Les résultats de la simulation montrent que l'algorithme Greedy fournit les meilleur résultats, suivi par l'algorithme Heuristic et ce, en dépit de l'exigence de Greedy à l'égard d'une connaissance préalable des mouvements d'utilisateurs. Par contre, l'algorithme Heuristic exploite la régularité temporelle et spatiale de la mobilité humaine et l'applique afin de délivrer les contenus dans le futur. Néanmoins, l'algorithme Heuristic crée un problème de confidentialité car les informations de contacts des utilisateurs doivent être téléchargées quotidiennement vers les fournisseurs de services. De même, Barbera et al. proposent un algorithme de sélection de sous-ensembles utilisant les arguments de réseaux sociaux, tels la *centralité* ou *classement des pages (page ranking)* [16]. Les deux méthodes de sélection sont : délégation VIP globale et délégation VIP de voisinage. Alors que le premier se concentre sur les utilisateurs qui sont globalement importants dans le réseau, le second sélectionne les utilisateurs importants dans leurs communautés sociales.

Li et al. étudient également le problème de la sélection du sous-ensemble initial, en le considérant comme un problème de maximisation d'une fonction d'utilité, sous plusieurs contraintes telles que l'hétérogénéité du trafic, la mobilité de l'utilisateur et la mémoire de stockage disponible sur son terminal [65].

D'autres types de solutions ont été également introduits dans la littérature pour le délestage de données [69, 79, 90]. Mayer et al. proposent un nouveau mécanisme de délestage prenant en charge la communication de bout-en-bout entre les utilisateurs mobiles [69]. L'idée sous-jacente ici est de partir sur des communications opportunistes (D2D) et de ne faire intervenir l'infrastructure cellulaire qu'à l'approche de la fin de validité du contenu, lorsque l'échange en D2D n'a pu se faire pour les utilisateurs non encore servis. Whiteback et al., se penchent également sur la dissémination d'un contenu en fonction de sa durée de validité [90]. Dans leur approche *Push-and-Track*, un premier sous-ensemble d'utilisateurs est satisfait par l'infrastructure cellulaire, ces utilisateurs reprenant ensuite le rôle d'émetteurs en faisant suivre le contenu en D2D. Un utilisateur ayant reçu l'information sur une liaison opportuniste envoie un accusé de réception vers l'infrastructure en cellulaire. Cet acquittement permet à la stratégie de suivre le contenu livré et d'évaluer la possibilité de réinjecter des copies. Étant donné que les acquittements sont plus petits que le contenu transmis, la charge sur l'infrastructure est considérablement allégée. De plus, les informations transportées par les accusés de réception peuvent contenir des informations de contrôle qui permettront à l'infrastructure de mettre à jour ses propres informations d'état, dans le but d'identifier le prochain sous-ensemble d'utilisateurs qui devront réinjecter le prochain contenu. Comme pour le travail de Mayer et al. [69], lorsqu'on arrive dans la zone de panique, l'infrastructure prend le relais pour servir les utilisateurs non encore satisfaits. Dans le même genre de solutions, Rebecchi et al. proposent une stratégie de réinjection capable de prendre en considération certaines caractéristiques du réseau opportuniste [79]. Leur stratégie *Droid* estime périodiquement et de façon adaptative où en est le processus de dissémination opportuniste, pour décider si l'infrastructure doit injecter une nouvelle copie du contenu ou non.

Pour récapituler ce que nous avons présenté comme travaux dans la littérature, nous résumons d'une manière générale pour chaque catégorie de stratégies quelques avantages et les inconvénients (Tableau2.1).

Table 2.1 – Tableau comparatif des différents travaux de la littérature

Catégorie	**Avantages**	**Inconvénients**
Déploiement d'infostations	(i) Couverture de zones ciblées ; (ii) faible coût de déploiement ; (iii) réduction du nombre de ressources consommées ; (iv) amélioration de l'expérience des utilisateurs.	(i) Difficulté de déploiement dans des endroits stratégiques ; (ii) problème de passage à l'échelle.
Délestage (Wi-Fi)	(i) Réduction de la charge sur le réseau cellulaire ; (ii) réduction du coût financier global ; (iii) meilleure exploitation de la diversité spatiale ; (iv) meilleure utilisation du spectre.	(i) Plan de déploiement des points d'accès sous-optimal ; (ii) sous-exploitation dans un contexte multi-cellules.
Délestage (Femtocellules)	(i) Exploitation du même spectre cellulaire ; (ii) Mitigation d'une mauvaise couverture cellulaire ; (iii) amélioration de la capacité cellulaire.	(i) nécessité de déployer des stations à courte portée ; (ii) interférences avec les macro-cellules ; (iii) déploiement planifié pour avoir une meilleure QoS ; (iv) déploiement par opérateur.
Coexistence LTE/Wi-Fi	(i) Meilleur usage des ressources cellulaires grâce au partage de la charge avec le Wi-Fi ; (ii) le Wi-Fi diffère les transmissions si besoin.	(i) possibilité d'interférences entre LTE et signaux Wi-Fi ; (ii) méthodes d'accès différentes selon technologie LTE ou Wi-Fi ; (iii) Modification nécessaires des méthodes d'accès et des méthodes de gestion de ressources.

Communications D2D overlay	(i) Pas de possibilité d'interférences et réduction de la complexité ; (ii) contrôle assuré par l'infrastructure cellulaire ; (iii) économie d'énergie ; (iv) amélioration du débit.	(i) Réduction de l'efficacité spectrale.
Communications D2D underlay (avec licence)	(i) Mutualisation des ressources radio ; (ii) amélioration attendue de l'efficacité spectrale ; (iii) contrôle assuré par l'infrastructure cellulaire ; (iv) efficacité énergétique découlant de l'utilisation du D2D.	(i) Impact des interférences sur la QoS ; (ii) complexité des méthodes d'allocation de ressources pour atténuer les interférences.
Communications D2D underlay (sans licence)	(i) Utilisation de différentes technologies ; (ii) absence d'interférences entre les deux technologies ; (iii) établissement de communications simultanées.	(i) Complexité de la gestion des interférences dans la bande sans licence ; (ii) augmentation de la consommation d'énergie ; (iii) accroissement du temps d'accès.

2.4 Positionnement de la thèse vis-à-vis de l'état de l'art

Il existe plusieurs différences entre les travaux précédemment cités et ce travail de thèse. Nous ne traitons pas le problème de la gestion des interférences, puisque dans notre contexte, nous avons une infrastructure cellulaire sous licence et des communications opportunistes utilisant un spectre sans licence. En outre, les critères de QoS considérés sont différents, puisque nous nous focalisons sur les contraintes temporelles. En particulier, notre objectif est de respecter la durée de vie d'un contenu, au terme de laquelle le contenu doit absolument avoir été délivré à l'ensemble (100%) de ses utilisateurs. Enfin, nous faisons face à des obstacles différents, car dans notre cas, la station de base (ou le nœud eNB) doit disposer d'informations sur les statuts des canaux de communication, qu'il s'agisse de communications entre utilisateurs D2D ou cellulaires, et sur les statuts des utilisateurs, connectés ou non à la station de base au moment de l'envoi du contenu. Or, ces informations ne sont pas toujours facilement accessibles

pour la station de base.

Tous les travaux que nous avons pu relever dans la littérature, et en particulier ceux concernant les communications D2D *underlay*, mettent en avant l'intérêt multiple à utiliser ces dernières : elles peuvent non seulement contribuer à augmenter le débit global et décharger le réseau cellulaire, mais aussi offrir à leurs propres utilisateurs une procédure allégée d'établissement de session et de meilleurs temps de réponse et débits, sans pour autant trop interférer sur les communications cellulaires. C'est pour cela que nous considérons ce type d'approche dans notre système de délestage de trafic. Les travaux que nous venons de citer ont pour même but de vouloir diminuer la charge du réseau cellulaire en introduisant des communications D2D overlay. Mais nous relevons néanmoins des spécificités de notre contexte. Tout d'abord, nous considérons qu'un contenu a une durée de vie propre qui lui est associée, et qu'il doit absolument être délivré avant que cette échéance ne soit atteinte. Ensuite, nous garantissons une qualité de service, dans la mesure où le contenu est délivré à chacun de ses destinataires (les utilisateurs intéressés).

Soit A un succès dans la vie. Alors A = x + y + z, où x = travailler, y = s'amuser, z = se taire.

Albert Einstein

3

SCoD : Livraison de contenu planifiée en environnement purement cellulaire

3.1 Introduction

L'augmentation sans fin de la capacité des réseaux cellulaires n'est pas une solution définitive au problème des masses de données auquel sont confrontés aujourd'hui les opérateurs cellulaires [29]. Actuellement, la plupart des opérateurs mettent en œuvre une stratégie d'affectation des ressources en deux phases, dans laquelle les clients voient leur bande passante limitée après un certain seuil de consommation de données (généralement, de l'ordre de quelques Gbits) [66, 92]. Des alternatives récemment proposées consistent à délester le trafic mobile vers des technologies complémentaires (telles que le Wi-Fi), ou à émettre dans le spectre sous licence avec des Femtocellules [80], ou encore à mettre en place des communications *device-to-device* [13].

Dans ce chapitre, nous présentons une toute autre approche pour résoudre le problème de la dissémination de données dans un réseau cellulaire chargé : nous profitons de la tolérance au délai inhérente à certains types de contenus afin de diminuer le nombre et le coût des transmis-

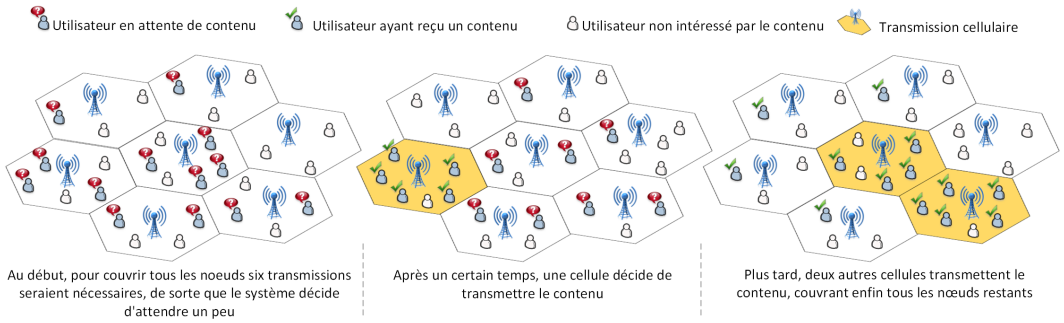


Figure 3.1 – Principe de fonctionnement d'une dissémination de contenu en environnement purement cellulaire.

sions cellulaires. Nous considérons un groupe d'utilisateurs mobiles intéressés par un contenu spécifique, groupe que nous appelons *population intéressée*. Le contenu demandé est supposé d'intérêt général, et doit être délivré à ses destinataires dans un délai borné et connu. Les utilisateurs sont dispersés dans une zone géographique desservie par un ensemble d'antennes.

Si l'on s'en remet au fonctionnement normal d'un réseau cellulaire, la demande du contenu par un seul utilisateur suffit pour déclencher instantanément une transmission. En conséquence, le réseau va procéder à autant de transmissions qu'il y a de cellules hébergeant au moins un utilisateur intéressé. Or, le fait de retarder légèrement une transmission contribue à augmenter les chances de satisfaire en même temps d'autres utilisateurs demandeurs situés dans la même cellule. Nous proposons donc SCoD (*Scheduled Content Delivery*), une stratégie de dissémination de contenu tolérant au délai dans les environnements mobiles qui permet l'économie de ressources radio. SCoD s'appuie sur la mobilité des utilisateurs pour programmer une transmission multicast dès lors qu'une cellule héberge un nombre suffisant d'utilisateurs intéressés. Le problème sous-jacent est alors de déterminer ce seuil, afin de réduire le nombre de points d'accès qui vont devoir émettre et le nombre total de transmissions nécessaires pour satisfaire tous les utilisateurs intéressés par le contenu considéré.

La figure 3.1 illustre le fonctionnement de SCoD. Dans le scénario considéré, les utilisateurs [1] intéressés par le contenu occupent six cellules au moment où ce dernier est disponible. Dans un système conventionnel, vu la position des utilisateurs, six transmissions sont nécessaires, puisque les utilisateurs intéressés sont positionnés dans six cellules. Avec SCoD, le système décide de différer les transmissions et suit la trace des utilisateurs : une transmission est déclenchée dans une cellule lorsqu'elle rassemble un nombre suffisant d'utilisateurs. Les différentes transmissions sont ainsi étalées dans le temps et sont réalisées aux instants où les utili-

1. Nous utilisons les termes "utilisateur", "nœud" ou "terminal" pour désigner tout destinataire d'un contenu.

sateurs sont, si possible, le plus regroupés. Dans le scénario de la figure 3.1, trois transmissions suffisent (au lieu de six).

Au cœur de SCoD se trouve une fonction de décision, dont le rôle est de déterminer si le nombre d'utilisateurs intéressés présents dans une cellule est suffisant ou non pour déclencher une transmission. Cette fonction de décision dépend nécessairement du délai maximum que tolèrent les utilisateurs pour la réception du contenu, et au terme duquel ils vont considérer la transmission en échec. Nous dressons ici les grandes lignes du comportement typique de la fonction de décision :

— Au début d'une période (la période commence à l'instant où le contenu est rendu disponible sur les points d'accès et elle s'achève avec le délai maximum toléré par les utilisateurs), SCoD recherche les grands rassemblements d'utilisateurs, ou en d'autres termes, les cellules qui concentrent de nombreux utilisateurs ;

— Au fur et à mesure que le temps s'écoule, SCoD devient de plus en plus pragmatique en relâchant de plus en plus son exigence sur la taille des rassemblements d'utilisateurs ;

— La taille du rassemblement requise pour déclencher une transmission va ainsi de plus en plus décroître jusqu'à ce que la fin de la période soit atteinte.

SCoD va ainsi alléger la charge sur l'infrastructure réseau, d'une façon similaire à ce qui est fait dans le *Cloud Radio Access Network* (C-RAN) [25]. En mettant en commun la coordination des ressources, le C-RAN améliore l'efficacité des stations de base lorsque le trafic est non uniforme dans les différentes cellules. De son côté, SCoD réduit les coûts pour l'opérateur en étalant les transmissions dans le temps et en réduisant leur nombre : le réseau économise ainsi des ressources qu'il peut attribuer à d'autres services.

L'efficacité réelle de SCoD dépend clairement de la précision de la fonction de décision et de la prévision des mouvements des utilisateurs entre les cellules. Nous évaluons les performances de SCoD en utilisant plusieurs heuristiques et en confrontant leurs résultats à ceux de deux autres alternatives, considérées comme des stratégies de référence. La première de ces stratégies est de type "transmettre dès que disponible" ; tel que décrite pour la figure 3.1, elle consiste, pour chaque point d'accès servant au moins un utilisateur intéressé, à diffuser le contenu au début de la période. De façon évidente, elle va donner lieu à autant de transmissions qu'il y a de points d'accès occupés. Dans la suite de ce chapitre, nous nous référerons à cette stratégie sous le nom de *Hot Potato*. La seconde stratégie de référence opère sous l'hypothèse d'un *Oracle*, qui connaît d'avance les trajectoires des utilisateurs intéressés. Parce qu'elle connaît les futurs déplacements des nœuds, cette stratégie permet de calculer le nombre minimum de transmis-

sions permettant de couvrir l'ensemble de la population intéressée. Bien que la stratégie Oracle ne soit pas applicable dans des contextes pratiques car non-causale (une stratégies irréalisable en réalité), elle nous fournit une borne supérieure quant aux attentes que nous pouvons avoir sur les performances de SCoD.

L'objet de ce chapitre est de présenter nos résultats de performance, obtenus par simulation :

— Les stratégies de référence Hot Potato et Oracle, ainsi que notre stratégie SCoD, ont été appliquées à des environnements correspondant à la ville de Bologne (Italie) et à la ville de Luxembourg.

— Nous créons différents plans de déploiement pour le réseau cellulaire couvrant chaque ville : chaque plan résulte de la variation de la taille des cellules et, par conséquent, du nombre total d'antennes nécessaires pour couvrir la ville considérée.

— Nous exploitons des traces de mobilité réelle de véhicules, obtenues sur le réseau routier de Bologne [61] et sur celui de Luxembourg [33].

— Nous envisageons plusieurs scénarios de dissémination en fonction du délai maximum que sont prêts à subir les utilisateurs pour recevoir le contenu (échéance).

Nous constatons que la stratégie SCoD réduit de façon significative le nombre de transmissions par rapport à la stratégie Hot Potato, et de plus, qu'elle conduit à des résultats proches de ceux de l'Oracle.

3.2 Modèle de système, hypothèses et illustration du problème

Nous supposons que le contenu doit être diffusé sous un système de publication-souscription (pub/sub) où les utilisateurs s'abonnent à un contenu. L'une des principales caractéristiques d'un tel système est que le nombre d'abonnés est connu à l'avance. De plus, la remise des contenus générés par ses applications peut être différée (le trafic n'est pas temps réel). Nous nous concentrons sur le contenu qui doit être livré à un grand nombre d'abonnés. Des exemples de tels contenus incluent les éditions numériques des journaux, les prévisions météorologiques, des événements liés à une ville, ou les contenus populaires que l'on trouve sur les réseaux sociaux. En conséquence, au lieu de pousser le contenu une fois disponible, nous cherchons à mesurer combien de temps un contenu peut être retardé afin de maximiser le nombre d'abonnés au sein d'une même cellule (et donc réduire le nombre de transmissions).

Nous définissons notre système par le tuple $[c, \mathbf{N}^c, \mathbf{A}, t_0^c, \tau]$, où c est le contenu à transmettre, $\mathbf{N}^c$ l'ensemble des utilisateurs intéressés à recevoir c, $\mathbf{A}$ l'ensemble des points d'accès

(par abus de langage, l'ensemble des cellules), t_0^c l'instant où c devient disponible, et τ le délai maximum que les utilisateurs tolèrent pour recevoir c. Nous nous référerons au temps $t_0^c + \tau$ comme le deadline du contenu c.

Le contenu c reste valide pendant l'intervalle $[t_0^c, t_0^c + \tau]$. Pour des raisons de clarté, nous considérons que les périodes de dissémination ne se chevauchent pas, ce qui signifie qu'il y a un seul et unique contenu présent dans le système à la fois. Cette hypothèse n'influence pas notre conclusion, car nous évaluons le système en fonction du contenu [2]. En outre, pour assurer la continuité de notre service de diffusion, nous considérons qu'une fois un contenu expiré, un autre est aussitôt généré.

