Infrastructure and device-to-device cellular data offloading

Vinicius Fernandes Soares Mota

Universidade Federal de Minas Gerais

Instituto de Ciências Exatas
Programa de Pós Graduação em Ciência da Computação

In International Joint Ph.D with

Université Paris-EST

École Doctorale MSTIC
Mathématiques et Sciences et Technologies de l'Information et de la Communication

PHD THESIS

Mention: Informatique

defended by

Vinícius Fernandes Soares Mota

Infrastructure and Device-to-Device Cellular Data Offloading

Advisors
José Marcos Silva Nogueira
Yacine Ghamri-Doudane

Jury:

M. Marcelo Dias de Amorim	Université Pierre et Marie Curie	Rapporteur
M. Célio Vinícius Albuquerque	Universidade Federal Fluminense	Rapporteur
M. Daniel Fernandes Macedo	Universidade Federal de Minas Gerais	Examinateur
Mdm. Lynda Mokdad	Université Paris-Est Cretéil	Examinateur
M. Yacine Ghamri-Doudane	Université de la Rochelle	Directeur de thèse
M. José Marcos Silva Nogueira	Universidade Federal de Minas Gerais	Directeur de thèse

Décembre, 2015

VINíCIUS FERNANDES SOARES MOTA

DESCARGA DE DADOS INFRA-ESTRUTURADA E OPORTUNÍSTICA EM REDES DE CELULAR

Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Doutor em Ciência da Computação.

Orientador: José Marcos Silva Nogueira
Yacine Ghamri-Doudane
Coorientador: Daniel Fernandes Macedo

Belo Horizonte

Dezembro de 2015

VINíCIUS FERNANDES SOARES MOTA

INFRASTRUCTURE AND DEVICE-TO-DEVICE CELLULAR DATA OFFLOADING

Thesis presented to the Graduate Program in Computer Science of the Federal University of Minas Gerais in partial fulfillment of the requirements for the degree of Doctor in Computer Science.

Advisors: José Marcos Silva Nogueira
Yacine Ghamri-Doudane
Co-Advisor: Daniel Fernandes Macedo

Belo Horizonte
December 2015

Mota, Vinícius Fernandes Soares

D1234p Infrastructure and Device-to-Device Cellular Data Offloading / Vinícius Fernandes Soares Mota. — Belo Horizonte, 2015
xxxv, 122 f. : il. ; 29cm

Tese (doutorado) — Federal University of Minas Gerais
Orientador: José Marcos Silva Nogueira

Yacine Ghamri-Doudane

1. Computação — Teses. 2. Redes — Teses. I. Orientador. II. Título.

CDU 519.6*82.10

[Folha de Aprovação]

Quando a secretaria do Curso fornecer esta folha,
ela deve ser digitalizada e armazenada no disco em formato gráfico.

Se você estiver usando o `pdflatex`,
armazene o arquivo preferencialmente em formato PNG
(o formato JPEG é pior neste caso).

Se você estiver usando o `latex` (não o `pdflatex`),
terá que converter o arquivo gráfico para o formato EPS.

Em seguida, acrescente a opção `approval=`{*nome do arquivo*}
ao comando \`ppgccufmg`.

Se a imagem da folha de aprovação precisar ser ajustada, use:
`approval=[`*ajuste*`][`*escala*`]`{*nome do arquivo*}
onde *ajuste* uma distância para deslocar a imagem para baixo
e *escala* é um fator de escala para a imagem. Por exemplo:
`approval=[-2cm][0.9]`{*nome do arquivo*}
desloca a imagem 2cm para cima e a escala em 90%.

I dedicate this work to my grandma and my mother.

Acknowledgments

To my family and their patience with me and my stress.

To Professor José Marcos for all enlightening talks and his comprehension to come up against difficulties during all steps of my PhD. Professor Daniel Macedo, for our long hours of discussions about networking, protocols and beers.

I would like to thank the Graduate Program in Computer Science (PPGCC) of UFMG, for providing a pleasant environment for work and the Department of Computer Science, where I met several friends. Among these friends, I highlight the "Cabrals' crew", with whom it has been a pleasure to talk about "philosophical issues" almost every Thursday nights in the last 7 years. Furthermore, I have the pleasure to work at the Winet lab, where brilliant minds helped me to elaborate this work. Winet team, thank you guys.

I would like to thank my advisor at the Université Paris-Est, Professor Yacine Ghamri-Doudane, for all the helpful discussions and contributions. Moreover, a big thank-you for the LIGM team, in which I could learn not only about networking and computer science, but also about cultures (and "*le vrai français*").

This work would not be possible without the financial support of FAPEMIG, CNPq, CAPES and CEMIG.

Besides the technical and financial support from those mentioned above, I'm grateful for those who indirectly helped me.

Although she is a pharmacist, in the last two years she motivated me every day to advance this dissertation. One of the key reasons for my increased productivity! Thanks a lot, Cris Giuberti :)

A special thanks to *República Notre Dame* and all my friends from Ouro Preto, my official resting place, for all those crazy moments.

I must mention my friends from Contagem, my hometown, with whom I have been sharing 30-years of friendship.

And finally, thank you to all those helped me to overcome this challenging PhD Course and to organize these ideas in any way.

"Find something you love to do, and you'll never have to work a day in your life."
(Unknown)

Abstract

This thesis addresses the overload problem of the Wireless Internet service Providers' (WISP) network. The growth of mobile broadband subscription has been leading several bottlenecks to WISPs, such as, bandwidth availability and resource sharing over a single cellular cell. WISPs can move off data traffic from its infrastructure by deploying small cells, such as femtocells, to public WiFi networks or, more recently, to device-to-device opportunistic networks. This work evaluates the feasibility to offload mobile data traffic using WiFi hotspots, proposes a framework to opportunistic data offloading and incentive mechanisms to encourage users cooperation.

We mapped 3G and WiFi coverage through several bus routes in Paris in order to evaluate how users and WISPs can benefit from the existing infrastructure. Our results indicate that the deployed WISPs access points can offload part of the data traffic, however restrictions such as association time and the authentication process may reduce the amount of offloaded data.

We propose a multi-criteria decision-making framework, called OppLite, to offload data from 3G networks using opportunistic device-to-device communications. Trace-driven simulations showed that opportunistic mobile offloading can expand coverage and network efficiency, offloading up to 36% of data in certain scenarios. Thus, the effectiveness of opportunistic mobile offloading depends mainly of the delay tolerance of the applications and whether the user cooperates.

Since opportunistic offloading depends on the user's willingness to offer his/her resources to others, we propose a message-based incentive mechanism that builds a reputation rank based on the source of messages received by the forwarding nodes, called MINEIRO. The network supports up to 60% of nodes with selfish behavior without performance degradation in a random mobility scenario. After this threshold, MINEIRO kept the delivery rate and the delay constant. Meanwhile, in a scenario with social-based mobility, selfish behavior degrades the network performance quickly.

Palavras-chave: Mobile Data offloading, Opportunistic Networking, cooperation.

Résumé

Cette thèse aborde le problème de la surcharge des réseaux des fournisseurs de service Internet sans fil. La croissance de l'abonnement au haut débit mobile a été conduit plusieurs goulots d'étranglement à WISP, tels que la disponibilité de la bande passante et le partage des ressources sur une station de base réceptrice, en raison de contraintes de spectre. Ainsi, les fournisseurs essaient de déplacer le trafic de données de son infrastructure en déployant de petites cellules tels que les femtocells ou réseaux WiFi publics ou, plus récemment, par la mise en place de réseaux opportunistes qui transmettent les données de dispositif à dispositif (les réseaux opportunistes).

Ce travail évalue la faisabilité de décharger le trafic de données mobile à l'aide des hotspots WiFi, propose un cadre pour le déchargement de données opportuniste et un mécanisme d'incitation pour encourager la coopération des utilisateurs.

Nous avons tracé la couverture 3G et WiFi à travers de plusieurs lignes de bus à Paris afin d'évaluer la façon dont les utilisateurs et les fournisseurs d'Internet peuvent bénéficier de l'infrastructure existante. Nos résultats indiquent que les points d'accès WiFI déployés par les fournisseurs de service Internet peuvent décharger une partie du trafic de données, cependant des restrictions telles que le temps de l'association et le processus d'authentification peuvent diminuer la quantité de données transmises.