Sauf indication contraire, nous ne faisons aucune hypothèse quant à la technologie radio utilisée et nous allons simplement nous référer à $a_i \in \mathbf{A}$ comme point d'accès sans fil (ou station de base) couvrant une zone géographique appelée *cellule*. Nous supposons que les cellules ne se chevauchent pas et qu'elles couvrent l'ensemble des zones cibles (c.-à-d. cellules juxtaposées). Cela signifie que tout utilisateur est associé à un point d'accès unique à tout moment. Nous considérons que le système effectue rapidement les transferts intercellulaires (*soft hand-offs*) tout au long du déplacement d'un utilisateur [57]. Nous supposons que l'infrastructure cellulaire est capable de procéder à des envois en *multicast*. En d'autres termes, une seule transmission est suffisante pour *pousser* le contenu vers tous les nœuds intéressés d'une même cellule. Ainsi, dans le reste de ce chapitre, chaque fois que nous utiliserons le terme *transmission*, nous entendrons par là une transmission en multicast.

Nous supposons enfin que la population $\mathbf{N}^c$ reste inchangée pendant une durée τ. Les nouveaux arrivants ne sont pas pris en compte dans le processus en cours. Les nœuds dans $\mathbf{N}^c$ appartiennent à l'un des deux sous-ensembles suivants :

— $\mathbf{N}_{a_i}^{c\oplus}(t)$, l'ensemble des utilisateurs associés au point d'accès a_i à l'instant t qui n'a pas encore reçu le contenu ;

— $\mathbf{N}_{a_i}^{c\ominus}(t)$, l'ensemble des utilisateurs associés au point d'accès a_i à l'instant t qui a déjà reçu le contenu.

Par conséquent :

$$\mathbf{N}^c = \bigcup_{i=1}^{|\mathbf{A}|} [\mathbf{N}_{a_i}^{c\oplus}(t) \bigcup \mathbf{N}_{a_i}^{c\ominus}(t)], \forall t \in [t_0^c, t_0^c + \tau] \tag{3.1}$$

2. En fait, nous obtenons une limite inférieure sur les gains que nous pouvons obtenir avec SCoD, car plusieurs contenus pourraient bénéficier du même rassemblements d'utilisateurs, conduisant à des gains additifs. Le traitement de plusieurs contenus dans le système en même temps fera l'objet de travaux futurs.

Nous définissons une *configuration* à un instant t comme étant une capture instantanée du système. Elle consiste à définir les utilisateurs associés à chaque point d'accès à l'instant t. Une configuration change chaque fois qu'au moins un utilisateur se déplace vers une autre cellule. Une seule transmission par un point d'accès est reçue par tous les utilisateurs de la cellule correspondante. Chaque point d'accès connaît à tout moment la liste des utilisateurs mobiles qu'il couvre.

Objectif. Notre objectif est de minimiser le nombre de transmissions en tenant compte de la mobilité des utilisateurs, d'une part, et de la tolérance au délai du contenu, d'autre part. En d'autres termes, étant donné le tuple $[c, \mathbf{N}^c, \mathbf{A}, t_0^c, \tau]$, nous devons trouver l'ensemble minimum de paires (t_i, a_i) qui couvre tous les nœuds intéressés à recevoir c, où $t_i \in [t_0^c, t_0^c + \tau]$ correspond à l'instant où le point d'accès a_i transmet une copie de c.

3.3 Jusqu'où la tolérance au délai peut-elle aider à économiser les ressources radio ?

Dans cette section, nous répondons à la question suivante : *est-il vraiment profitable de retarder la transmission d'un contenu ?*

3.3.1 Les Stratégies de transmission

Pour évaluer les avantages des transmissions retardées, nous comparons la stratégie Hot Potato qui consiste à pousser un contenu dès que disponible à la stratégie Oracle qui s'appuie sur la connaissance des mouvements futurs des utilisateurs, et ce afin de déterminer les meilleurs moments pour transmettre. Notons que la stratégie Oracle est un système non-causal que nous utilisons uniquement à des fins de comparaison.

Hot Potato. Cette stratégie transmet un contenu dès qu'il est disponible. En d'autres termes, dès qu'un contenu est disponible (à t_0^c), chaque point d'accès vérifie la liste des utilisateurs intéressés sous son contrôle. Un point d'accès déclenche une transmission si au moins un utilisateur intéressé est présent dans sa cellule. En utilisant la notation que nous avons introduite dans la section 3.2, le point d'accès $a_i \in \mathbf{A}$ transmet une copie du contenu c à $t = t_0^c$ seulement si $\mathbf{N}_{a_i}^{c\oplus}(t_0) \neq \emptyset$ (c'est-à-dire seulement si au moins un abonné est dans la cellule a_i). En conséquence, le nombre de transmissions est égal au nombre de cellules où les abonnés se trouvent à t_0^c. La stratégie Hot Potato est agnostique à la mobilité des utilisateurs et à leur tolérance à la réception retardée d'un contenu.

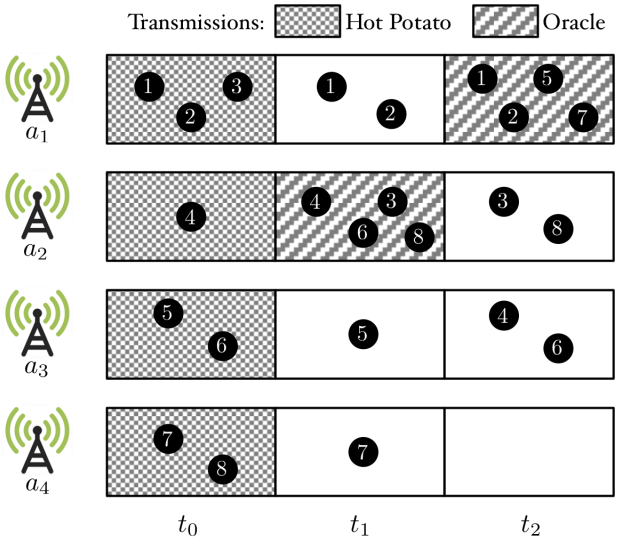


Figure 3.2 – Les stratégies Hot Potato et Oracle.

Oracle. Contrairement à la stratégie Hot Potato, l'Oracle repose sur la connaissance complète, *a priori*, de la mobilité des utilisateurs pour toute la période τ. En bref, on définit un graphe orienté et biparti où les sommets sont composés d'utilisateurs et de points d'accès. Les arêtes illustrent l'association d'un utilisateur avec un point d'accès. Un arête entre un point d'accès et un nœud utilisateur signifie que le point d'accès a hébergé l'utilisateur au moins une fois pendant la durée de vie du contenu (deadline τ). L'Oracle calcule alors un ensemble de couverture minimum qui correspond au nombre minimum de transmissions qui permet au contenu d'atteindre tous les utilisateurs au moins une fois. En conséquence, cette stratégie s'appuie sur ces connaissances pour déterminer précisément quels points d'accès doivent transmettre et à quels moments afin de maximiser les avantages de chaque transmission. La stratégie Oracle exploite pleinement la tolérance des abonnés à recevoir un contenu avec un léger retard, afin de tirer le meilleur parti de chaque transmission en termes d'abonnés satisfaits.

Trouver les meilleurs moments pour transmettre un contenu revient à résoudre le problème bien connu du *minimum set cover* [55] qui est également jugé comme NP-complet. Nous utilisons alors une heuristique (*greedy*) qui va nous fournir une approximation acceptable de la meilleure solution en un temps raisonnable [28, 85].

La figure 3.2 montre un exemple simple, dans lequel nous considérons une population de $|\mathbf{N}^c|= 8$ utilisateurs et $|\mathbf{A}|= 4$ points d'accès. Les huit utilisateurs sont intéressés à recevoir un contenu c mis à disposition à $t = t_0$. La date limite pour la transmission de c est fixée à $t_0 + \tau$.

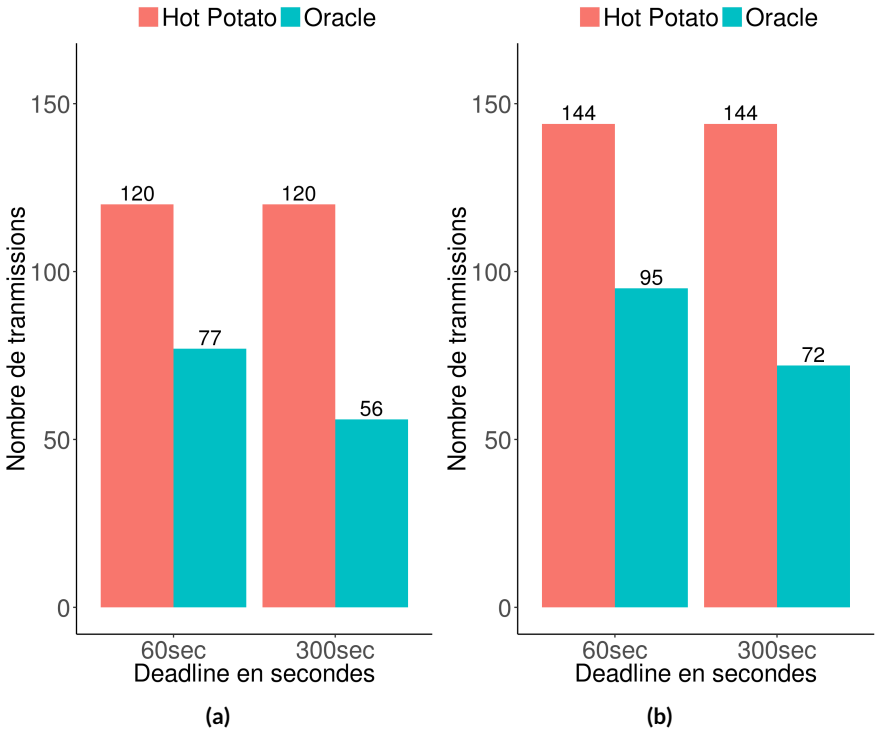


Figure 3.3 – Hot Potato et Oracle : Le nombre de transmissions réalisées en fonction de la tolérance au délai et de la population (a) $|\mathbf{N}^c| = 500$ (b) $|\mathbf{N}^c| = 1000$.

Nous discrétisons le temps en divisant l'intervalle $[t_0, t_0 + \tau]$ en trois sous-intervalles définissant quatre instants t_i avec $i \in [0,\ 3]$. L'instant t_i reflète une nouvelle *configuration* du réseau. Nous considèrerons, sans perte de généralité, que tous les sous-intervalles durent une unité de temps.

Hot Potato transmet le contenu dès qu'il est disponible, c'est-à-dire à t_0, et il y a autant de copies du continu que le nombre de cellules hébergeant des abonnés à $t = t_0$. Dans ce scénario, quatre transmissions sont nécessaires pour couvrir tous les abonnés, chacune des quatre cellules a_1, a_2, a_3, et a_4 hébergeant au moins un abonné à t_0. De son côté, l'Oracle utilise la distribution de contenu différée pour déterminer l'ensemble minimal de transmissions. Ici, l'Oracle ne nécessite que deux transmissions (par les cellules a_1 et a_2 aux instants t_2 et t_1, respectivement).

3.3.2 Évaluation avec des traces de mobilité réelle

Dans cette section, nous évaluons combien de ressources radio nous pouvons économiser à l'échelle de la ville de Bologne en Italie. Nous profitons de traces de mobilité véhiculaire disponibles pour Bologne qui décrivent les mouvements de plus de 18 000 véhicules [61]. Nous considérons chaque véhicule comme un utilisateur mobile demandant un contenu distribué selon deux échéances différentes, τ= 60 s and τ= 300 s. Nous créons un déploiement réaliste (déploiement de points d'accès avec un processus poissonnien) d'un réseau cellulaire composé de $|\mathbf{A}| = 300$ points d'accès couvrant Bologne. Nous considérons deux tailles pour la population des nœuds intéressés par un contenu, présents dans le système jusqu'à l'expiration du deadline du contenu : $|\mathbf{N}^c|$ = 500 et $|\mathbf{N}^c|$ = 1 000.

Les résultats sont présentés dans la figure 3.3. Nous évaluons le nombre de transmissions pour les deux stratégies au cours de la période τ. Nous observons, comme prévu, que l'Oracle nécessite moins de transmissions pour atteindre tous les utilisateurs que la stratégie Hot Potato. Nous notons également que le nombre de transmissions diminue lorsque nous augmentons le délai de validité du contenu. Ce comportement s'explique par le fait que certaines cellules deviennent plus encombrées à certains moments de la trace, comme une conséquence naturelle de la mobilité des utilisateurs. [3]

En résumé, nous concluons qu'il est possible de réduire significativement le nombre de transmissions en utilisant la mobilité des nœuds et leur tolérance à recevoir des contenus retardés. La question maintenant est de savoir comment s'approcher le plus possible des performances de l'Oracle, sachant que cette stratégie est un système non causal. Dans ce qui suit, nous décrivons SCoD, notre proposition qui repose sur une fonction de décision utilisant une heuristique. Cette fonction est à la fois simple et très puissante pour prédire les rassemblements de nœuds à travers le réseau, et modulée par le temps restant avant le deadline.

3.4 SCoD : Livraison de contenu planifiée

Nous proposons SCoD (*Scheduled Content Delivery*), une stratégie qui décide quand déclencher une transmission. SCoD base sa décision sur le nombre d'abonnés attendant toujours le contenu c ainsi que sur le temps restant avant le deadline τ. Dans le cas où le deadline est atteint, la solution est triviale – une transmission multicast est déclenchée dans chaque cellule hébergeant au moins un abonné qui n'est pas encore satisfait. Dans le cas où il reste du temps, SCoD

3. En fait, nous rencontrons ce type de situation dans plusieurs scénarios de la vie réelle, comme le trafic sur le réseau routier ou dans les transports en commun pendant les heures de pointe.

vérifie si le nombre d'utilisateurs non encore satisfaits dans une cellule est assez grand pour déclencher une transmission multicast dans cette dernière. La question est de savoir comment déterminer ce qui est «assez grand». Ceci fait l'objet de la section suivante.

3.4.1 Déclenchement des transmissions

D'un point de vue formel, le point d'accès a_i transmet le contenu c à l'instant t si l'une des deux conditions ci-dessous est vérifiée :

$$\begin{cases} |\mathbf{N}_{a_i}^{c\oplus}(t)| \geq \omega\beta(t) \sum_{j=1}^{|A|} |\mathbf{N}_{a_j}^{c\oplus}(t)|, & \forall t \in [t_0^c, t_0^c + \tau[, \\ |\mathbf{N}_{a_i}^{c\oplus}(t)| > 0, & t = t_0^c + \tau, \end{cases} \tag{3.2}$$

où $|\mathbf{N}_{a_i}^{c\oplus}(t)|$, comme précédemment mentionné, est le nombre de nœuds intéressés à recevoir le contenu c et rassemblés sous la couverture du point d'accès a_i à l'instant t. La fonction $\beta(t)$, appelée fonction de retardement (*lateness function*), définit la tolérance de SCoD concernant la taille d'un rassemblement. L'idée principale derrière cette fonction est d'éviter les transmissions à faible utilité au début de l'intervalle de diffusion. Comme le temps passe, SCoD devient de moins en moins conservateur. Cela signifie que SCoD est de moins en moins exigeant sur le nombre requis d'utilisateurs non satisfaits dans une cellule pour déclencher une transmission. Pour assurer ce comportement, nous définissons $\beta(t)$ comme une fonction décroissante de façon monotone et normalisée sur l'intervalle $[0, 1]$, c.-à-d., $0 \leq \beta(t) \leq 1, \forall t \in [t_0, t_0 + \tau]$.

Comme on le voit dans l'équation 3.2, $\beta(t)$ est associée à un paramètre ω qui détermine *l'agressivité* du processus décisionnel au début de l'intervalle. En fait, $\omega \in [0, 1]$ est un coefficient de pondération qui reflète le niveau de besoin de la diffusion. Si $\omega = 1$, cela signifie que le contenu c n'est transmis que si, au début de la période de diffusion, 100% des nœuds sont rassemblés sous un seul point d'accès (trop agressif). De même, $\omega = 0.5$ signifie qu'au début de la dissémination, il faut au moins 50% de la population $\mathbf{N}^c$ rassemblés sous le même point d'accès pour déclencher une transmission.

Dans la figure 3.4, nous illustrons le fonctionnement de SCoD, en précisant les conditions pour que SCoD déclenche des transmissions. La ligne en gras est l'enveloppe fournie par SCoD tandis que la ligne en pointillés donne le nombre de nœuds mobiles associés au point d'accès a_i et qui n'ont pas encore reçu le contenu. Une transmission est effectuée lorsque les deux courbes se rencontrent.

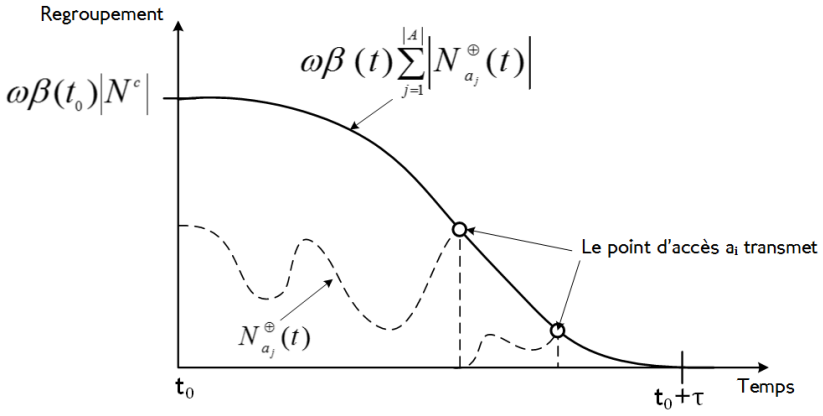


Figure 3.4 – Fonctionnement de SCoD.

3.4.2 Déterminer ω et $\beta(t)$

L'efficacité de SCoD dépend du bon choix des deux paramètres, le coefficient ω et la fonction de retardement $\beta(t)$. Le paramètre ω donne l'amplitude de début de la fonction de retardement, comme illustré dans la figure 3.4. Un compromis doit être pris en compte lors de la définition de la valeur de ω. Si ω est trop proche de 1, nous ne pouvons jamais déclencher une transmission si les nœuds ne tendent pas à se rassembler sous le même point d'accès. Si ω est trop petit, SCoD fonctionnera comme la solution Hot Potato et aucune amélioration ne sera observée. Trouver un bon compromis est la clé pour un comportement performant de SCoD.

En pratique, comme la mobilité des nœuds ne peut être connue à l'avance, nous devons nous appuyer sur des observations antérieures pour définir les paramètres de SCoD. À ce stade, nous recueillons pour chaque intervalle de durée τ le nombre maximal d'utilisateurs réunis au même instant sous le même point d'accès (nous appelons cela le *taux de regroupement*). Nous utilisons l'observation la plus récente de ce nombre pour définir la valeur de ω. Étant donné que chaque intervalle concerne un contenu spécifique, le groupe de nœuds diffère d'un intervalle à l'autre. Nous divisons le nombre total d'utilisateurs par le nombre d'utilisateurs intéressés afin d'avoir une valeur normalisée. Nous pourrions appliquer d'autres approches de prédiction, mais nous avons remarqué que la considération de l'intervalle précédent suffit pour obtenir de bons résultats, comme nous le verrons dans la section 3.6.