Dans une tentative d'offrir une nouvelle approche pour le déchargement mobile, nous proposons un cadre de décision multi-critères, appelé OppLite, pour décharger les données de réseaux 3G grâce à des communications dispositif à dispositif opportunistes. Nous avons montré par des simulations que le déchargement mobile opportuniste dispositif à dispositif peut étendre la couverture et l'efficacité des réseaux cellulaires, déchargement jusqu'à 36% des données dans certains scénarios. Ainsi, l'efficacité de déchargement mobile par les réseaux opportunistes dépend principalement de la tolérance de délai par l'application et de coopération de l'utilisateurs dans le réseau.

Une fois que le déchargement opportuniste dépend de la volonté de l'utilisateur d'offrir ses ressources aux autres, nous avons proposé un mécanisme d'incitation qui construit un rang de réputation basée sur la source des messages reçus par les nœuds

de intermediaires, appelé MINEIRO. Le réseau prend en charge jusqu'à 60% de nœuds avec un comportement égoïste sans dégradation des performances dans un scénario de mobilité aléatoire. Après ce seuil, MINEIRO a maintenu le taux de livraison et le retard constante. Pendant ce temps, dans un scénario avec une mobilité basée sociale, du comportement égoïste dégrade rapidement les performances du réseau.

Les sections suivantes résument chacune des contributions. Elles suivent la structure de la thèse, afin que le lecteur puisse identifier les contributions de chaque chapitre de ce manuscrit.

Chapitre 2: Notions de Base

Ce chapitre présente les notions de base nécessaires à la compréhension de cette thèse de doctorat. Nous décrivons les concepts de base de l'Internet mobile à large bande, les réseaux opportunistes, la théorie des jeux et la théorie utilitaire.

Les technologies fournissant l'Internet mobile à large bande ont évolué ces dernières années. L'avance principale concerne le débit de données offert par le réseau, qui a évolué de quelques kilobits par seconde à des gigabits par seconde au cours des années. Les réseaux de troisième génération (3G), qui sont définis par l'Union International de Télécommunications (UIT), offrent un débit d'au moins 200 Kbits/s et jusqu'à 84 Mbps. De nos jours, les réseaux 3G sont remplacés par la quatrième génération, aussi appelée 4G, qui fournit des taux maximaux de 100 Mbit/s pour la communication en grande mobilité et 1 Gbit/s dans les scénarios en mobilité réduite. Aujourd'hui, les universités et l'industrie travaillent sur la cinquième génération de réseaux mobiles (5G), qui vise à améliorer l'évolutivité en termes de coût, d'économie d'énergie et de ressources.

Nous avons utilisé les réseaux opportunistes (OppNet) comme une solution pour soulager le trafic de données du réseau cellulaire. Les OppNets sont des réseaux qui, contrairement aux réseaux classiques, sont sujettes à des déconnexions fréquentes et des hauts retards de communication. De plus, les OppNets se caractérisent par l'utilisation du paradigme *store-carry-forward*, où les messages sont stockés dans la mémoire secondaire (p. ex. disques durs et cartes-éclair), et ces messages sont envoyées à chaque fois qu'un lien de communication est établi. Nous avons proposé une taxonomie pour classer les scénarios d'application qui utilisent les concepts d'OppNets. Parmi ces applications, nous mettons en évidence les réseaux d'appareils portatifs (*Pocket Switch Networks* – PoSNet), formés par des personnes portant des appareils portatifs (par exemple smartphones, assistants numériques personnels), fournissant ainsi une communication point-à-point. Un PoSNet peut supporter des applications sociales, des

jeux et en plus, aider les réseaux infrastructurés. Dans ce cas, les appareils mobiles interagissent de manière opportuniste pour réduire la charge subie par l'infrastructure.

La *théorie d'utilité* a été utilisée pour décider quand et qui devrait passer de communication en mode infrastructuré à la communication en mode opportuniste. Cette théorie quantifie l'ensemble des préférences d'un client dans une échelle numérique. L'utilité d'un bien ou d'un service (nommée x) peut être désignée comme une fonction mathématique $u(x)$.