À titre d'illustration, nous calculons ce rapport en utilisant la trace de Bologne avec un nombre variable de points d'accès (100, 200, 300, 400 et 500). Nous supposons que $\tau = 300$s. Les résultats sont présentés dans la figure 3.5. Ce chiffre représente les valeurs du rapport de rassemblement pour des intervalles consécutifs de longueur τ. Bien que la variation du rapport

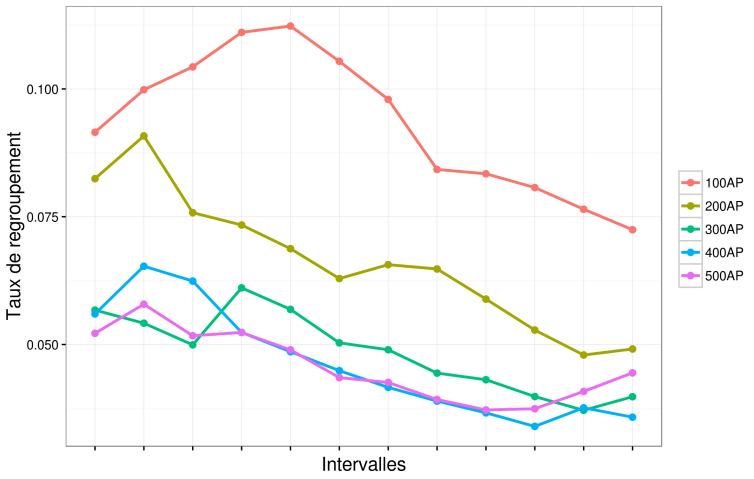


Figure 3.5 – Variation du taux de regroupement pour des intervalles consécutifs de durée τ = 300 s.

soit erratique, les variations du rapport d'un intervalle à l'autre ne sont pas significatives dans la plupart des cas. Cela est encore plus vrai pour les valeurs plus grandes de τ. Ces résultats indiquent que nous n'avons pas besoin de regarder plus loin dans le passé ; observer l'intervalle précédent semble suffire pour avoir une idée de la façon dont la population est susceptible de se rassembler dans l'intervalle suivant.

Table 3.1 – Les fonctions de retardement utilisées pour $\beta(t)$.

Fonction	**Expression**
Linear (linéaire)	$\beta(t) = \frac{-(t-t_0)}{\tau} + 1$
Square-root (racine)	$\beta(t) = \sqrt{\frac{-(t-t_0)+\tau}{(t-t_0)+\tau}}$
Power (puissance)	$\beta(t) = (\frac{-(t-t_0)}{\tau} + 1)^3$

Pour la fonction de retardement, nous adoptons une approche similaire. Nous testons plusieurs fonctions sur une observation récente sur une fenêtre de taille τ et réutilisons la même fonction (avec le ω qui correspond) pendant l'intervalle courant. Dans la pratique, nous considérons trois fonctions monotonement décroissantes différentes dans nos tests. Elles sont décrites dans le tableau 3.1

Afin d'évaluer le SCoD, nous utilisons deux ensembles de données véhiculaires réalistes générés avec le simulateur de trafic SUMO [60]. Le premier représente le trafic de véhicules dans la ville de Luxembourg [33]. Il couvre une surface de 155 km^2 et 931 km de routes. La mobilité s'appuie sur les statistiques publiées par le STATEC (Luxembourg National Institute of Statistic and Economic studies) [2]. La population compte environ 140 000 véhicules et la durée est de 24 heures. Le deuxième ensemble correspond au déplacement de plus de 18 000 véhicules dans la ville de Bologne (le même que nous avons utilisé dans la section 3.3) [61]. La zone est d'environ 10 km^2, traversée par plus de 191 km de routes [90] sur une période d'une heure.

A des fins d'évaluation, et afin de comparer équitablement les résultats, nous avons réduit la taille de la zone cible dans la trace de Luxembourg. Nous avons choisi la zone la plus fréquentée sur une superficie de 10 km^2, ce qui implique une population de 16 000 véhicules. Nous suivons l'emplacement des véhicules pendant une durée d'une heure dans les deux ensembles de données.

Les traces décrivent les mouvements des véhicules mais ne fournissent aucune information sur la connectivité que ces véhicules peuvent avoir avec l'Internet. Nous émulons un réseau cellulaire en utilisant un diagramme de Voronoi pour chacune des deux villes et où chaque cellule de Voronoi contient seulement un point d'accès [14]. Cela nous donne une certaine flexibilité pour évaluer notre proposition dans des conditions de densité multiples et des changements de cellules fréquents. À titre d'illustration, la figure 3.6 montre la tessellation de Voronoi de 100 points d'accès dans la zone centrale de Bologne.

Nous comparons SCoD avec les stratégies présentées dans la section 3.3 : Oracle (O) et Hot Potato (HP). Nous définissons une stratégie de base que nous appelons *couverture totale* (Full Coverage, FC), où l'infrastructure distribue systématiquement le contenu à chaque cellule ayant reçu au moins un utilisateur intéressé pendant l'intervalle $[t_0^c, t_0^c + \tau]$.

Pour donner plus de confiance à nos résultats, nous simulons de multiples plans de déploiement suivant une distribution uniforme. Nous générons 10 topologies différentes pour chaque nombre de points d'accès $|\mathbf{A}| \in \{100, 200, 300, 400, 500\}$, ce qui fait 50 scénarios au total. En ce qui concerne le temps de validité (deadline) du contenu, nous considérons plusieurs scénarios avec les valeurs suivantes $\tau \in \{60\text{ s}, 120\text{ s}, 180\text{ s}, 300\text{ s}\}$. Les résultats que nous présenterons dans la section suivante sont la moyenne de 10 simulations effectuées pour chaque topologie.

Comparaison. Nous considérons le *gain* $\mathcal{G}$ pour comparer les solutions. Ce paramètre mesure les économies réalisées par une solution donnée par rapport à la stratégie de couverture totale

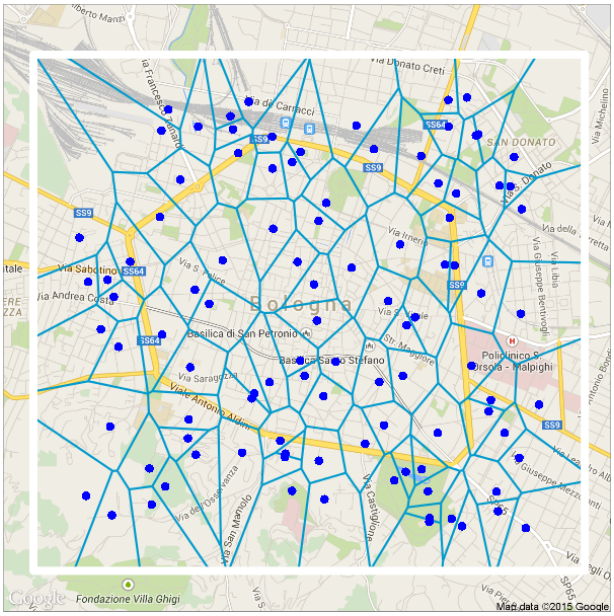


Figure 3.6 – Exemple d'un diagramme de Voronoi avec 100 cellules déployées dans la ville de Bologne.

(FC). Il est formellement défini comme :

$$\mathcal{G}_{\text{strategie}} = \frac{N_{\text{FC}} - N_{\text{strategie}}}{N_{\text{FC}}}, \tag{3.3}$$

où N_{FC} est le nombre de transmissions lorsque la stratégie FC est utilisée et $N_{\text{strategie}}$ est le nombre de transmissions pour la "stratégie" considérée.

Real run vs ideal run. Nous comparons également le nombre moyen de transmissions pour notre stratégie par rapport à son cas d'exécution idéal, c'est-à-dire si les valeurs les plus appropriées de $\beta(t)$ et ω étaient connues à l'avance. Il montre jusqu'à quel point la stratégie est proche de son propre meilleur cas possible.

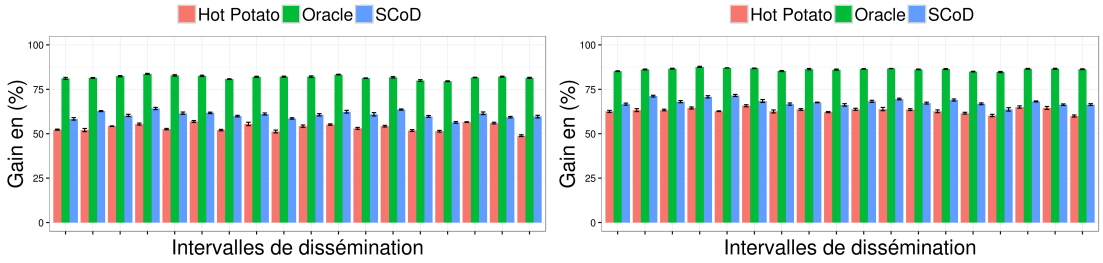
La fonction de retardement $\beta(t)$ et la fréquence d'utilisation. Nous considérons les trois fonctions présentées dans le tableau 3.1 pour la fonction de retardement $\beta(t)$. Par conséquence, avant de commencer la transmission de chaque contenu, notre stratégie choisit quelle fonction sera utilisée au cours de l'intervalle suivant. Nous mesurons le pourcentage de l'utilisation des différentes fonctions afin d'avoir une estimation statistique de la façon dont le réseau se comporte. Faisons varier plusieurs paramètres du réseau pour montrer s'ils ont un impact sur les résultats ou non. Si une fonction était amenée à apparaître clairement comme meilleure, elle pourrait par exemple être favorisée dans les futurs déploiements de réseau.

3.6 Résultats d'évaluation

Pour évaluer les avantages de SCoD, nous la comparons avec les autres stratégies décrites dans la section précédente. Nous verrons que notre proposition fonctionne bien pour économiser les ressources radio en diminuant le nombre de transmissions. En outre, nous soulignons l'efficacité de notre système en termes d'adaptabilité à l'arrivée d'un nouveau contenu.

3.6.1 Dataset de Luxembourg

Gain moyen

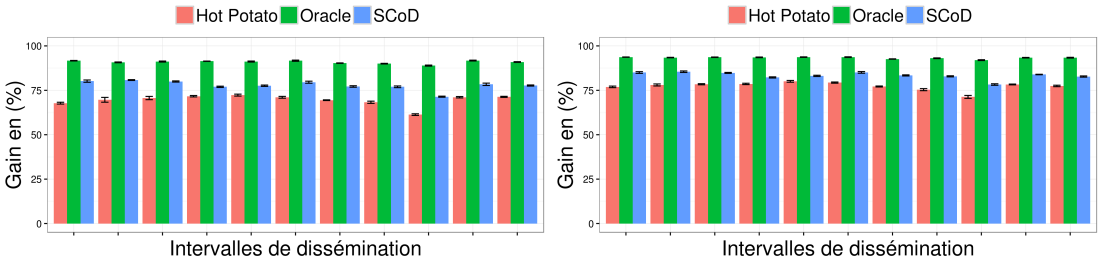


(a) $|\mathbf{A}|$= 300 AP, τ = 180 s **(b)** $|\mathbf{A}|$= 500 AP, τ = 180 s

Figure 3.7 – Le gain moyen de transmission pour un deadline $\tau = 180\ s$ à Luxembourg : SCoD vs les stratégies de base.

Nous évaluons d'abord le gain moyen de transmission pour la ville de Luxembourg, selon la définition de l'équation 3.3. Chaque point sur l'axe des abscisses représente un intervalle de durée τ – cela signifie que chaque barre de la figure correspond à la diffusion d'un contenu. L'axe des ordonnées donne le gain $\mathcal{G}_{\text{strategie}}$ par stratégie. Les graphes montrent les résultats pour deux valeurs du nombre de points d'accès ($|\mathbf{A}|$={300,500}) et deux valeurs du deadline (τ={180, 300}). Des intervalles de confiance de 95% sont indiqués. Il est important de noter que lorsque le délai augmente, la population moyenne diminue, car nous ne considérons que les utilisateurs qui restent dans le système pendant l'intégralité de l'intervalle de transmission.

Dans la figure 3.7, nous illustrons l'effet d'un petit délai sur les trois stratégies, avec deux valeurs différentes du nombre de points d'accès. Augmenter le nombre de points d'accès revient à diminuer la taille des cellules. D'une part, les rassemblements ont tendance à être plus petits (plus petits groupes par cellule). D'autre part, les utilisateurs sont plus exposés à des transferts intercellulaires fréquents (hand-offs), de sorte qu'il y a probablement plus de configurations pendant la période de transmission. Le comportement des stratégies de diffusion dépend



(a) $|\mathbf{A}|$ = 300 AP, τ = 300 s **(b)** $|\mathbf{A}|$ = 500 AP, τ = 300 s

Figure 3.8 – Le gain moyen de transmission pour un deadline $\tau = 300\ s$ à Luxembourg : SCoD vs les stratégies de base.

alors d'une combinaison de la taille des cellules et de la mobilité des nœuds. Nous observons que SCoD conduit à des résultats légèrement meilleurs lorsque nous augmentons le nombre de points d'accès. Une des raisons de ce phénomène est que la période de transmission n'est pas suffisamment longue pour donner aux utilisateurs la flexibilité de fréquenter différents rassemblements.

Nous montrons des résultats pour un deadline plus long dans la figure 3.8. Nous observons à nouveau que SCoD dépasse la stratégie Hot Potato dans tous les cas. Ici, SCoD est encore plus proche d'Oracle. La raison de ce comportement est que les utilisateurs disposent de plus de temps pour se déplacer et pour participer à la dynamique de formation de groupes.

Real run vs ideal run

Nous évaluons dans cette section l'adaptabilité de SCoD en comparant plusieurs configurations de points d'accès et de deadline. Nous limitons notre analyse uniquement au cas de SCoD. Nous appelons *real run* le cas où SCoD s'exécute en temps réel (c.-à-d. la cas ayant donné les résultats de la section précédente). Nous appelons *ideal run* le meilleur résultat qui pourrait être obtenu avec SCoD, c'est-à-dire avec les meilleurs coefficients de pondération possibles. Ce dernier cas est obtenu grâce à plusieurs exécutions de SCoD. Chaque exécution est lancée avec une variante des fonctions de retard présentées dans le tableau 3.1. Bien évidemment, la variante *ideal run* est non réalisable en pratique, car elle exige que les mouvements futurs soient connus à l'avance (en d'autres termes, elle consiste en une version de type Oracle de SCoD). Mais elle nous permet d'appréhender le potentiel de notre proposition.

Dans la figure 3.9, nous montrons le nombre de transmissions pour différentes valeurs de $|\mathbf{A}|$ et τ. Nous constatons que le nombre de transmissions est quasi identique et ce, quelle que

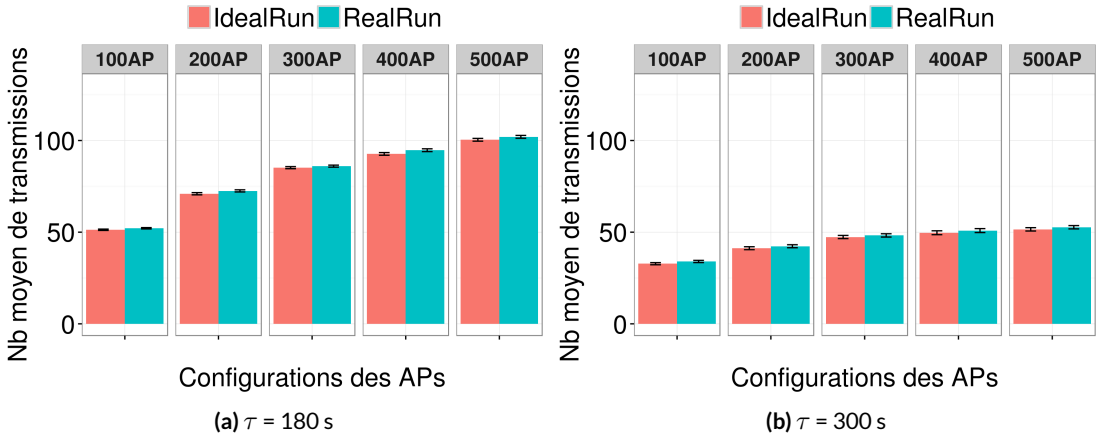


Figure 3.9 – Nombre moyen de transmissions des variantes real run et ideal run pour Luxembourg.

soit la configuration des AP ou la valeur du deadline. Cela signifie que SCoD (real run) prend les décisions appropriées par rapport à son meilleur cas (ideal run).

La fonction de retardement (fréquence d'utilisation)

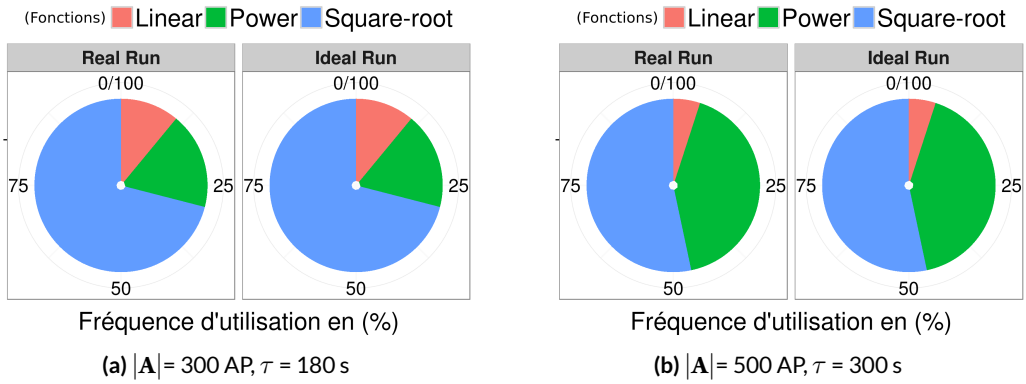


Figure 3.10 – Les graphiques montrent les trois fonctions de retardement $\beta(t)$ utilisées par l'ensemble de données real run et ideal run pour la trace de Luxembourg.

Nous affinons notre évaluation en mesurant la fréquence d'utilisation de chaque fonction de retardement $\beta(t)$. Cela pourra aider un opérateur de réseau à choisir de façon appropriée sa propre fonction en se basant sur la répartition de son trafic et le comportement de mobilité de ses utilisateurs. La figure 3.10 montre la fréquence de choix et d'utilisation de chacune des

trois fonctions de retardement possibles, et ce pour les deux variantes de SCoD (real run et ideal run). Nous illustrons les observations cumulatives de chaque fonction pendant les différents intervalles de diffusion. Nous observons une corrélation totale entre les deux variantes pour les deux configurations simulées. Ces résultats confirment notre observation précédente (dans la section 3.6.1), comme quoi SCoD réussit à choisir les paramètres appropriés.

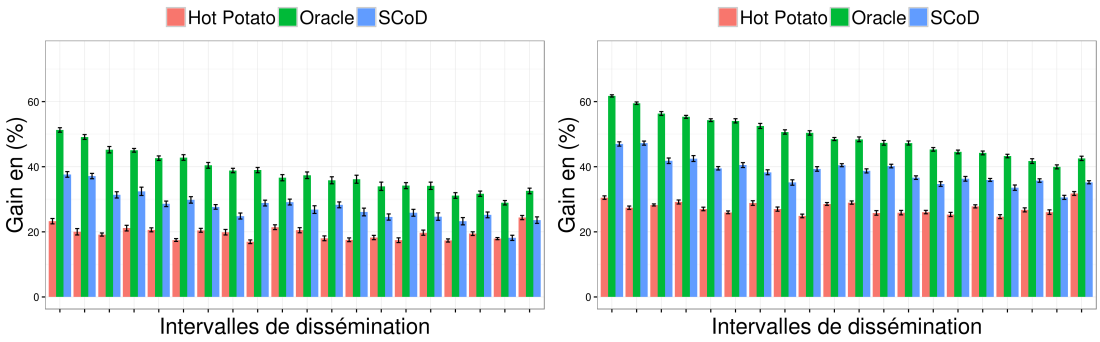
L'observation la plus importante est que, lorsque nous changeons $|\mathbf{A}|$ et τ, le ratio des différentes fonctions de latence change considérablement. Lorsque le délai est allongé, la population diminue puisque les utilisateurs ne restent pas suffisamment longtemps dans la trace. Dans le même temps, nous avons augmenté le nombre de cellules, ce qui réduit leur taille. Encore une fois, nous observons des rassemblements plus petits dans toute la topologie. En conséquence, SCoD adapte sa fonction de retardement en adoptant la fonction *Power* plus fréquemment. Cette fonction a en effet la caractéristique d'aller rapidement aux plus petites valeurs de regroupement pendant un intervalle et donc, de capturer des grappes plus petites.

3.6.2 Dataset de Bologne

Gain moyen

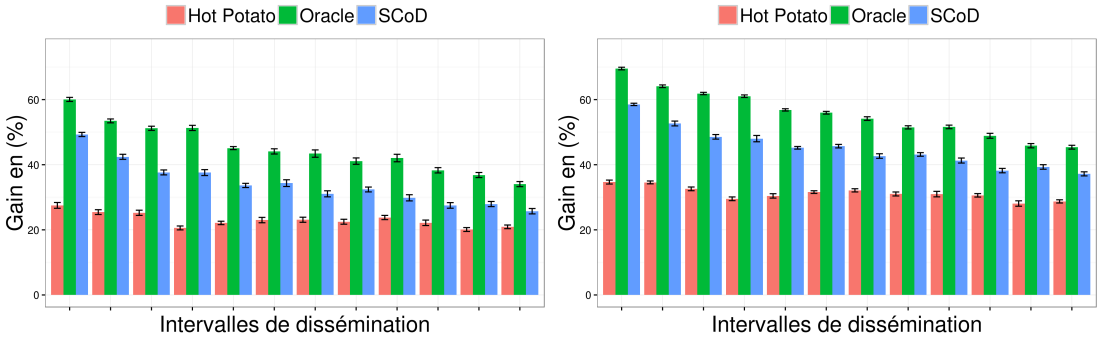
Selon la même approche que pour l'ensemble de données précédentes (trace de Luxembourg), nous montrons dans les figures 3.11 et 3.12, le gain moyen de transmission tel que défini dans l'équation 3.3, obtenu ici pour la trace de Bologne. Nous séparons également ici les résultats pour différencier la variation des paramètres (τ , A) et leurs impacts sur SCoD. Nous avons remarqué au cours de notre analyse que la taille de la population servie à chaque intervalle dans la ville de Bologne est environ $4\sim8$ fois plus grande que celle sur la ville de Luxembourg. En outre, la spécificité du modèle de mobilité des utilisateurs à Bologne leur permet de rester plus longtemps dans la ville, ce qui jouera en faveur de notre stratégie.

Nous observons dans la figure 3.11 que, comme prévu, SCoD dépasse la stratégie Hot Potato dans presque tous les cas (sauf pour deux intervalles de la figure 3.11a). Même avec un petit délai, SCoD conduit à des améliorations significatives lorsque nous changeons la taille de la topologie. Comme cela a été expliqué précédemment, le nombre de points d'accès déployés augmente (c'est-à-dire que les cellules sont plus petites) et les mouvements des utilisateurs entraînent des transferts plus fréquents (mais pas autant que pour les utilisateurs dans la ville de Luxembourg). Ceci se traduit par une variation très rapide de la densité des cellules. SCoD exploite bien un tel phénomène et réalise des gains qui se rapprochent de plus en plus du meilleur résultat possible, à savoir l'Oracle.



(a) $|\mathbf{A}|$ = 300 AP, τ = 180 s **(b)** $|\mathbf{A}|$ = 500 AP, τ = 180 s

Figure 3.11 – Le gain moyen de transmission pour un deadline $\tau = 180\ s$ à Bologne : SCoD vs. les stratégies de base.



(a) $|\mathbf{A}|$ = 300 AP, τ = 300 s **(b)** $|\mathbf{A}|$ = 500 AP, τ = 300 s

Figure 3.12 – Le gain moyen de transmission pour un deadline $\tau = 300\ s$ à Bologne : SCoD vs les stratégies de base.

Nous illustrons dans la figure 3.12 le gain obtenu grâce à une variation de topologie avec un délai étendu ($\tau = 300$ s). Le raisonnement présenté dans les résultats précédents pour la trace de Luxembourg est confirmé. Ceci est intrinsèquement lié à la philosophie derrière notre stratégie, où nous profitons autant que possible de la mobilité des nœuds. Lorsque le délai augmente, cela donne plus de temps à notre stratégie pour s'adapter et atteindre un meilleur regroupement de nœuds. De plus, au fur et à mesure que les utilisateurs mobiles restent plus longtemps dans la trace de Bologne, la population qui sera servie est plus grande ; par conséquent, le taux de regroupement des utilisateurs est positivement influencé.

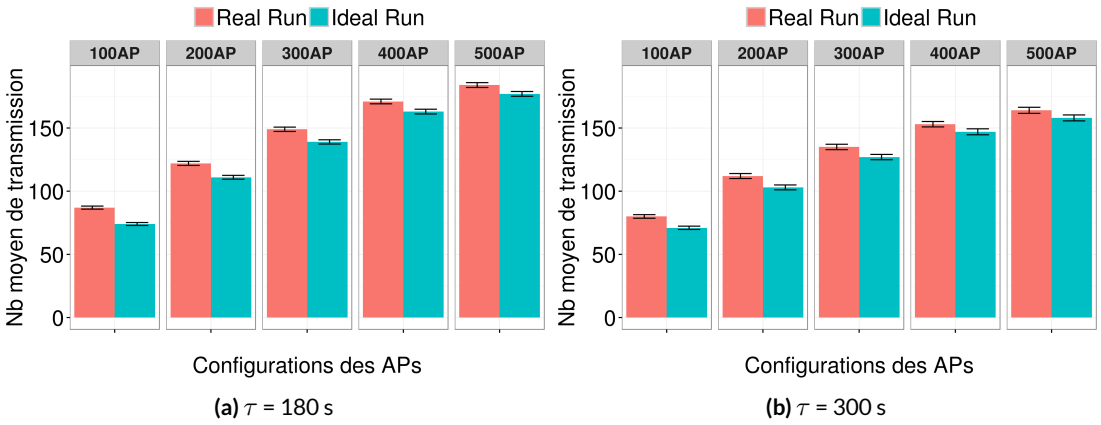


Figure 3.13 – Nombre moyen de transmissions des variantes real run et ideal run pour Bologne.

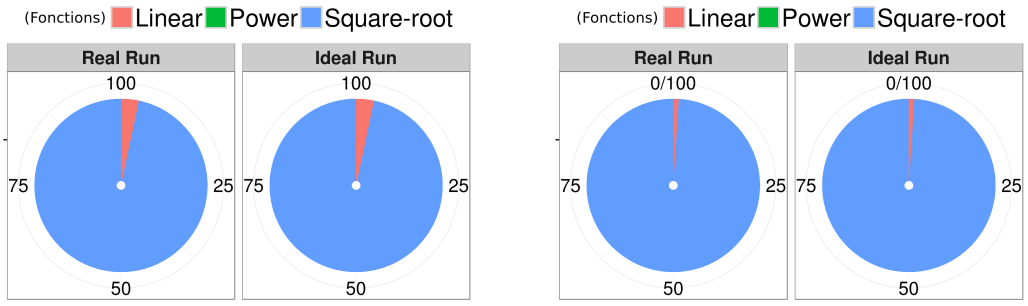
De même que dans la section 3.6.1, nous nous sommes intéressés à évaluer l'adaptabilité de notre système dans la ville de Bologne. Pour ce faire, nous comparons le nombre de transmissions générées par notre stratégie en temps réel (*real run*) au meilleur cas, ce dernier correspondant correspond au cas où SCoD s'exécute avec les meilleurs coefficients de pondération possibles (*ideal run*).

Nous observons dans la figure 3.13 que la taille de la configuration et le nombre de transmissions sont fortement corrélés, comme prévu : si nous augmentons le nombre de points d'accès, le nombre de transmissions augmente également. Nous notons également que la différence entre le cas réel et le cas idéal n'est finalement pas si importante. Cela confirme en quelque sorte que l'utilisation des mesures de l'intervalle le plus récent est suffisante pour paramétrer SCoD et obtenir de bons résultats.

3.6.3 La fonction de retardement (fréquence d'utilisation)

Afin d'aider un concepteur de réseau à choisir les valeurs les plus appropriées pour ses déploiements, nous étudions également comment la nature de la fonction de retardement, $\beta(t)$, influe sur les résultats. La figure 3.14 montre la fréquence à laquelle chaque fonction conduit au meilleur gain. Les résultats représentent les observations cumulatives sur toute la période de simulation (c'est-à-dire pour plusieurs intervalles). Nous comparons également les stratégies real run et ideal run pour confirmer les résultats mentionnés ci-dessus. Dans l'ensemble, la fonc-

tion *Square-root* surpasse les autres fonctions dans presque tous les cas (sauf pour la fonction *Linear* qui est la meilleure dans environ $1\% \sim 3\%$ des cas).



(a) $|\mathbf{A}|$ = 300 AP, τ = 180 s **(b)** $|\mathbf{A}|$ = 500 AP, τ = 300 s

Figure 3.14 – Comparaison des fonctions de retardement $\beta(t)$ pour la trace de Bologne.

Nous remarquons une légère diminution du pourcentage de la fonction *Square-root* lorsque nous augmentons τ. Lorsque le deadline augmente, le système favorise une fonction plus conservatrice, car les utilisateurs ont plus de possibilités d'être regroupés. Dans notre cas, l'utilisation de la fonction *Square-root* (qui diminue très lentement par rapport aux autres fonctions) se comporte exactement de cette façon. La nature de la fonction caractérise si la stratégie est égoïste, c'est-à-dire, combien de temps elle gardera le contenu avant de le libérer pour des transmissions potentielles.

De plus, la fonction *Power* tombe très rapidement à des petites valeurs, ce qui permet à SCoD de transmettre sans attendre des rassemblements possibles. Dans ce cas, notre stratégie a un comportement proche de celui Hot Potato. En fait, la fonction *Power* n'est jamais sélectionnée.

3.7 Conclusion

Dans ce chapitre, nous avons proposé et évalué SCoD (*Scheduled Content Delivery*), une nouvelle stratégie d'économie de ressources pour la dissémination de contenus dans les environnements cellulaires. Notre stratégie profite de la mobilité des utilisateurs qui ont tendance à se regrouper et de leur tolérance à recevoir des contenus retardés. Le but de notre stratégie est de retarder la transmission d'un contenu donné dans les cellules hébergeant un nombre suffisant d'utilisateurs pour limiter le coût des transmissions. En maximisant l'avantage attendu d'une transmission en termes de récepteurs, SCoD réussit à réduire le nombre total de transmissions nécessaires pour satisfaire tous les utilisateurs.

Nous avons comparé SCoD à d'autres stratégies : la stratégie Hot Potato qui recrée l'opération de base d'un réseau cellulaire standard et l'Oracle qui nécessite la connaissance des futurs mouvements des utilisateurs. Notre évaluation a porté sur deux ensembles de données véhiculaires recueillies dans les villes de Bologne et Luxembourg. Nous avons émulé plusieurs déploiements de réseaux cellulaires couvrant les deux villes en faisant varier le nombre de points d'accès et par conséquent la taille des cellules. Nous avons également examiné différents types de contenu, chacun avec un délai différent. Nos résultats d'évaluation montrent que SCoD réduit considérablement les transmissions cellulaires, en particulier lorsque le deadline du contenu augmente et lorsque la taille des cellules diminue.

Il ne faut pas compter sur ceux qui ont créé les problèmes pour les résoudre.

Albert Einstein

4

Résister au déferlement de données cellulaires au moyen de communications opportunistes

4.1 Introduction

Dans ce chapitre, nous traitons de nouveau le problème de la diffusion de contenus dans un réseau cellulaire sous très forte charge, mais avec une approche différente de celle adoptée dans le chapitre 3. Malgré la surcharge de trafic, le réseau se doit de continuer à satisfaire les demandes de ses utilisateurs. Pour faire face à ce challenge, nous étudions l'intérêt d'une utilisation conjointe de communications opportunistes (*device-to-device*, D2D) et de communications cellulaires. Les communications opportunistes ont été rendues possibles par la disponibilité de plusieurs technologies d'accès dans les terminaux mobiles. Pour des raisons de simplicité, et sans perte de généralité, nous considérerons ici que le Wi-Fi est l'alternative à la technologie de l'opérateur. Cela évitera toute gestion supplémentaire des interférences et préservera le précieux spectre cellulaire, ce qui est notre objectif principal. Dans le cas où le trafic est commuté, cela signifie que le délai de livraison peut augmenter. Pour cette raison, nous rappelons que notre application cible concerne les services qui acceptent un délai supplémentaire

avant d'obtenir leur contenu (contenu différé).

À ce stade, nous proposons d'étendre le domaine d'analyse des communications opportunistes en étudiant leur efficacité sous un autre angle. En particulier, plusieurs travaux dans la littérature ont considéré que les utilisateurs, ayant reçu le contenu demandé, peuvent le transférer à leurs voisins situés à portée [50, 90]. Cependant, notre idée est d'exploiter un *sous ensemble* des utilisateurs en tant que mules de données (*data mules*, DM) d'un contenu demandé. Ce groupe d'utilisateurs est le seul autorisé à transmettre une information de manière opportuniste. Le choix des DM doit être validé par l'infrastructure pour éviter des problèmes de sécurité. Par ailleurs, les communications D2D dépendent de la bonne volonté des utilisateurs qui devraient en retour recevoir une compensation. Notre travail s'inspire du rôle des mules de données dans l'étude réalisée par Rebecchi et al., où les auteurs ont pris en compte la récompense des participants au processus de transmission dans leur stratégie de diffusion [81].

Dans ce chapitre, nous considérons que chaque groupe de mules de données est valide pour un intervalle de temps. [1] À la fin de l'intervalle de temps considéré, l'infrastructure doit arroser les cellules contenant encore au moins un utilisateur sans le contenu demandé. Étant donné que les utilisateurs sont mobiles et qu'ils peuvent communiquer pendant leur mobilité, une mule de données peut joindre directement d'autres utilisateurs sans l'intervention de l'infrastructure. En d'autres termes, les mules de données agissent comme un pont entre l'infrastructure cellulaire et le reste des abonnés, pour chaque contenu demandé.

Nous étudions notre problème au travers de plusieurs stratégies afin d'enrichir l'analyse, en suivant une méthodologie similaire à celle présentée dans le chapitre 3. Plus précisément, nous définissons *Hot Potato* (HP) comme la stratégie de base qui transmet le contenu par l'infrastructure dès qu'il est disponible. Cette stratégie ne considère aucune notion d'économie des ressources puisque le contenu est immédiatement transmis quel que soit le coût. Les cellules qui hébergent au moins un utilisateur sont impliquées, ce qui signifie que le nombre de transmissions dépend du nombre de cellules occupées. Ensuite, nous appelons *Oracle* (O) la stratégie qui suppose une connaissance *préalable* de chaque contact utilisateur pendant un intervalle de temps (période de validité d'un contenu). La stratégie est censée connaître toutes les configurations de réseau [2] pour le même intervalle. Selon ces informations, la stratégie peut choisir les mules de données les plus importantes et optimales qui répondent aux besoins des utilisateurs avec moins d'émissions. Cette stratégie est utilisée comme une borne supérieure car elle

1. Rappel : il s'agit de l'intervalle de temps entre l'instant où le contenu a été rendu disponible et l'instant où le contenu n'est plus valide, car nous traitons des contenus tolérants au délai.

2. Rappel : une configuration de réseau est une capture instantanée du réseau à un instant spécifique, comprenant les utilisateurs et les AP qui les hébergent à cet instant.

n'est pas causale (non réalisable). Enfin, *Wait-and-See* est une stratégie causale qui s'appuie sur le lien D2D et le lien cellulaire pour gérer le processus de diffusion. Dans cette troisième stratégie, la communication opportuniste est prioritaire devant les transmissions cellulaires. Cela veut dire que l'infrastructure est conservée comme solution de repli pour garantir à 100% la livraison de contenu.

Nous confrontons nos résultats grâce à une évaluation sur la trace de données véhiculaires de la ville de Bologne [61] déjà utilisée dans le Chapitre 3. Tout d'abord, nous évaluons le gain en termes de transmissions que notre stratégie pourrait atteindre par rapport aux autres stratégies (O, HP), puis nous illustrons l'efficacité du D2D par rapport à plusieurs paramètres. Nous observons que la stratégie Wait-and-See surpasse la stratégie HP dans tous les cas. En outre, elle se rapproche du gain de la stratégie Oracle en améliorant les paramètres D2D (portée des nœuds et taille du groupe des mules de données). Deuxièmement, nous étudions le compromis qui devrait être considéré par l'opérateur lorsque les communications D2D sont récompensées. Ces résultats montrent que les communications opportunistes sont moins bénéfiques, voire pas du tout, lorsque leur coût dépasse une certaine fraction du coût des transmissions cellulaires.

4.2 Modèle de système et hypothèses

Nous considérons un même type de service que dans le Chapitre 3, à savoir un service pub/sub [42]. Dans un tel système, les utilisateurs s'abonnent en manifestant leur intérêt pour des contenus particuliers. Quand un tel contenu est disponible, les utilisateurs abonnés sont notifiés de manière asynchrone. Dans cette section, nous décrivons les grandes lignes de notre problème, ce pour bien comprendre l'objectif de l'analyse que nous allons effectuer plus tard. De plus, nous introduirons les hypothèses posées dans le cadre notre étude, hypothèses nécessaires pour limiter la complexité du problème.