Enfin, comme le propriétaire d'un dispositif mobile peut être égoïste et refuser le transfert de messages de tiers à l'aide de son appareil, nous avons utilisé la théorie des jeux pour modeler le comportement de l'utilisateur. La théorie des jeux a pour but de remédier des situations dont la récompense d'un participant du jeu est affectée par sa décision, ainsi que les décisions prises par d'autres intervenants qui sont en interaction avec lui. Les participants du jeu sont appelés *joeurs*, tandis que leurs décisions sont connues comme *stratégies*. Un joueur reçoit un *gain* basé sur la stratégie choisie par ce dernier et les stratégies choisies par les autres joueurs. Deux concepts fondamentaux en théorie des jeux sont la *Meilleure réponse* et *l 'équilibre de Nash*. La *meilleure réponse* est la stratégie, choisie par un joueur, qui maximise son gain indépendamment de la stratégie choisie par les autres joueurs. L'équilibre de *Nash* est atteint lorsque tous les joueurs choisissent leurs meilleures réponses.

Chapitre 3: Travaux connexes

Dans ce chapitre, nous avons procédé à une recherche systématique sur les travaux connexes afin de présenter les techniques de l'état-de-l'art dans le déchargement de l'infrastructure cellulaire et dans les mécanismes de motivation proposés pour engager la coopération des utilisateurs. Nous divisons notre recherche systématique en trois étapes: tout d'abord, nous avons effectué une revue profonde de la littérature sur les réseaux opportunistes. Ensuite, nous avons réduit notre recherche visant à mettre l'accent sur des propositions de déchargement dans les réseaux cellulaires. Enfin, nous avons examiné l'état-de-l'art sur la coopération et les mécanismes d'encouragement pour motiver les utilisateurs à adopter le déchargement opportuniste.

Trois approches peuvent être trouvées dans la littérature pour le déchargement de données mobiles: Femtocell, le déchargement assisté par réseaux WiFi et le déchargement dispositif-à-dispositif opportuniste.

Femtocells sont des petites stations de base, à faible puissance, connectées au réseau par une liaison filaire. Elles ont comme but d'améliorer la couverture cellulaire à l'intérieur des résidences ou bureaux. L'interférence résultant du grand déploiement

des femtocells avec l'infrastructure cellulaire est le problème le plus important de cette technologie.

L'étude de la disponibilité des réseaux WiFi dans les villes permet que les WISPs (*Wireless Internet Service Provider*) estiment combien ils peuvent décharger de leur infrastructure par les réseaux WiFi. Les recherces antérieurs montrent que les villes sont largement couvertes par les points d'accès WiFi. Toutefois, la majorité de ces points d'accès ne fournit pas d'accès au public.

Concernant l'utilisation du femtocell ou réseaux WiFi afin de décharger le trafic, les fournisseurs sont susceptibles d'adopter les deux solutions. Ainsi, le déchargement WiFi et les femtocells sont des solutions complémentaires pour le déchargement du trafic de données. Néanmoins, la performance du réseau WiFi se dégrade alors que la densité des points d'accès augmente.

Ensuite, nous avons présenté les travaux qui utilisent le déchargement opportuniste pour soulager la congestion du trafic de donnés sur les réseaux 3G en utilisant les dispositifs des utilisateurs. En utilisant le déchargement opportuniste, les dispositifs mobiles de l'utilisateur peuvent également jouer le rôle d'un *relay* entre autres dispositifs et l'infrastructure.

Enfin, nous avons présenté les travaux qui abordent les mécanismes de coopération et d'encouragement. Les mécanismes pour engager la coopération des êtres humains ont été étudiés et encore, il n'y a pas de solution qui garantit l'engagement d'une façon efficace. Dans le contexte des réseaux des données, plusieurs mécanismes ont été proposés pour les réseaux P2P, les réseaux ad hoc, réseaux opportunistes et, récemment pour les scénarios de déchargement opportunistes. Ceux-ci proposent de la punition des utilisateurs égoïstes au paiement des utilisateurs coopératifs.

Chapitre 4: La faisabilité de déchargement sur WiFi

Ce chapitre caractérise la connectivité WiFi et 3G à Paris. Notre principal objectif était d'évaluer le potentiel du déchargement WiFi dans les lignes d'autobus de la ville, à l'aide des points d'accès (Aps) déployées par les opérateurs (ici appelés WISPs du à son nom en anglais). Nous tentons de répondre à la question suivante : *Est-il possible pour les WISPs de décharger les données par l'infrastructure WiFi déjà déployé dans les villes?*

Afin d'atteindre notre objectif, nous avons mis en œuvre une Application Android et nous avons mesuré la couverture 3G et WiFi à travers de plusieurs itinéraires de bus à Paris.

Catégorisation de connectivité

Nous effectuons une analyse de la couverture WiFi, séparant les points d'accès publics fournis par le gouvernement, ceux fournis par des WISPs et finalement les points d'accès privés.

Les points d'accès publics fournis par le gouvernement ont été recensés par une base de données publique. Cette base de données liste 312 points d'accès répartis sur la ville. Pendant ce temps, nous avons implémenté une application pour les appareils Android pour collecter et consigner les informations sur la disponibilité du réseau 3G et du WiFi dans les rues. Nous avons recensés 21,649 APs, parmi eux 55.4% étant des WISPs. Les trois principaux WISPs représentent 27,9, 15,85 et 11.6%, respectivement des APs obtenus.

Nous avons obtenu une couverture 3G à 90% du temps, tandis que la couverture WiFi par les feux follets atteint jusqu'à 99% du temps dans les 82 km d'itinéraires. En fait, nous avons constaté une moyenne de 27,5 points d'accès, avec un minimum d'un et maximum de 66 points d'accès. Nous observons qu'en moyenne, le client mobile est capable de scanner un point d'accès sur une distance de 60m et 90% d'APs sont découvrables au-dessus de 130m. Ensuite, nous montrons que 27% des APs ne nécessitent pas d'authentification de couche de liaison (par exemple WPA2-PSK). Cependant, 57% d'entre eux sont des APs des WISP, ce qui signifie qu'ils sont exclusifs pour les clients et l'étape d'authentification sera effectuée au sein d'un *proxy*.

Déchargement de données générées par l'utilisateur

Afin de déterminer la quantité de données que les WISPs pourraient décharger par les réseaux WiFi déployés, nous avons comparé les points d'accès publics de Paris et les points d'accès obtenus dans notre expérience avec des milliers de données générés par l'utilisateur dans un service de partage de localisation.

Nous supposons que si un utilisateur est plus proche qu'une certaine distance d'un point d'accès, le client mobile peut utiliser le WiFi au lieu du réseau 3G. Nous avons examiné toutes les points d'accès comme étant publics, ce qui représente t le maximum de déchargement de données. Après, nous avons observé la capacité de déchargent des données des points d'accès WiFi de chaque WISP.

Seulement les points d'accès publics du gouvernement pourraient décharger jusqu'à 10% du trafic total. Dans le meilleur cas, si tous les points d'accès étaient ouverts et disponibles pour les clients mobiles, il serait possible de décharger de 10 à 30% du trafic.

Puisqu'un réseau WiFi ouvert n'est pas une réalité, nous montrons que les WISPs jouent un rôle important dans le déchargement par WiFi. Ils peuvent décharger près de 30% du trafic sur l'ensemble de données évaluée. La couverture WiFi par les WISPS

est supérieure à celle des points d'accès gouvernementaux. Ainsi, les WISPs devraient offrir des incitations à leurs clients pour utiliser leurs points d'accès.

Enfin, comme le temps et la distance qu'un dispositif reste dans le rayon d'un AP est faible, même à basse vitesse, nous croyons que le principal goulot d'étranglement pour l'efficacité du déchargement par WiFi est le temps excessif requis pour l'authentification et l'association dans ces réseaux.

Par ailleurs, dans des environnements surpeuplés, par exemple dans les événements populaires, même le WiFi n'est pas capable de supporter des milliers d'utilisateurs connectés au même temps. Dans le chapitre suivant, nous proposons une solution à l'aide de la communication opportuniste, c'est-à-dire la communication dispositif à dispositif, pour décharger le réseau 3G.

Chapitre 5 : Un cadre pour déchargement opportunistes des données mobiles

Les solutions existantes de déchargement exigent des routeurs spéciaux ou nouveaux déploiements de routeurs, qui impliquent des changements importants dans l'infrastructure. Nous nous concentrons sur le déchargement s'appuyant sur la communication opportuniste de dispositif à dispositif. Cela permet le déchargement à bas cout à l'operateur, car il ne demande aucune infrastructure supplémentaire.