4.2.1 Définition du problème

Nous abordons le problème de diffusion d'un contenu populaire **c** au sein d'un réseau cellulaire à capacité limitée. Dans un souci de simplification, nous ne nous intéressons dans un premier temps qu'à la liaison descendante. L'étude de la liaison ascendante fera l'objet de travaux futurs. Pour formaliser notre problème, nous définissons notre système avec les paramètres : $[c, \mathbf{N}^c, \mathbf{M^c}, \mathbf{A}, t_0^c, \tau, r]$ où :

— c : le contenu pour lequel des abonnés ont manifesté leur intérêt. c est diffusé soit directement par des transmissions cellulaires, soit par l'intermédiaire de mules de données

en D2D. Le contenu est considéré comme tolérant au délai, car il permet un délai supplémentaire avant d'être remis aux abonnés mobiles intéressés ;

— $\mathbf{N}^c$: toute la population d'abonnés ayant manifesté le même intérêt pour le contenu c ;

— $\mathbf{M^c}$: le sous-ensemble des nœuds mobiles mandatés pour agir comme (*mules de données*). Une mule de données $m_i \in \mathbf{M^c}$ est caractérisée par sa capacité à communiquer simultanément sur la fréquence sous-licence (cellulaire) et sur la fréquence sans-licence (Wi-Fi). L'utilisation de la fréquence sans licence permet, d'une part, d'éviter toute interférence avec le trafic cellulaire, et d'autre part, de diminuer l'utilisation des ressources cellulaires limitées, ce qui est l'objectif de notre travail. Nous supposons que le groupe de mules de données est un sous-ensemble de la population inscrite pour un service ($\mathbf{M^c} \subset \mathbf{N^c}$) ;

— $\mathbf{A}$: l'ensemble des points d'accès (AP) déployés dans tout le réseau. Nous désignons par cellule une zone du réseau gouvernée par un AP $a_i \in \mathbf{A}$ (sans référence aucune à une technologie en particulier) ;

— t_0^c : le moment où le contenu c est disponible à la fois dans l'infrastructure et dans les mules de données définies. Dès cet instant, le processus de diffusion peut démarrer ;

— τ : le deadline (durée de validité) du contenu c. Cette valeur est la durée maximale tolérée par un utilisateur pour obtenir son contenu. Si le système dépasse cette période avec des abonnés insatisfaits (utilisateurs sans contenu), le processus a échoué (nous ne traitons pas ce cas dans ce chapitre). L'opérateur devra assurer la satisfaction totale des abonnés à la fin de τ, même si cela aura un coût supplémentaire ;

— r : la portée de communication d'une mule de données, dans laquelle elle peut effectuer une transmission réussie avec un ou plusieurs utilisateurs mobiles (transmissions en multicast). Tout utilisateur intéressé dans cette zone peut recevoir le contenu directement sans avoir impliqué l'infrastructure. La portée de la transmission dépend de la technologie utilisée pour les transmissions.

Notre objectif principal est de transmettre un contenu c à une population intéressée N^c avec le moindre coût possible, par rapport à son échéance τ. En particulier, nous cherchons à minimiser le nombre de transmissions effectuées par l'infrastructure cellulaire en transmettant une partie du trafic via les liens opportunistes. Cela améliorera les performances et atténuera la charge sur le lien cellulaire. En sus des communications D2D, l'infrastructure pourra avoir à intervenir à certains moments, afin de combler les manques laissés par les mules de données

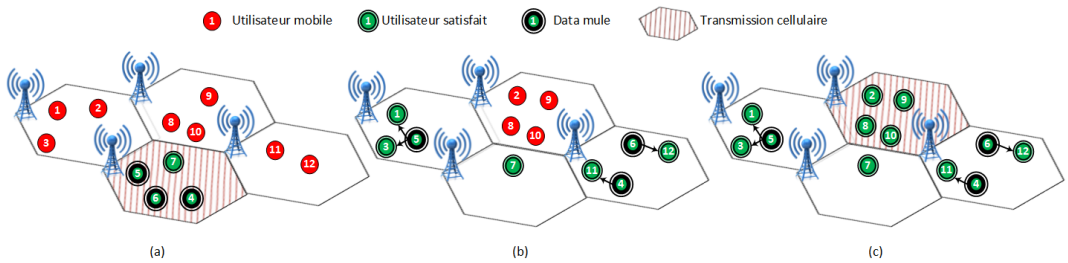


Figure 4.1 – (a) La première transmission satisfait quatre utilisateurs, trois d'entre eux sont choisis comme mules de données (4, 5 et 6). (b) La mobilité des utilisateurs, y compris les mules de données, permet la satisfaction d'un certain nombre d'utilisateurs intéressés de manière opportuniste. (c) Une transmission cellulaire supplémentaire est nécessaire lorsque le deadline du contenu est atteinte afin de satisfaire les utilisateurs n'ayant pas encore reçu le contenu.

(fournir le contenu aux utilisateurs non encore satisfaits). Pour ce faire, nous cherchons à servir autant que possible les abonnés mobiles via le lien D2D. En conséquence, le nombre d'utilisateurs insatisfaits à l'échéance sera réduit, ce qui devrait réduire les émissions cellulaires. En d'autres termes, nous réduisons le nombre de paires (t_i, a_i), où $t_i \in [t_0^c, t_0^c + \tau]$ est l'instant auquel le point d'accès a_i transmet une copie de c et couvre le reste des utilisateurs insatisfaits.

4.2.2 Hypothèses sur la livraison de contenu

Pour atteindre notre objectif et aborder correctement les problèmes mentionnés ci-dessus, il existe plusieurs points clés que nous devons vérifier préalablement. Nous supposons que les utilisateurs restent connectés en permanence à l'interface cellulaire. En outre, ils utilisent leur deuxième interface en tant que lien opportuniste pour toute la période d'activité (pendant τ). Selon la nature de la mobilité du réseau, les utilisateurs peuvent alterner entre les différentes cellules au fil du temps. Cela crée nécessairement des déconnexions avec l'infrastructure lors des changements. Toutefois, étant donné que le temps pour passer d'une cellule à une autre est faible [78], nous supposons que les utilisateurs sont immédiatement associés à un nouvel AP après un transfert de cellule.

Le contenu c est disponible à t_0^c et reste valable pendant l'intervalle $[t_0^c, t_0^c + \tau]$. Nous considérons un seul contenu à la fois (les contenus arrivent de façon séquentielle). En outre, nous supposons que la population des abonnés (N^c) et le groupe des mules de données (M^c) restent inchangés sur le laps de temps $[t_0^c, t_0^c + \tau]$ (les nouveaux arrivants et les départs ne sont pas pris en compte). Nous partons du principe que le système doit assurer la satisfaction totale des utilisateurs à la fin de chaque intervalle (c.-à-d. 100% de livraison de contenu).

Les utilisateurs N^c forment trois sous-ensembles :

— $\mathbf{N}_{a_i}^{c\oplus}(t)$, l'ensemble des utilisateurs satisfaits attachés au point d'accès a_i à l'instant t.

— $\mathbf{N}_{a_i}^{c\ominus}(t)$, le sous-ensemble d'utilisateurs insatisfaits associés au point d'accès a_i à l'instant t.

— $\mathbf{M}^c$ le sous-ensemble des mules de données qui serviront le contenu c d'une manière opportuniste. Notez que $\mathbf{M}^c \subset \mathbf{N}_{a_i}^{c\oplus}(t)$.

Formellement, nous représentons les trois ensembles comme suit :

$$\begin{cases} \mathbf{N}^c = \bigcup_{i=1}^{|\mathbf{A}|} [\mathbf{N}_{a_i}^{c\oplus}(t) \bigcup \mathbf{N}_{a_i}^{c\ominus}(t)], & \forall t \in [t_0^c, t_0^c + \tau], \\ \mathbf{M}^c \subset \mathbf{N}^{c\oplus}(t), & \forall t \in [t_0^c, t_0^c + \tau]. \end{cases} \tag{4.1}$$

Le sous-ensemble des mules de données est choisi après une ou plusieurs transmissions à t_0. Cette ou ces transmissions satisferont un certain nombre d'utilisateurs, parmi lesquels les mules de données seront élues. Le coût de service de ces utilisateurs sera ajouté au coût global ultérieur, calculé à la fin du processus de transmission. Une fois le processus effectué, un nouveau contenu devient disponible, l'infrastructure suit la même procédure pour élire un autre groupe de mules avec des fonctionnalités équivalentes à l'intervalle précédent.

4.3 Les stratégies de diffusion

Dans cette section, nous répondons à la question suivante : *Est-ce que nous pouvons vraiment bénéficier des communications D2D pour diminuer le trafic sur le réseau cellulaire ?* Pour mesurer le bénéfice de l'intégration de ce type de communications dans un réseau cellulaire, nous proposons de comparer différentes stratégies de diffusion à travers leurs comportements de transmission.

4.3.1 La stratégie Hot Potato (HP)

Cette stratégie transmet un contenu dès qu'il est disponible là où les abonnés sont. Le coût d'une telle stratégie dépend du nombre de cellules occupées au moment de la transmission. En utilisant la notation que nous avons introduite dans la section 4.2.1, un point d'accès $a_i \in \mathbf{A}$ transmet une copie du contenu c en mode multicast à $t = t_0^c$ si et seulement si $\mathbf{N}_{a_i}^{c\ominus}(t_0) \neq \emptyset$ (c'est-à-dire uniquement si au moins un abonné est dans la cellule a_i). En conséquence, le nombre de transmissions est égal au nombre de cellules où les abonnés sont situés à t_0^c. Notez encore une fois que la stratégie Hot Potato est agnostique à la mobilité des utilisateurs et à leur tolérance au délai.

4.3.2 La stratégie Wait-and-See (WS)

Contrairement à Hot Potato, cette stratégie profite des aspects sociaux des utilisateurs et, en particulier, de la mobilité et la capacité de communiquer via deux interfaces simultanément (c.-à-d. cellulaire et Wi-Fi). Ces fonctions aident à gérer une partie du trafic cellulaire par des liens opportunistes. La figure 4.1 illustre le fonctionnement de cette stratégie. À l'arrivée du contenu (à t_0^c), l'infrastructure définit le nombre des mules de données, $|M^c|$ (trois dans l'exemple), qui serviront le contenu c en D2D pour l'intervalle en cours. Les mules choisies devront obtenir le contenu avant le commencement du processus de transmission. Pour ce faire, l'infrastructure sélectionne une ou plusieurs cellules qui répondent au nombre requis des mules. Une ou plusieurs diffusions se déclenchent à t_0^c, ce qui affecte tous les utilisateurs des cellules impliquées. Par conséquent, l'infrastructure pourra élire les mules à l'intérieur des cellules satisfaites. Nous notons que le choix des mules de données sera *aléatoire* puisque nous supposons que l'opérateur n'a pas d'information sur les meilleurs utilisateurs qu'il devra sélectionner.

Comme le temps passe, les utilisateurs se déplacent d'un endroit à l'autre et changent de cellules. Des rencontres entre ces utilisateurs et les mules de données peuvent se produire, ce qui permet aux transmissions D2D d'atteindre les utilisateurs intéressés à portée. Au cours de cette période, l'infrastructure surveille le niveau du processus de diffusion. Une fois la durée de vie du contenu (τ) écoulée, s'il reste des utilisateurs encore insatisfaits présents dans le système, une ou plusieurs transmissions multicast supplémentaires auront lieu pour assurer une satisfaction de 100%.

4.3.3 La stratégie Oracle (O)

Comme défini dans la section 4.1, Oracle est une stratégie censée connaître les contacts de chaque utilisateur dans le futur. En outre, la stratégie a une connaissance préalable des différentes configurations du réseau. Selon cette information, elle peut, d'une part, choisir les mules de données les plus importantes qui satisferont la plus grande partie de la population en D2D, et d'autre part, combler les lacunes émanant du D2D pour assurer une satisfaction totale de tous les utilisateurs intéressés. Qui plus est, elle fournit les meilleurs AP concernés qui transmettent le contenu aux utilisateurs non satisfaits. Ceci est analogue au problème bien connu du *Minimum Weighted Set Cover* [**?**]. Malheureusement, trouver la couverture minimale s'avère être un problème NP-complet. Nous utilisons donc une heuristique simple (*greedy*) qui fournit une approximation raisonnable de la meilleure solution, et ce en un temps raisonnable [28, 85]. Le poids le plus élevé est associé au lien cellulaire pour éviter la transmission par l'AP au début

de l'intervalle. Inversement, le poids le plus faible est affecté aux mules de données pour favoriser les transmissions D2D. Nous caractérisons ainsi, grâce aux poids, la priorité d'utiliser un lien spécifique pour transmettre le contenu. En d'autres termes, Oracle favorisera les liens D2D avec des poids faibles puis les liens cellulaires avec des poids élevés. Nous notons que Oracle est un système non causal (non réaliste) ; il envoie le contenu exclusivement à un ensemble d'utilisateurs pré-calculés. Nous appliquons néanmoins cette stratégie afin d'obtenir une limite supérieure pour notre analyse.

4.4 La compagne d'évaluation

4.4.1 Trace de données

Afin d'évaluer l'efficacité de l'utilisation de communications D2D dans un réseau cellulaire, nous utilisons une des traces de données déjà exploitées dans le chapitre 3 ; cette trace de données véhiculaire a été générée avec le simulateur de trafic SUMO [60] et représente le trafic de plus de 18000 véhicules dans la ville de Bologne (Italie) [61]. La zone géographique sous-jacente à nos simulations est d'environ 10 km^2 et comporte plus de 191 km de routes [90].

Nous émulons un réseau cellulaire dans cette zone à l'aide d'un diagramme de Voronoi qui couvre l'ensemble de la zone géographique ciblée, où chaque cellule de Voronoi contient un seul point d'accès [14]. Cela nous donne une certaine souplesse pour évaluer notre proposition sous différentes conditions de densités. En outre, d'autres paramètres tels que la portée d'une mule de données (r) et la taille du groupe des mules de données (M^c) seront introduits.

Nous simulons plusieurs plans de déploiement obtenus avec une distribution uniforme. Nous générons cinq topologies différentes pour chaque nombre de points d'accès $|\mathbf{A}| \in \{100, 200, 300, 400, 500\}$, ce qui nous donne 25 scénarios au total. En ce qui concerne le deadline du contenu, nous considérons plusieurs scénarios avec les valeurs suivantes de $\tau \in \{60\text{ s}, 120\text{ s}, 180\text{ s}, 300\text{ s}\}$. Les résultats que nous présenterons dans la prochaine section sont la moyenne de cinq séries de simulations effectuées pour chaque taille de topologie.

4.4.2 Opérateur de réseau égoïste vs utilisateurs altruistes

Nous comparons la stratégie Wait-and-See avec les autres stratégies présentées dans la section 4.3 : Oracle (O) et Hot Potato (HP) et ce, dans un premier temps, en l'absence de considération de coûts de communication D2D. Cela signifie que nous supposons que le coût d'une transmission entre une mule de données et un utilisateur ordinaire est négligeable par rapport à

celui d'une transmission cellulaire – c'est d'ailleurs l'hypothèse communément admise dans la littérature. En outre, nous définissons une stratégie de base que nous appelons "*Full Coverage*" (FC), où l'infrastructure distribue systématiquement le contenu dans chaque cellule qui a reçu la visite d'au moins un utilisateur intéressé pendant l'intervalle $[t_0^c, t_0^c + \tau]$. Cette stratégie est clairement trop naïve pour être mise en œuvre et ne sert qu'à fournir une limite inférieure pour notre évaluation.

Gain en terme de transmissions. Nous suivons les mêmes étapes que dans le chapitre 3 pour cette partie d'évaluation. Pour cela, nous considérons le *gain* $\mathcal{G}$ pour comparer les solutions. Ce paramètre mesure les économies offertes par une solution donnée par rapport à la stratégie de couverture complète FC. Il est défini comme suit :

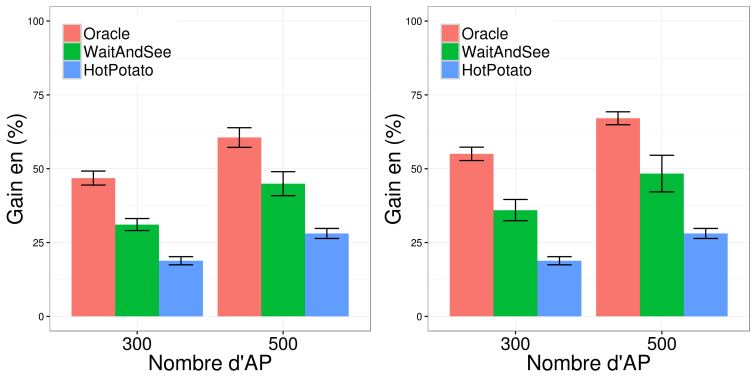
$$\mathcal{G}_{\text{stratégie}} = \frac{N_{\text{FC}} - N_{\text{stratégie}}}{N_{\text{FC}}}, \tag{4.2}$$

où N_{FC} est le nombre de transmissions lorsque la stratégie de couverture complète est utilisée et $N_{\text{stratégie}}$ est le nombre de transmissions pour la "stratégie" considérée.

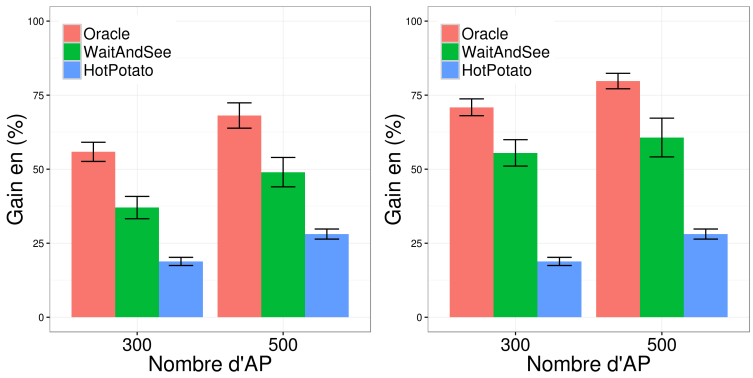
Couverture D2D vs couverture AP. Nous comparons également le degré de satisfaction des utilisateurs. Cette évaluation donne le nombre d'utilisateurs couverts par les mules de données et le nombre de ceux couverts par l'infrastructure cellulaire. Ce critère de performance illustre à quel point une stratégie profite des échanges opportunistes pour satisfaire les utilisateurs intéressés. Nous évaluons cette métrique sous plusieurs combinaisons des paramètres présentés dans la section 4.4.1, en particulier la portée et le nombre des mules de données.

4.4.3 Le D2D est-il bénéfique tout le temps ?

Dans la section précédente 4.4.2, les utilisateurs qui participent aux échanges opportunistes (les mules de données) ne sont pas récompensés. L'opérateur se comporte comme une entité égoïste qui profite des utilisateurs sans leur donner de compensation. En pratique, intégrer le D2D dans un système cellulaire doit être soigneusement analysé, sous différentes dimensions, y compris le coût. En conséquence, nous allons dans cette section mettre en lumière le cas où l'opérateur doit prendre en considération le coût des échanges D2D. Dans cette éventualité, les mules de données ne sont pas traitées comme les autres nœuds puisqu'elles doivent être récompensées pour permettre l'utilisation des capacités de leurs appareils dans le processus de diffusion. Néanmoins, l'opérateur devra s'assurer que le coût gratifiant de ces derniers (φ_{D2D})



(a) τ = 180 s, M^c = 1%, r = 10m **(b)** τ = 180 s, M^c = 1%, r = 100m



(c) τ = 180 s, M^c = 10%, r = 10m **(d)** τ = 180 s, M^c = 10%, r = 100m

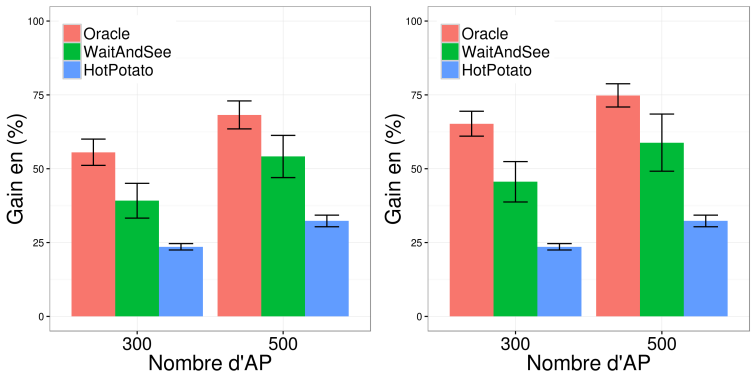
Figure 4.2 – Gain moyen de transmission pour les stratégies étudiées : Wait-and-See, Hot Potato et Oracle. Le deadline considéré est τ = 180 s.

ne doit pas excéder le coût d'une transmission cellulaire (φ_{AP}). Nous étudions ici le compromis qui doit être fait en ce qui concerne les deux types de coût précédents.

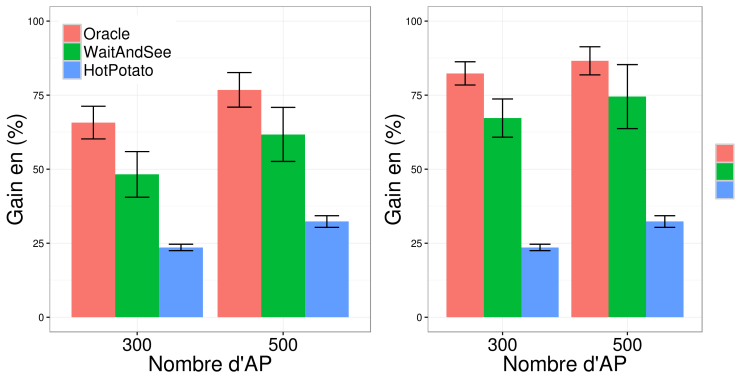
Étant donné que la stratégie Wait-and-See permet et des transmissions réalisées par le réseau cellulaire et des transmissions opportunistes en D2D, nous définissons son coût total pour transmettre un contenu c comme suit :

$$x^c.\varphi_{\text{D2D}} + y^c.\varphi_{\text{AP}} = \Phi_c^{\text{WS}}, \tag{4.3}$$

où x^c et y^c sont, respectivement, les nombres moyens de transmissions réalisées par les mules de données et par les points d'accès pour diffuser le contenu c. φ_{D2D} est la récompense qu'un mule de données doit recevoir de l'opérateur pour une transmission effectuée. De même, φ_{AP} est le coût d'une transmission cellulaire via un point d'accès. Enfin, Φ_c^{WS} est le coût total que le



(a) τ = 300 s, M^c = 1%, r = 10m **(b)** τ = 300 s, M^c = 1%, r = 100m



(c) τ = 300 s, M^c = 10%, r = 10m **(d)** τ = 300 s, M^c = 10%, r = 100m

Figure 4.3 – Gain moyen de transmission pour les stratégies étudiées : Wait-and-See, Hot Potato et Oracle. Le deadline considéré est τ = 300 s.

réseau a subi pour transmettre un contenu c selon Wait-and-See. En même temps, nous définissons dans l'équation 4.4 le coût de la stratégie Hot Potato comme suit :

$$z^c.\varphi_{\text{AP}} = \Phi_c^{\text{HP}}, \tag{4.4}$$

où z^c est le nombre moyen de transmissions effectué par la stratégie HP. Φ_c^{HP} est le coût total associé à la stratégie HP pour atteindre toute la population intéressée par le contenu c.

D'après les équations 4.3 et 4.4, nous extrayons la relation entre φ_{D2D} et φ_{AP} :

$$\frac{\varphi_{\text{D2D}}}{\varphi_{\text{AP}}} = \frac{x^c + y^c}{z^c}. \tag{4.5}$$

Nous considérons ici que $\varphi_{\text{AP}} = 1$ unité de mesure. La section suivante illustre le point à

partir duquel l'incorporation du D2D devient inutile car le montant des compensations versées aux mules par l'opérateur devient supérieur au coût des transmissions cellulaires. Nous mettons en évidence la variabilité du compromis en fonction des différents paramètres de réseau.

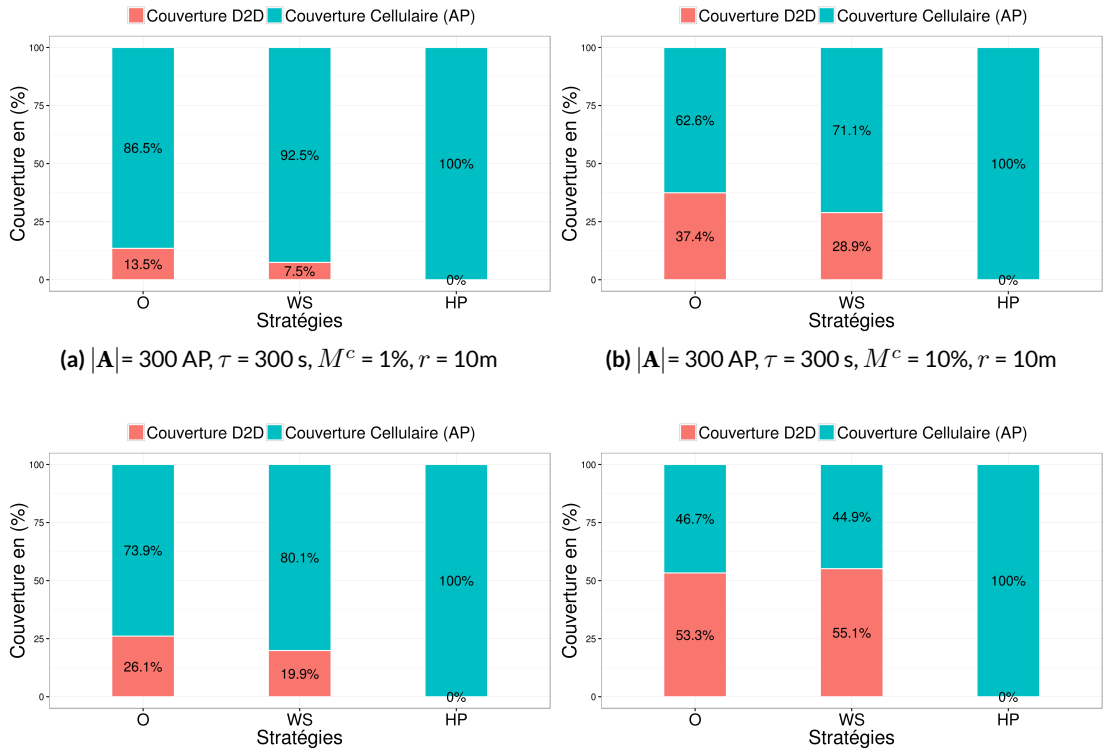
4.5 Les résultats numériques

4.5.1 Gain en terme de transmissions

Nous évaluons d'abord le gain moyen de transmissions pour notre stratégie, conformément à la définition donnée dans l'équation 4.2. Chaque point de l'axe des abscisses représente la moyenne de plusieurs intervalles de même durée τ – en d'autres termes, chaque barre dans la figure correspond à la diffusion d'un contenu dans plusieurs intervalles. L'axe des ordonnées donne le gain $\mathcal{G}_{\text{stratégie}}$ par stratégie. Dans le même temps, les figures montrent les résultats pour deux valeurs des configurations des points d'accès ($|\mathbf{A}|$={300, 500}) et deux valeurs de deadline (τ={180 s, 300 s}). Nous avons fait varier la portée des communications opportuniste avec les valeurs $r = \{10\ m, 100\ m\}$. Nous avons exprimé le nombre de mules de données en pourcentage de la population totale présentée lors de la diffusion de chaque intervalle ($M^c = \{1\%, 10\%\}$).

Dans la figure 4.2, nous illustrons l'effet de plusieurs paramètres sur les trois stratégies. En particulier, l'effet d'un petit délai pour des nombres de points d'accès différents et l'impact des paramètres D2D sont montrés afin d'appréhender le comportement des stratégies Oracle et Wait-and-See, car elles sont influencées uniquement par ces communications opportunistes. Nous remarquons que la stratégie Wait-and-See surpasse la stratégie Hot Potato et qu'elle se rapproche de l'Oracle dans tous les cas, indépendamment du paramètre que l'on fait varier : ceci était attendu puisque le coût du D2D n'est pas pris en compte.

Nous observons que l'augmentation de la portée des mules de données (figure 4.2a, 4.2b) a un léger impact positif sur le gain. C'est parce que les mules de données ont un espace plus large pour effectuer des transmissions, mais cela ne signifie pas systématiquement beaucoup plus de contacts. La portée dépend absolument de la technologie utilisée par les mules (c'est-à-dire Wi-Fi ou Bluetooth), où même de la version d'une même technologie. Cet argument ne s'applique pas au cas de la figure 4.2c, où le nombre plus grand de mules de données se traduit forcément par plus de contacts. Ce comportement est illustré avec les deux configurations de point d'accès. Enfin, dans la figure 4.2d, pour la stratégie Wait-and-See, nous obtenons près de 20% d'amélioration par rapport au premier cas. Nous pensons que cette amélioration est significativement importante car nous utilisons un deadline de petite valeur $\tau = 180$ s).



(a) $|\mathbf{A}|$ = 300 AP, τ = 300 s, M^c = 1%, r = 10m

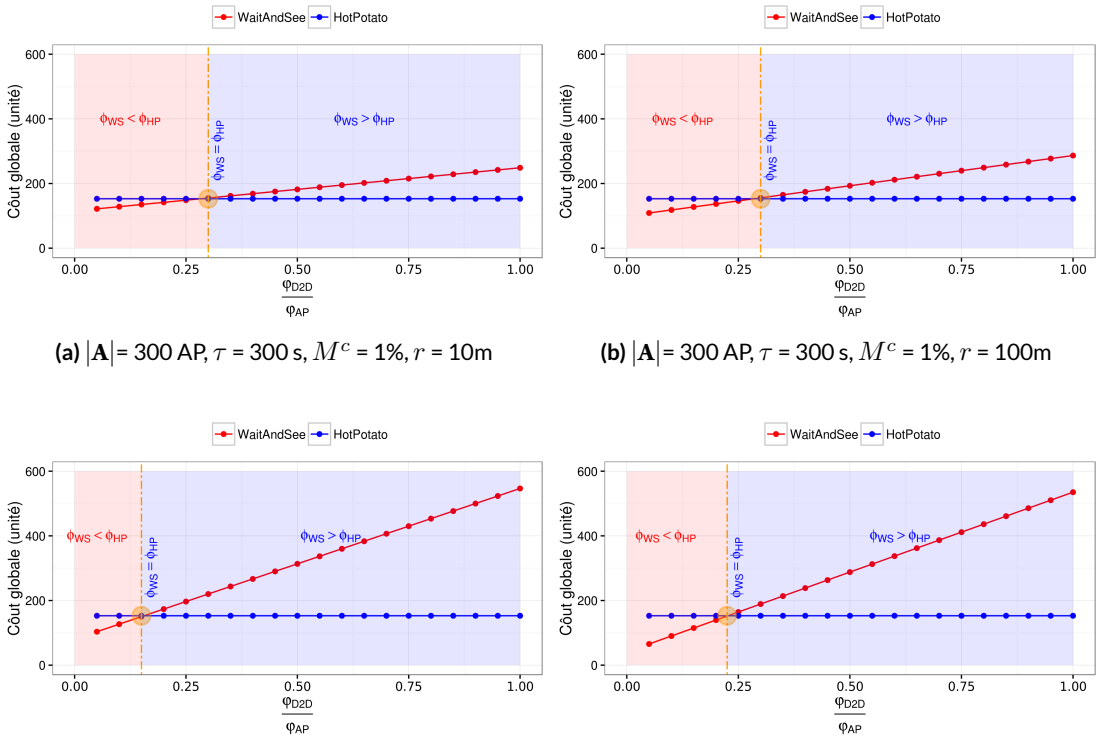
(b) $|\mathbf{A}|$ = 300 AP, τ = 300 s, M^c = 10%, r = 10m

(c) $|\mathbf{A}|$ = 300 AP, τ = 300 s, M^c = 1%, r = 100m

(d) $|\mathbf{A}|$ = 300 AP, τ = 300 s, M^c = 10%, r = 100m

Figure 4.4 – Le pourcentage de couverture des utilisateurs intéressés avec les mules de données et les points d'accès cellulaires. Plusieurs paramètres sont illustrés.

Dans la figure 4.3, nous montrons les résultats pour un délai plus important. Selon notre hypothèse, nous considérons uniquement les utilisateurs qui restent présents sur l'intégralité de l'intervalle. Nous observons à nouveau que Wait-and-See surpasse la stratégie Hot Potato dans tous les cas. Ici, notre stratégie est encore plus proche de l'Oracle. Nous notons que toutes les stratégies ont économisé beaucoup plus de ressources, car le nombre de transmissions a diminué, ce qui va de pair avec la diminution du nombre d'utilisateurs. Les mules de données ont beaucoup plus de temps pour déclencher des transmissions et diminuer la charge de l'infrastructure cellulaire en fin d'intervalle. En suivant la même tendance que la figure 4.2, la figure 4.3 confirme le raisonnement présenté dans les résultats déjà mentionnés (τ =180 s) lorsque nous avons fait varier la portée des nœuds et la taille du groupe des mules de données. Néanmoins, nous augmentons aussi le temps de validité, ce qui est traduit par plus de temps pour les utilisateurs y compris les mules afin d'effectuer des déplacements et rencontrer d'autres utilisateurs.



(a) $|\mathbf{A}|$= 300 AP, τ = 300 s, M^c = 1%, r = 10m

(b) $|\mathbf{A}|$= 300 AP, τ = 300 s, M^c = 1%, r = 100m

(c) $|\mathbf{A}|$= 300 AP, τ = 300 s, M^c = 10%, r = 10m

(d) $|\mathbf{A}|$= 300 AP, τ = 300 s, M^c = 10%, r = 100m

Figure 4.5 – La relation entre le coût cellulaire et le coût des transmissions opportunistes.

4.5.2 Couverture D2D vs couverture AP

Nous raffinons nos évaluations en mesurant le *pourcentage de couverture* pour les deux modes de transmission : opportuniste et cellulaire. Cela signifie que nous montrons le degré d'activation des deux types de communication. Étant donné que les résultats ont presque le même aspect, nous évitons la répétition en illustrant un extrait des résultats avec un deadline $\tau = 300$ et une configuration de points d'accès $|A| = 300\ AP$. Nous conservons les mêmes paramètres pour les communications D2D.

La figure 4.4 montre le pourcentage d'utilisateurs couverts, respectivement, par les mules de données et les points d'accès. L'observation la plus importante est que le pourcentage de couverture augmente au fur et à mesure que nous étendons le nombre de mules de données et leur portée de transmission. Cependant, nous constatons que l'augmentation du nombre de mules a beaucoup plus d'impact sur la participation au D2D, ce qui confirme ce que nous avons expliqué dans la section précédente 4.5.1.

En même temps, nous observons dans la figure 4.4d que Wait-and-See atteint une couverture D2D légèrement meilleure que la stratégie Oracle. Ce dernier ayant une connaissance préalable des mouvements et des contacts futurs de chaque nœud, cela lui permet de sélectionner les meilleurs candidats pour les groupes de mules de données. L'Oracle prend également en compte le reste des utilisateurs qui n'ont pas rencontré de mules de données pendant leur trajet. Par conséquent, ces utilisateurs doivent être satisfaits par les points d'accès, ce qui est coûteux et doit donc être utilisé le moins possible. L'Oracle a beaucoup plus de souplesse et de contrôle sur le choix des mules et les AP impliqués dans les transmissions. Par ailleurs, Wait-and-See est une stratégie causale qui permet aux mules de données d'effectuer leurs transmissions pendant tout l'intervalle de durée τ, sans aucune connaissance du futur. À la fin, le reste des utilisateurs doivent être satisfaits quel que soit le coût via la partie cellulaire.

4.5.3 Le D2D est-il bénéfique tout le temps ?

Dans cette section, nous abordons le bénéfice des communications opportunistes sous un autre angle. Nous considérons que le coût des transmissions D2D n'est plus négligeable par rapport au coût cellulaire. Par conséquent, l'opérateur doit contrôler le processus de communication afin de fournir une bonne expérience utilisateur (QoE) tout en continuant à obtenir un bénéfice financier. Ce contrôle comprend le coût pour compenser les mules de données pour leur coopération dans le processus de diffusion. Nous pensons que cette analyse pourra fournir quelques lignes directrices aux opérateurs pour choisir le montant de la compensation financière à donner aux mules tout en restant gagnants.

Dans ce but, nous montrons dans la figure 4.5 la relation entre le D2D et le coût cellulaire. Les différentes régions représentent le coût global des stratégies Wait-and-See et Hot Potato selon les équations 4.3 et 4.4, respectivement. Chaque figure est divisée en deux régions principales ; *Région 1 (rouge)* est la région où $\varphi_{\text{D2D}} < \varphi_{\text{AP}}$, ce qui signifie que l'incorporation D2D est utile et que les mules de données aideront à économiser les ressources sans perte. À l'inverse, *Région 2 (bleu)* est la région où l'opérateur doit éviter les communications D2D, car non rentable : $\varphi_{\text{D2D}} > \varphi_{\text{AP}}$. Nous remarquons que le point de croisement entre les deux stratégies signifie que $\varphi_{\text{D2D}} = \varphi_{\text{AP}}$.

Notre objectif, à partir de cette analyse, est de déterminer le point de croisement à partir duquel le D2D n'est plus utile et bénéfique pour l'opérateur. Nous examinons toute variation potentielle de ce point de croisement selon plusieurs paramètres. Nous observons que l'augmentation de la portée des mules a un impact important sur le point de croisement (légère-

ment décalé vers l'arrière). Ce comportement n'a pas joué en faveur de l'opérateur. D'un côté, la portée étendue signifie que la mule de données pourrait atteindre (mais pas de façon systématique) plus d'utilisateurs intéressés avec une même transmission. De l'autre côté, elle pourrait donner lieu à des transmissions supplémentaires si les utilisateurs sont très dispersés géographiquement dans le réseau. Cette observation est confirmée par la petite élévation de la pente de la ligne Wait-and-See.