Dans ce chapitre, nous décrivons *OppLite*, notre cadre de prise de décision multicritère basée sur la théorie de l'utilité, qui permet de basculer entre les modes infrastructuré et opportuniste basées uniquement sur des décisions locales effectuées dans les dispositifs. *OppLite* utilise le nombre de voisins, la durée de vie de la batterie et la puissance du signal comme critères pour prendre la décision d'envoyer un message directement à l'infrastructure ou à l'aide de la communication opportuniste.

La solution proposée

OppLite surveille le réseau autour de l'utilisateur, mesurant des informations telles que le nombre de voisins. Le *User profile* définit le poids de chaque information capturée dans la décision et la tolérance de retard pour les applications en cours d'exécution.

Après avoir recueilli des informations du réseau, la gestion de *communication-mode module* applique une fonction utilitaire pour chacun des critères observés et agrège les résultats de toutes les fonctions de l'utilitaire. Dans le mode *standard*, un périphérique peut passer à *Relay* ou *opportuniste* mode en fonction des résultats de l'agrégation des fonctions utilitaires.

Nous avons utilisé une fonction d'utilité sigmoïde pour évaluer chaque critère.

Nous avons supposé une fonction utilitaire $u(x) \in [0, 1]$, où x c'est le critère évalué. *OppLite* fait remarquer les critères suivant: nombre de voisins, qualité de lien et vie de batterie.

Les nœuds prennent des décisions parmi trois alternatives : devenir un nœud *relay*, devenir un nœud *opportunistic* ou rester un nœud *standard*. En considérant la fonction d'utilité multicritères, le module de gestion de mode de communication décide si le dispositif doit basculer réseaux opportunistes, déchargeant ainsi l'infrastructure.

Ensuite, nous avons décrit comment *OppLite* gère deux décisions, lorsqu'un nœud doit devenir un nœud *relay*; et quand un nœud doit devenir un nœud *Opportunistic*. En raison des exigences de délai des demandes, *OppLite* définit un seuil de délai maximal pour la communication opportuniste , retourner sur le mode *standard* lorsqu'un message non remis atteint ce seuil de retard.

Chapitre 6: L'Évaluation de L'OppLite

Dans ce chapitre, nous avons présenté et discuté des évaluations approfondies de l'OppLite. Nous avons proposé trois applications pour évaluer L'OppLite : *i*) *Opportunistic Relaying* (OpR), les nœuds opportunistes qui transmettent leurs messages aux infrastructures à travers les nœuds de transmission ; *Cache-and-Forward ii)* (CaF), les nœuds de transmission qui transfèrent tout le contenu reçu aux nœuds opportunistes ; et *iii*) *Relay as Cache* (RaF), nœuds opportunistes qui cherchent le contenu dans le cache des autres nœuds de transmission.

Toutes les applications ont été analysées à l'aide de deux traits : INFOCOM et ROLLERNET. INFOCOM représente un scénario de conférence et ROLLERNET représente un groupe qui fait du roller à Paris. Nous avons les caractérisé et montré qu'ils ont de différentes caractéristiques de connectivité. INFOCOM présente un graphique stable ayant peu des composantes connexes avec plusieurs nœuds, tandis que ROLLERNET a des graphiques dynamiques ayant des moments de composants plus connectés avec quelques nœuds.

Dans l'application *OpR*, 1/3 des nœuds en mode de transmission peuvent transférer jusqu'à 45% du trafic dans un scénario de conférence avec un retard de moins de 20 minutes. Dans un scénario qui représente le mouvement de la foule dans les rues de Paris, OppLite transmet jusqu'à 70% du trafic avec 42 nœuds parmi 98 nœuds en mode de transmission. Ainsi, OppLite réduit le nombre de nœuds qui utilisent l'infrastructure.

Dans l'application *CaF*, OppLite peut décharger jusqu'à 80% du trafic des données dans l'INFOCOM et 94% dans le ROLLERNET avec un délai de 20 minutes.

Si les nœuds en mode de transmission ne transfèrent les messages qu'aux alentours, ne tolérant qu'un retard d'une seconde, OppLite peut décharger jusqu'à 30% et 52% du trafic des données dans l'INFOCOM et dans ROLLERNET, respectivement.

Ces résultats dans l'application *CaF*, sont issus du fait que les nœuds de transmission transfèrent tout message reçu à tous les contacts de nœuds opportunistes lors d'un retard configuré. Puisque les nœuds opportunistes sont passifs, les nœuds de transmission qui ont plus de voisinage dans le mode opportuniste atteignent des proportions supérieures de déchargement. Différemment de L'Algorithme *Random*, OppLite commute les nœuds avec un plus grand voisinage en mode de transmission, ce qui provoque un meilleur déchargement. En outre, l'application *Cache-and-Forward* ne considère que la réplication de la transmission au mode *one-hop*. Étendre *CaF* pour permettre l'envoi de messages au mode *n-hop*, ce qui permettrait d'améliorer le taux de déchargement du réseau.

Dans l'application *RaC*, OppLite décharge de 32 à 44% du trafic de données lorsque les nœuds opportunistes tolèrent 20 minutes de retard dans l'ensemble de données de l'INFOCOM. En cas de nœuds opportunistes, n'attendez qu'une seule seconde jusqu'à recevoir une réponse d'un relais, OppLite décharge de 5 à 35% du trafic. En raison des caractéristiques du trait ROLLERNET, le taux de déchargement variait de 13 à 17% avec 10 minutes de retard. Dans les scénarios de tolérance de retard plus faibles, OppLite décharge jusqu'à 7% des données du trafic dans le ROLLERNET. Vu que les nœuds demandent le contenu à partir d'un ensemble fini, la quantité de contenu disponible à la demande affecte la performance de l'OppLite.

Dans toutes les applications évaluées, OppLite attend la coopération de l'utilisateur de changer son dispositif pour les modes de transmission ou opportuniste. Bien que l'OppLite permette aux utilisateurs de configurer leur volonté de coopération, il n'est pas sûr si l'utilisateur serait coopératif. Des mécanismes de motivation peuvent gérer ces questions qui seront discutées dans le chapitre suivant.

Chapitre 7: Vers les mécanismes de motivation

Dans ce chapitre, nous nous sommes concentrés dans la question suivante: *Comment peut-on motiver les utilisateurs à partager leurs ressources avec les autres?* Nous avons proposé un mécanisme de motivation pour les réseaux opportunistes génériques qui classifie les nœuds basés sur les messages qu'ils envoient. Ensuite, nous avons mis en contexte la coopération de réseaux sans fil, en mettant l'accent sur le déchargement de données mobiles opportunistes.

Le mécanisme de motivation proposé

Notre motivation était de fournir un mécanisme de motivation visant à augmenter l'envie de l'utilisateur de transmettre des messages à d'autres utilisateurs.

Nous avons proposé un mécanisme de motivation, appelé MINEIRO - ***M****essage-based* ***IN****centive mechanism for* ***E****nd-user* ***I****mprovement of* ***R****outing* ***O****pportunities* en réseau opportuniste. MINEIRO construit une réputation bien classée basée sur la source des messages reçus par les nœuds de transmission.

MINEIRO classifie les nœuds d'acheminement selon la source des messages qu'ils envoient, punissant ainsi les nœuds qui ne transmettent que leurs propres messages. Par conséquent, si les nœuds souhaitent augmenter leurs chances d'avoir leurs messages livrés à la destination, ils doivent donc faire parvenir des messages des autres nœuds.

En plus, nous avons modelé MINEIRO comme un jeu Bayésien et nous avons montré les conditions dans lesquelles le profil stratégique de coopération, autrement dit, la transmission des messages de tiers, conduit à un équilibre Bayésien.

Nous avons évalué MINEIRO dans un environnement de simulation et nous avons montré que MINEIRO encourage les utilisateurs à continuer à transmettre les messages de tiers.

Mécanismes de motivation pour le Déchargement Opportuniste

Vers les mécanismes de motivation permettant le déchargement opportuniste, nous avons montré comment le WISP peut fournir des récompenses pour la transmission coopérative dans un coût limite supérieur. Notre mécanisme de motivation basé en récompense des données indique clairement les avantages pour les nœuds de transmission et les opportunistes, tout en évitant une grande augmentation de coûts pour le WISP. WISPs ont besoin d'examiner les coûts pour fournir des récompenses, du point de vue de la transmission, il faut considérer la limite du plan de données et les avantages de coopérer. Du point de vue de l'opportuniste, il faut considérer combien de temps il peut attendre pour obtenir un contenu souhaité.