L'augmentation de la quantité de mules de données a une influence significatif sur le point de croisement entre les lignes de coûts. Ce comportement est illustré par une augmentation du nombre de transmissions D2D résultant de la stratégie Wait-and-See dans la figure 4.5b. Par conséquent, la région 1 se déplace considérablement en arrière en laissant la région 2 gagner plus d'espace. Cela signifie que le nombre de transmissions effectuées par les mules de données a augmenté, ce qui diminue automatiquement les coûts pour l'opérateur. Nous constatons néanmoins que, en combinant à la fois le nombre de mules de données et leur portée, l'impact est moins agressif que faire varier un seul paramètre à la fois. En d'autres termes, les inconvénients d'un paramètre sont comblés par les avantages de l'autre. Par conséquent, l'opérateur peut tirer parti de ces paramètres pour trouver le compromis le plus approprié pour son réseau.

4.6 Conclusion

Nous avons abordé, dans cette partie du document de thèse, le problème de la diffusion de contenu sous une lourde contrainte de ressources cellulaires, et sous un autre angle d'étude : le réseau est surchargé par le trafic transmis, et il doit toutefois répondre aux demandes des utilisateurs. Pour ce faire, nous avons proposé une stratégie Wait-and-See qui profite des communications D2D pour transférer une partie du trafic initialement destiné à être pris en charge par les points d'accès cellulaires. La stratégie utilise certains utilisateurs comme mules de données à l'arrivée d'un contenu. Ces nœuds sont les seuls autorisés à transmettre de manière opportuniste, en D2D. Les mules de données tirent parti de leur mobilité et de la tolérance au délai des utilisateurs pour effectuer des transmissions chaque fois qu'elles entrent en contact avec des utilisateurs intéressés. Lorsque l'échéance de validité de contenu touche à sa fin, l'infrastructure cellulaire vérifie s'il y a encore des utilisateurs n'ayant pas leur contenu. Si c'est le cas, une transmission est déclenchée dans chaque cellule contenant un utilisateur intéressé et non satisfait.

Nous avons comparé la stratégie Wait-and-See avec deux autres stratégies : Hot Potato qui émule l'opération de base d'un réseau cellulaire et l'Oracle qui nécessite la connaissance a priori des futurs mouvements d'utilisateurs. Notre évaluation a porté sur une trace de données véhiculaires dans la ville de Bologne. Nous avons simulé plusieurs déploiements de réseaux cellulaires couvrant la ville en faisant varier le nombre de points d'accès et par conséquent la taille des cellules. Nous avons également considéré différents types de contenu, chacun avec un délai différent. En outre, nous avons fait varier la portée des mules de données et leur nombre pour étudier le comportement du réseau sous plusieurs conditions. Nos premiers résultats ont porté sur le gain que nous avons obtenu en appliquant le D2D dans le réseau.

Nous avons examiné le cas où l'opérateur est égoïste : il bénéficie des avantages du D2D sans donner de contre-partie pour la participation des mules de données. Dans ce cas, nous avons montré que le gain obtenu par notre stratégie est remarquable par rapport à la stratégie Hot Potato et qu'il se rapproche du meilleur cas (Oracle).

Nous analysons un deuxième cas où nous considérons que les utilisateurs doivent être indemnisés pour leur participation. Pour ce faire, l'opérateur doit déterminer combien coûte une transmission par une mule de données. En conséquence, nous avons fourni une analyse de compromis qui compare la relation entre le coût D2D et le coût cellulaire. Nous avons mis en lumière le point de croisement des deux coûts qui détermine l'égalité entre eux. Plus le coût D2D est proche de ce point de croisement, mieux c'est. Au-delà de ce point, le D2D n'est plus utile et bénéfique pour l'opérateur.

Si vous ne pouvez pas l'expliquer simplement, vous ne le comprenez pas assez bien.

Albert Einstein

5 Conclusion et Perspectives

5.1 Conclusion

Les dix dernières années ont connu une croissance explosive du nombre d'abonnés pour les réseaux mobiles [30]. Cette croissance est principalement due à l'évolution de la technologie et à l'apparition de nouveaux dispositifs (téléphones intelligents et tablettes). En même temps, la prolifération des appareils sans fil a considérablement fait exploser le trafic échangé sur le réseau. Cela a causé une perte en termes de revenu pour les opérateurs, en raison en partie au trafic redondant.

Pour faire face à l'explosion de trafic, les opérateurs ont introduit la quatrième génération des réseaux mobiles à haut débit (4G) [1]. Cependant, même avec cette amélioration, les ressources ne vont pas suffire à répondre aux besoins des utilisateurs dans un avenir très proche. Le même scénario sera probablement le même pour la 5G. Une première réponse pour augmenter la capacité cellulaire a consisté à mettre en place notamment des stations de base supplémentaires. Hélas, cela crée tout d'abord des problèmes d'interférences supplémentaires. De plus, cela génère une augmentation des dépenses (CAPEX/OPEX) pour la gestion et le déploiement des ressources. D'autres alternatives, a priori attrayantes, ont été développées pour

faire face à de tels problèmes. Certaines augmentent la capacité du réseau en réduisant la taille de la cellule. Cela augmentera la fréquence et le taux de transfert [96], car les cellules sont plus petites. Une stratégie plus récente consiste à dévier une partie du trafic vers une technologie complémentaire (par exemple, du cellulaire vers le Wi-Fi [35], les technologies sous bandes avec licence telles que les Femtocellules [39] ou les communications du type dispositif-à-dispositif [13, 80]).

Dans cette thèse, nous avons étudié le problème de la dissémination des contenus dans un environnement mobile sous très forte charge de trafic, et en particulier, celui de la diffusion des contenus populaires et tolérants au délai. Cela signifie que l'infrastructure cellulaire est surchargée par le trafic et qu'un manque de ressources se fait lourdement ressentir, mais qu'elle doit tout de même continuer à répondre aux requêtes des utilisateurs. Tout au long de ce travail, nous nous sommes intéressés à garantir une satisfaction totale des utilisateurs de différentes manières possibles (100% de taux de livraison).

Dans la première partie de cette thèse, nous avons traité le problème via une architecture purement cellulaire. Dans un tel contexte, nous avons proposé SCoD (*Scheduled Content Delivery*), une stratégie de dissémination de contenus dans les environnements cellulaires permettant d'économiser des ressources. Au lieu de percevoir la mobilité des utilisateurs comme un défi supplémentaire à traiter, nous avons choisi au contraire de l'exploiter afin de réduire la charge sur l'opérateur. Pour cela, SCoD tire parti de la mobilité des nœuds ainsi que de la tolérance aux délais pour déterminer les meilleurs moments où un point d'accès devrait transmettre une copie d'un contenu demandé.

De façon plus précise, la stratégie profite de la tendance au regroupement de ces utilisateurs mobiles, dans le but de retarder et de planifier des transmissions dans le temps. La tolérance aux délais des utilisateurs mobiles est une clé très importante dans la conception de SCoD. Le but sous-jacent de notre proposition est de minimiser le nombre d'antennes qui transmettent le même contenu tout en maximisant le nombre d'utilisateurs résidant dans la même cellule. En d'autres termes, son objectif est de permettre aux utilisateurs de se rassembler dans le moins de cellules possible, ce pour réduire le nombre de transmissions multicast.

Dans une seconde partie, nous exploitons une communications du type dispositif-à-dispositif (D2D) comme technologie prometteuse pour la prochaine génération de réseaux mobiles et de systèmes sans fil (5G). Nous présentons une étude du bénéfice que les communications opportunistes peuvent apporter au réseau cellulaire. Nous avons proposé une stratégie intitulée *Wait-and-See*; cette stratégie profite des communications directes entre les terminaux mobiles afin de transmettre des contenus. L'objectif principal de cette stratégie est de faire endosser à

un petit nombre d'utilisateurs mobiles un rôle de *mules de données*. Ces dernières sont les seuls nœuds à pouvoir transmettre le contenu demandé en D2D. Cette partie du travail suppose que l'infrastructure ne récompense pas les mules pour leur participation au processus de diffusion. Par la suite, nous avons examiné un deuxième cas de figure, dans lequel l'opérateur indemnise les utilisateurs qui partageant les capacités de leurs appareils pour le processus de dissémination. Dans cette étude, nous avons mis en avant le compromis entre le coût des transmissions cellulaires et celui des mules de données. Nous avons analysé la sensibilité de ce compromis aux différentes valeurs des paramètres (nombre de mules de données, deadline, portée des nœuds).

Nos différentes propositions ont mis en exergue l'intérêt d'utiliser à la fois la mobilité et la tolérance aux délais pour réduire le trafic redondant dans le canal cellulaire. Nous sommes persuadés que ce type d'approche deviendra inévitable dans les futures générations de réseaux mobiles.

5.2 Perspectives

L'objectif principal de ce travail de thèse a été de proposer des stratégies de dissémination de contenus tolérants au délai, et en particulier, les contenus les plus échangés sur le réseau cellulaire, autrement dit "populaires". Les stratégies que nous avons présentées dans les chapitres précédents diminuent sensiblement la charge sur les ressources cellulaires. Elles reposent essentiellement sur un mécanisme de transmission mettant à profit, d'une part, la mobilité des utilisateurs, et d'autre part, la tolérance à recevoir des contenus d'une manière différée. Autrement dit, nous avons géré la mobilité comme une ressource qui peut être exploitée pour rendre des services aussi bien à l'opérateur cellulaire qu'à ses utilisateurs. Compte tenu de cela, plusieurs orientations pour de futurs travaux peuvent être envisagées afin de relever les défis dans un même axe de recherche, toujours dans le but d'aider les opérateurs à mieux utiliser leurs ressources, tout en satisfaisant la demande de leurs utilisateurs. Nous donnons ici quelques suites possibles à nos travaux :

— **Direction 1.** Comme les utilisateurs sont mobiles par nature, ils peuvent subir fréquemment des transferts intercellulaires (hands-off) tout au long de leurs parcours. En conséquence, des opportunités de regroupements d'utilisateurs avec les mêmes intérêts peuvent se présenter. Dans le même temps, de multiples opérations vont avoir lieu, et en particulier, l'opération de *désassociation* d'un point d'accès et celle de *réassociation* avec le point d'accès le plus proche en terme de signal. L'effet direct de ces deux opérations est la génération d'un délai supplémentaire afin de rétablir une connexion avec l'infrastructure

cellulaire. Pour cette raison, nous souhaitons étendre notre travail avec une étude sur l'impact de la réassociation sur les regroupements des utilisateurs et par conséquent sur notre stratégie SCoD, puisqu'elle exploite les rassemblements des nœuds mobiles.

— **Direction 2.** Le problème de la gestion simultanée de plusieurs contenus nous semble une direction intéressante. Cela nécessite la proposition d'un mécanisme de priorisation pour équilibrer l'évolution de la diffusion des différents contenus. Ce dernier serait couplé avec un système dynamique, qui accepterait les nouveaux arrivants ainsi que les départs avant la terminaison du processus de dissémination. De plus, les utilisateurs devraient avoir à exprimer leurs intérêts à des contenus via un lien ascendant vers l'infrastructure. Nous pensons que la prise en compte d'un tel lien est intéressante pour pouvoir tirer pleinement parti du spectre cellulaire.

— **Direction 3.** L'étude des utilisateurs du point de vue de leur centralité dans le réseau aidera à la diffusion efficace des contenus. Une telle étude passe par la modélisation du réseau à l'aide de deux graphes bipartis $G = (V, U)$, où V est l'ensemble des points d'accès et U l'ensemble des utilisateurs intéressés. Un premier graphe peut être obtenu, si une arête de l'ensemble U vers l'ensemble V signifie qu'un utilisateur mobile dans U a reçu un contenu du point d'accès de V – ce premier graphe peut être vu comme un *graphe de transmission*. Un second graphe peut également être défini, en établissant une arête entre un utilisateur et un AP si cet utilisateur a été au moins connecté une fois sur l'AP – ce second graphe peut être vu comme un *graphe de présence*. L'étude simultanée des deux graphes permettra d'extraire les utilisateurs les plus centrés pour comprendre leur comportement. Cela aidera à améliorer notre stratégie en surveillant les utilisateurs centrés pour une transmission plus efficace des ceontenus.

— **Direction 4.** Nous avons souligné, dans cette thèse, l'importance des solutions opportunistes dans un réseau cellulaire. Du point de vue de l'utilisateur, un frein majeur est la consommation excessive de la batterie en cas d'activation simultanée de plusieurs interfaces sans fil, même en mode veille. Cela pourrait affecter la coopération des utilisateurs dans le processus de dissémination en partageant leur connectivité avec les autres utilisateurs. Des réponses possibles sont alors de concevoir des interfaces à faible consommation (ce qui est coûteux) ou de développer des politiques d'économie d'énergie (par exemple, avec la mise en place de cycles de service pour allumer et éteindre une interface de l'appareil). Notons que l'implantation de cette dernière approche nécessite a priori un accord opérateur/utilisateur, sachant que la politique de confidentialité empêche l'opé-

rateur d'avoir des droits sur l'appareil de l'utilisateur.

— **Direction 5.** En ce qui concerne l'indemnisation des utilisateurs, nous ne devrons pas oublier que la coopération des utilisateurs dans l'approche opportuniste est essentielle. En particulier, les utilisateurs doivent accepter le partage de leurs ressources mobiles (p. ex. batterie, stockage, interfaces réseaux actives). Le problème majeur qui se pose alors est un problème de sécurité et de confidentialité, étant donné que les utilisateurs acceptent rarement de recevoir ou stocker des données provenant d'un appareil étranger. Le défi sera donc de proposer une infrastructure de confiance qui fournira un degré de sécurité rassurant pleinement les utilisateurs.

— **Direction 6.** Notre contribution cible clairement une problématique orientée contenu. Le paradigme de communication CCN (*Content Centric Networking*) peut s'avérer adapté à la problématique de la dissémination de contenu en milieu cellulaire. CCN exploite des caches pour rendre le contenu disponible à plusieurs endroits du réseau. En tirant parti de la transmission multi-sauts, ces caches pourront contribuer à la diffusion efficace des contenus. La simplicité de CCN à demander un contenu pourra également simplifier notre tâche pour proposer un système d'adressage approprié.

Bibliographie

[1] URL http://www.3gpp.org/release-13.

[2] Luxembourg national institute of statistic and economic studies. URL http://www.statistiques.public.lu.

[3] URL http://www.verizon.com/about/news/vzw/2014/01/lte-multicast-verizon-power-house.

[4] ETSI TS 126 346. Multimedia broadcast/multicast service (MBMS) ; protocols and codecs. Technical report, 2012.

[5] 3rd Generation Partnership Project. Technical Specification Group Radio Access Networks : 3G Home NodeB Study Item Technical Report (Release 8), 2008.

[6] Fuad M Abinader, Erika PL Almeida, Fabiano S Chaves, André M Cavalcante, Robson D Vieira, Rafael CD Paiva, Angilberto M Sobrinho, Sayantan Choudhury, Esa Tuomaala, Klaus Doppler, et al. Enabling the coexistence of LTE and WiFi in unlicensed bands. *IEEE Communications Magazine*, 52(11) :54–61, 2014.

[7] Joon Ahn, Maheswaran Sathiamoorthy, Bhaskar Krishnamachari, Fan Bai, and Lin Zhang. Optimizing content dissemination in vehicular networks with radio heterogeneity. *IEEE Transactions on Mobile Computing*, 13(6) :1312–1325, 2014.

[8] Adnan Aijaz, Hamid Aghvami, and Mojdeh Amani. A survey on mobile data offloading : technical and business perspectives. *IEEE Wireless Communications*, 20(2) :104–112, 2013.

[9] Jeffrey G Andrews, Holger Claussen, Mischa Dohler, Sundeep Rangan, and Mark C Reed. Femtocells : Past, present, and future. *IEEE Journal on Selected Areas in Communications*, 30(3) :497–508, 2012.

[10] ANR. DataTweet - un service public de communications opportunistes, 2013-2017. URL http://www.agence-nationale-recherche.fr.

[11] Arash Asadi and Vincenzo Mancuso. On the compound impact of opportunistic scheduling and D2D communications in cellular networks. In *Proceedings of the 16th ACM international conference on Modeling, analysis & simulation of wireless and mobile systems*, pages 279–288. ACM, 2013.

[12] Arash Asadi and Vincenzo Mancuso. Energy efficient opportunistic uplink packet forwarding in hybrid wireless networks. In *Proceedings of the fourth international conference on Future energy systems*, pages 261–262. ACM, 2013.

[13] Arash Asadi, Qing Wang, and Vincenzo Mancuso. A survey on device-to-device communication in cellular networks. *IEEE Communications Surveys & Tutorials*, 16(4) : 1801–1819, 2014.

[14] Franz Aurenhammer. Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Computing Surveys (CSUR)*, 23(3) :345–405, 1991.

[15] Aruna Balasubramanian, Ratul Mahajan, and Arun Venkataramani. Augmenting mobile 3G using WiFi. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*, pages 209–222. ACM, 2010.

[16] Marco Valerio Barbera, Julinda Stefa, Aline Carneiro Viana, Marcelo Dias de Amorim, and Mathias Boc. VIP delegation : Enabling vips to offload data in wireless social mobile networks. In *2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS)*, pages 1–8. IEEE, 2011.

[17] Eyuphan Bulut and Boleslaw K Szymanski. WiFi access point deployment for efficient mobile data offloading. *ACM SIGMOBILE Mobile Computing and Communications Review*, 17(1) :71–78, 2013.

[18] Han Cai, Irem Koprulu, and Ness B Shroff. Exploiting double opportunities for deadline based content propagation in wireless networks. In *INFOCOM, 2013 Proceedings IEEE*, pages 764–772. IEEE, 2013.

[19] Lin X Cai, Xuemin Sherman Shen, Jon W Mark, Lin Cai, and Yang Xiao. Voice capacity analysis of WLAN with unbalanced traffic. *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*, 55(3), 2006.

[20] Andre M Cavalcante, Erika Almeida, Robson D Vieira, Sayantan Choudhury, Esa Tuomaala, Klaus Doppler, Fabiano Chaves, Rafael CD Paiva, and Fuad Abinader. Performance evaluation of LTE and WiFi coexistence in unlicensed bands. In *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*, pages 1–6. IEEE, 2013.

[21] Prabhu Chandhar and Suvra Sekhar Das. Analytical evaluation of offloading gain in macrocell-femtocell OFDMA networks. In *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*, pages 1–6. IEEE, 2013.

[22] Vikram Chandrasekhar and Jeffrey G Andrews. Uplink capacity and interference avoidance for two-tier cellular networks. In *IEEE GLOBECOM 2007-IEEE Global Telecommunications Conference*, pages 3322–3326. IEEE, 2007.

[23] Vikram Chandrasekhar, Jeffrey G Andrews, and Alan Gatherer. Femtocell networks : a survey. *IEEE Communications magazine*, 46(9) :59–67, 2008.

[24] Vikram Chandrasekhar, Jeffrey G Andrews, Tarik Muharemovic, Zukang Shen, and Alan Gatherer. Power control in two-tier femtocell networks. *IEEE Transactions on Wireless Communications*, 8(8) :4316–4328, 2009.

[25] Aleksandra Checko, Henrik L Christiansen, Ying Yan, Lara Scolari, Georgios Kardaras, Michael S Berger, and Lars Dittmann. Cloud RAN for mobile networks — a technology overview. *IEEE Communications surveys & tutorials*, 17(1) :405–426, 2015.

[26] Qimei Chen, Guanding Yu, Amine Maaref, Geoffrey Li, and Aiping Huang. Rethinking mobile data offloading for LTE in unlicensed spectrum. *IEEE Transactions on Wireless Communications*, 15(7) :4987–5000, 2016.

[27] Qimei Chen, Guanding Yu, Hangguan Shan, Amine Maaref, Geoffrey Ye Li, and Aiping Huang. Cellular meets WiFi : Traffic offloading or resource sharing ? *IEEE Transactions on Wireless Communications*, 15(5) :3354–3367, 2016.

[28] Vasek Chvatal. A greedy heuristic for the set-covering problem. *Mathematics of Operations Research*, 4(3) :233–235, 1979.

[29] Cisco. Cisco visual networking index : Forecast and methodology, 2015–2020 white paper, 2016.

[30] Cisco. Cisco visual networking index : Global mobile data traffic forecast update, 2015 2020 white paper. Technical report, 2016.

[31] Holger Claussen. Performance of macro-and co-channel femtocells in a hierarchical cell structure. In *2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–5. IEEE, 2007.

[32] Holger Claussen and Doru Calin. Macrocell offloading benefits in joint macro-and femtocell deployments. In *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 350–354. IEEE, 2009.

[33] Lara Codeca, Raphaël Frank, and Thomas Engel. Luxembourg SUMO traffic (LuST) scenario : 24 hours of mobility for vehicular networking research. In *Vehicular Networking Conference (VNC), 2015 IEEE*, pages 1–8. IEEE, 2015.

[34] Alejandro de la Fuente, Ana Garcia Armada, and Raquel Pérez Leal. Joint multicast/unicast scheduling with dynamic optimization for LTE multicast service. In *European Wireless 2014 ; 20th European Wireless Conference ; Proceedings of*, pages 1–6. VDE, 2014.

[35] Savio Dimatteo, Pan Hui, Bo Han, and Victor OK Li. Cellular traffic offloading through WiFi networks. In *2011 IEEE Eighth International Conference on Mobile Ad-Hoc and Sensor Systems*, pages 192–201. IEEE, 2011.

[36] Irit Dinur and Samuel Safra. On the hardness of approximating minimum vertex cover. *Annals of mathematics*, pages 439–485, 2005.

[37] Klaus Doppler, Mika Rinne, Carl Wijting, Cássio B Ribeiro, and Klaus Hugl. Device-to-device communication as an underlay to LTE-advanced networks. *IEEE Communications Magazine*, 47(12) :42–49, 2009.

[38] Klaus Doppler, Mika P Rinne, Pekka Janis, Cssio Ribeiro, and Klaus Hugl. Device-to-device communications ; functional prospects for LTE-advanced networks. In *2009 IEEE International Conference on Communications Workshops*, pages 1–6. IEEE, 2009.

[39] Hesham ElSawy, Ekram Hossain, and Dong In Kim. Hetnets with cognitive small cells : user offloading and distributed channel access techniques. *Communications Magazine, IEEE*, 51(6) :28–36, 2013.

[40] Ericsson. The ericsson mobility report, 2016.

[41] Jakob Eriksson, Hari Balakrishnan, and Samuel Madden. Cabernet : vehicular content delivery using WiFi. In *Proceedings of the 14th ACM international conference on Mobile computing and networking*, pages 199–210. ACM, 2008.

[42] Patrick Th Eugster, Pascal A Felber, Rachid Guerraoui, and Anne-Marie Kermarrec. The many faces of publish/subscribe. *ACM Computing Surveys (CSUR)*, 35(2) :114–131, 2003.

[43] Kevin Fall and Stephen Farrell. DTN : an architectural retrospective. *IEEE Journal on Selected Areas in Communications*, 26(5) :828–836, 2008.

[44] Richard H Frenkiel and Tomasz Imielinski. Infostations : The joy of many-time many-where communications. *Journal on Mobile Computing*, 1996.

[45] Richard H Frenkiel, BR Badrinath, Joan Borras, and Roy D Yates. The infostations challenge : Balancing cost and ubiquity in delivering wireless data. *IEEE Personal Communications*, 2000.

[46] Laura Galluccio, Giacomo Morabito, Aris L Moustakas, and Sergio Palazzo. Opportunistic communications in infostation systems : Delay and stability analysis. In *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, pages 1–5. IEEE, 2010.

[47] Ivan Ganchev, Stanimir Stojanov, Mairtin O'Droma, and Damien Meere. An infostation-based university campus system for the provision of mlearning services. In *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)*, pages 195–199. IEEE, 2006.

[48] Ivan Ganchev, Stanimir Stojanov, Mairtin O'Droma, and Damien Meere. Adaptable infostation-based mlecture service provision within a university campus. In *Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007)*, pages 165–169. IEEE, 2007.

[49] Silvia Giordano et al. Mobile ad hoc networks. *Handbook of wireless networks and mobile computing*, pages 325–346, 2002.

[50] Bo Han, Pan Hui, VS Anil Kumar, Madhav V Marathe, Jianhua Shao, and Aravind Srinivasan. Mobile data offloading through opportunistic communications and social participation. *IEEE Transactions on Mobile Computing*, 11(5) :821–834, 2012.

[51] Liang Hu, Claudio Coletti, Nguyen Huan, István Z Kovács, Benny Vejlgaard, Ralf Irmer, and Neil Scully. Realistic indoor WiFi and femto deployment study as the offloading solution to lte macro networks. In *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, pages 1–6. IEEE, 2012.

[52] Ana Lucia Iacono and Christopher Rose. Bounds on file delivery delay in an infostations system. In *Vehicular Technology Conference Proceedings, 2000. VTC 2000-Spring Tokyo. 2000 IEEE 51st*, pages 2295–2299. IEEE, 2000.

[53] LK Saliya Jayasinghe, Praneeth Jayasinghe, Nandana Rajatheva, and Matti Latva-aho. MIMO physical layer network coding based underlay device-to-device communication. In *PIMRC*, pages 89–94, 2013.

[54] Jeongho Jeon, Qian Clara Li, Huaning Niu, Apostolos Papathanassiou, and Geng Wu. Lte in the unlicensed spectrum : A novel coexistence analysis with wlan systems. In *2014 IEEE Global Communications Conference*, pages 3459–3464. IEEE, 2014.

[55] Richard M Karp. Reducibility among combinatorial problems. In *Complexity of computer computations*, pages 85–103. Springer, 1972.

[56] Brett Kaufman and Behnaam Aazhang. Cellular networks with an overlaid device to device network. In *2008 42nd Asilomar Conference on Signals, Systems and Computers*, pages 1537–1541. IEEE, 2008.

[57] Duk Kyung Kim and Dan Keun Sung. Characterization of soft handoff in CDMA systems. *IEEE Transactions on Vehicular Technology*, 48(4) :1195–1202, 1999.

[58] Troels Kolding, Pawel Ochal, Przemyslaw Czerepinski, and Klaus Pedersen. Impact of carrier configuration and allocation scheme on 3G femtocell offload effect. In *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, pages 1–5. IEEE, 2011.

[59] Timo Koskela, Sami Hakola, Tao Chen, and Janne Lehtomaki. Clustering concept using device-to-device communication in cellular system. In *2010 IEEE Wireless Communication and Networking Conference*, pages 1–6. IEEE, 2010.

[60] Daniel Krajzewicz and Christian Rossel. Simulation of urban mobility (SUMO). *Centre for Applied Informatics (ZAIK) and the Institute of Transport Research at the German Aerospace Centre*, 2007.

[61] Vineet Kumar, Lan Lin, Daniel Krajzewicz, Fatma Hrizi, Oscar Martinez, Javier Gozalvez, and Ramon Bauza. itetris : Adaptation of its technologies for large scale integrated simulation. In *Proceedings of the 71st Vehicular Technology Conference (VTC 2010-Spring)*, pages 1–5. IEEE, 2010.

[62] Kyunghan Lee, Joohyun Lee, Yung Yi, Injong Rhee, and Song Chong. Mobile data offloading : how much can WiFi deliver ? In *Proceedings of the 6th International COnference*, page 26. ACM, 2010.

[63] Kyunghan Lee, Joohyun Lee, Yung Yi, Injong Rhee, and Song Chong. Mobile data offloading : How much can WiFi deliver ? *IEEE/ACM Transactions on Networking (TON)*, 21(2) :536–550, 2013.

[64] Xu Li, Chunming Qiao, Yunfei Hou, Yunjie Zhao, Aditya Wagh, Adel Sadek, Liusheng Huang, and Hongli Xu. On-road ads delivery scheduling and bandwidth allocation in vehicular CPS. In *INFOCOM, 2013 Proceedings IEEE*, pages 2571–2579. IEEE, 2013.

[65] Yong Li, Guolong Su, Pan Hui, Depeng Jin, Li Su, and Lieguang Zeng. Multiple mobile data offloading through delay tolerant networks. In *Proceedings of the 6th ACM workshop on Challenged networks*, pages 43–48. ACM, 2011.

[66] Ken Lo. Mobile internet : How many gigabytes do you need ?, 2015. URL http://kenstechtips.com/.

[67] David López-Pérez, Alvaro Valcarce, Guillaume De La Roche, Enjie Liu, and Jie Zhang. Access methods to WiMAX femtocells : A downlink system-level case study. In *Communication Systems, 2008. ICCS 2008. 11th IEEE Singapore International Conference on*, pages 1657–1662. IEEE, 2008.

[68] Francesco Malandrino, Claudio Casetti, Carla-Fabiana Chiasserini, and Marco Fiore. Offloading cellular networks through ITS content download. In *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2012 9th Annual IEEE Communications Society Conference on*, pages 263–271. IEEE, 2012.

[69] Christoph P Mayer and Oliver P Waldhorst. Offloading infrastructure using delay tolerant networks and assurance of delivery. In *Wireless Days (WD), 2011 IFIP*, pages 1–7. IEEE, 2011.

[70] Steven McCanne. Scalable multimedia communication using IP multicast and lightweight sessions. *IEEE Internet Computing*, 3(2) :33–45, 1999.

[71] Damien Meere, Ivan Ganchev, Máirtín O'Dróma, Stanimir Stojanov, and Veselina Valkanova. Adaptation for assimilation : Shaping context-sensitive m-learning services within a multi-agent environment. In *Telecommunications (AICT), 2010 Sixth Advanced International Conference on*, pages 74–79. IEEE, 2010.

[72] Damien Meere, Ivan Ganchev, and Mairtin O'Droma. Utilising a multi-agent system to support the deployment of contextualised mlearning services. In *2012 6th IEEE International Conference Intelligent Systems*, pages 088–094. IEEE, 2012.

[73] Soumen Mitra, Srinjoy Chattopadhyay, and Suvra Sekhar Das. Deployment considerations for mobile data offloading in LTE-femtocell networks. In *2014 International Conference on Signal Processing and Communications (SPCOM)*, pages 1–6. IEEE, 2014.

[74] Nokia. Nokia LTE for unlisenced spectrum, White Paper, 2014.

[75] PE Omiyi and H Haas. Maximising spectral efficiency in 3G with hybrid ad-hoc UTRA TDD/UTRA FDD cellular mobile communications. In *Spread Spectrum Techniques and Applications, 2004 IEEE Eighth International Symposium on*, pages 613–617. IEEE, 2004.

[76] Bo Peng, Tao Peng, Ziyang Liu, Yufeng Yang, and Chunjing Hu. Cluster-based multicast transmission for device-to-device (D2D) communication. In *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th*, pages 1–5. IEEE, 2013.

[77] Tao Peng, Qianxi Lu, Haiming Wang, Shaoyi Xu, and Wenbo Wang. Interference avoidance mechanisms in the hybrid cellular and device-to-device systems. In *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 617–621. IEEE, 2009.

[78] Ayaskant Rath and Shivendra Panwar. Fast handover in cellular networks with femtocells. In *Communications (ICC), 2012 IEEE International Conference on*, pages 2752–2757. IEEE, 2012.

[79] Filippo Rebecchi, Marcelo Dias de Amorim, and Vania Conan. DROiD : Adapting to individual mobility pays off in mobile data offloading. In *Proceedings of the IFIP Networking 2014 Conference*, pages 1–9. IFIP, 2014.

[80] Filippo Rebecchi, Marcelo Dias De Amorim, Vania Conan, Andrea Passarella, Raffaele Bruno, and Marco Conti. Data offloading techniques in cellular networks : a survey. *IEEE Communications Surveys & Tutorials*, 17(2) :580–603, 2015.

[81] Filippo Rebecchi, Marcelo Dias de Amorim, and Vania Conan. Should I seed or should I not : On the remuneration of seeders in D2D offloading. In *IEEE WoWMoM 2016 Seventeenth International Symposium on a World of Wireless, Mobile and Multimedia Networks*, 2016.

[82] Nikodin Ristanovic, Jean-Yves Le Boudec, Augustin Chaintreau, and Vijay Erramilli. Energy efficient offloading of 3G networks. In *2011 IEEE Eighth International Conference on Mobile Ad-Hoc and Sensor Systems*, pages 202–211. IEEE, 2011.

[83] Durga P Satapathy and Jon M Peha. Performance of unlicensed devices with a spectrum etiquette. In *Global Telecommunications Conference, 1997. GLOBECOM'97., IEEE*, volume 1, pages 414–418. IEEE, 1997.

[84] Jaakko Seppälä, Timo Koskela, Tao Chen, and Sami Hakola. Network controlled device-to-device (D2D) and cluster multicast concept for LTE and LTE-A networks. In *2011 IEEE Wireless Communications and Networking Conference*, pages 986–991. IEEE, 2011.

[85] Petr Slavík. A tight analysis of the greedy algorithm for set cover. In *Proceedings of the 28th ACM Symposium on Theory of Computing (STOC)*, pages 435–441. ACM, 1996.

[86] Giuseppe Sollazzo, Mirco Musolesi, and Cecilia Mascolo. TACO-DTN : a time-aware content-based dissemination system for delay tolerant networks. In *Proceedings of the 1st international MobiSys workshop on Mobile opportunistic networking*, pages 83–90. ACM, 2007.

[87] Vladimir Vukadinović and Gunnar Karlsson. Spectral efficiency of mobility-assisted podcasting in cellular networks. In *Proceedings of the Second International Workshop on Mobile Opportunistic Networking*, pages 51–57. ACM, 2010.

[88] Jerry Chun-Ping Wang, Hossam ElGindy, and Justin Lipman. On cache prefetching strategies for integrated infostation-cellular network. In *Proceedings. 2006 31st IEEE Conference on Local Computer Networks*, pages 185–192. IEEE, 2006.

[89] Ning Wang and Jie Wu. Opportunistic WiFi offloading in a vehicular environment : Waiting or downloading now? In *Proc. of the 35th IEEE International Conference on Computer Communications (IEEE INFOCOM 2016)*, 2016.

[90] John Whitbeck, Yoann Lopez, JéRéMie Leguay, Vania Conan, and Marcelo Dias De Amorim. Push-and-track : Saving infrastructure bandwidth through opportunistic forwarding. *Pervasive and Mobile Computing*, 8(5) :682–697, 2012.

[91] Hongyi Wu, Chunming Qiao, Swades De, and Ozan Tonguz. Integrated cellular and ad hoc relaying systems : iCAR. In *IEEE Journal on Selected Areas in Communications*. IEEE, 2001.

[92] Edward Wyatt. AT&T accused of deceiving smartphone customers with unlimited data plans, 2014. URL *http://www.nytimes.com/*.

[93] Yongping Xiong, Rui Tian, Jian Ma, Yan Liu, and Limin Sun. Optimal infostation deployment for spatio-temporal information dissemination. In *Computers and Communications (ISCC), 2010 IEEE Symposium on*, pages 120–122. IEEE, 2010.

[94] Mehmet Yavuz, Farhad Meshkati, Sanjiv Nanda, Akhilesh Pokhariyal, Nick Johnson, Balaji Raghothaman, and Andy Richardson. Interference management and performance analysis of UMTS/HSPA+ femtocells. *IEEE Communications Magazine*, 47(9) : 102–109, 2009.

[95] Shu-ping Yeh, Shilpa Talwar, Seong-choon Lee, and Heechang Kim. WiMAX femtocells : a perspective on network architecture, capacity, and coverage. *IEEE Communications Magazine*, 46(10) :58–65, 2008.

[96] Shu-Ping Yeh, Shilpa Talwar, Geng Wu, Nageen Himayat, and Kerstin Johnsson. Capacity and coverage enhancement in heterogeneous networks. *Wireless Communications, IEEE*, 18(3) :32–38, 2011.

[97] Seokhoon Yoon, Duc T Ha, Hung Q Ngo, and Chunming Qiao. MoPADS : A mobility profile aided file downloading service in vehicular networks. *IEEE Transactions on Vehicular Technology*, 58(9) :5235–5246, 2009.

[98] C-H Yu, Olav Tirkkonen, Klaus Doppler, and Cássio Ribeiro. Power optimization of device-to-device communication underlaying cellular communication. In *2009 IEEE International Conference on Communications*, pages 1–5. IEEE, 2009.

[99] Chia-Hao Yu, Klaus Doppler, Cassio B Ribeiro, and Olav Tirkkonen. Resource sharing optimization for device-to-device communication underlaying cellular networks. *IEEE Transactions on Wireless Communications*, 10(8) :2752–2763, 2011.

[100] Ran Zhang, Miao Wang, Lin X Cai, Zhongming Zheng, Xuemin Shen, and Liang-Liang Xie. LTE-unlicensed : the future of spectrum aggregation for cellular networks. *IEEE Wireless Communications*, 22(3) :150–159, 2015.

[101] Rongqing Zhang, Xiang Cheng, Liuqing Yang, and Bingli Jiao. Interference-aware graph based resource sharing for device-to-device communications underlaying cellular networks. In *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 140–145. IEEE, 2013.

[102] Bin Zhou, Honglin Hu, Sheng-Qiang Huang, and Hsiao-Hwa Chen. Intracluster device-to-device relay algorithm with optimal resource utilization. *IEEE transactions on vehicular technology*, 62(5) :2315–2326, 2013.

[103] Bin Zhou, Saisai Ma, Jing Xu, and Zhenhong Li. Group-wise channel sensing and resource pre-allocation for LTE D2D on ISM band. In *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 118–122. IEEE, 2013.