Par ailleurs, OppLite peut être facilement étendu pour supporter les nouveaux critères qui encouragent les utilisateurs à agir comme nœuds de transmission ou opportunistes. Nous avons montré une intégration entre MINEIRO et l'OppLite d'une façon centralisée et décentralisée. En supposant que les utilisateurs sont rationnels, ils collaborent quand ils ont assez de ressources nécessaires pour gagner une réputation positive ; ils en utilisent comme un nœud opportuniste pour sauver leurs ressources ou améliorer la bande large lorsque cela se fait nécessaire.

Chapitre 8: Conclusions et Recherches futures

Finalement, nous avons conclu ce travail en résumant nos contributions et don-

nant des directions futures de la recherche. Nos trois contributions majeures sont : une étude de la faisabilité du déchargement par l'infrastructure WiFi ; un cadre pour le déchargement de l'infrastructure moins opportuniste; et vu que le réseau opportuniste demande la coopération des utilisateurs, nous avons développé des mécanismes de motivation pour encourager des comportements coopératifs.

Chaque contribution exige une enquête profonde pour affiner les résultats présentés dans cette thèse. Nous avons divisé les futures recherches dans trois étapes:

WiFi infrastructure de déchargement Nous avons pour objectif la publication de l'application développée pour recueillir des informations sur les points d'accès WiFi comme une ressource ouverte. Par conséquent, nous pouvons recruter plus de bénévoles pour recueillir des données dans plusieurs villes. De là, la même analyse peut être faite dans d'autres villes avec des propriétés différentes comme la taille de la population, Produit Intérieur Brut (PIB) ou d'autres conditions de trafic.

Déchargement des données mobile opportuniste Étendre notre cadre de déchargement des données mobiles opportunistes pour permettre une Auto-Configuration Dynamique, Nouveaux Critères et des nouvelles applications en plus de l'OppLite. Par ailleurs, nous envisageons d'étudier les questions sur la confidentialité des données et la sécurité concernant le déchargement opportuniste. Enfin, mettre en place un banc d'essai pour comprendre les limitations pratiques du déchargement opportuniste de dispositif à dispositif.

Au-delà des défis techniques inhérentes pour engager la coopération, convaincre les utilisateurs à changer leurs dispositifs en mode de transmission ou opportuniste apporte aussi des défis sociaux et psychologiques. Un mécanisme de motivation est efficace s'il recrute plus de participants et maintient ceux-ci ayant un comportement coopératif pour plus long temps.

Pour l'efficacité des mécanismes de motivation qui récompensent, il est nécessaire de considérer les coûts pour les WISPs (ou toute autre entité centrale), ainsi comme le coût pour les participants. Cependant, il s'agit de trouver et décider, une valeur qui minimise le coût pour le WISP et au même temps, qui motive l'utilisateur à exiger d'autres enquêtes.

Des expériences réelles peuvent expliquer le comportement de l'utilisateur. Nous avons l'intention de mettre en place et de déployer l'OppLite intégré avec MINEIRO pour analyser et caractériser les éléments de motivation tels que la réputation, les récompenses, entre autres.

List of Figures

1.1 The network model. 6
1.2 A scenario of opportunistic offloading: Bob acts as relay, other devices can leave the infrastructure network and to use the relay node opportunistically. 7
1.3 A scenario of opportunistic mobile data offloading: Bob acts as relay and other devices request the content opportunistically from the relay node. . . 8

2.1 Taxonomy of Opportunistic Networks. 16
2.2 Protocol stack of a DTN node. 17

3.1 A classification of incentive mechanisms. 35

4.1 Concentration of APs in Paris (200 x 200m grid cells). Darker cells indicate more hotspots available in a region. 43
4.2 Distribution of access points through the measured bus routes. 45
4.3 WiFi connectivity Properties. 46
4.4 Histogram of access points found in each scan operation. 46
4.5 WiFi Graph for each WISP. WISP A shows a high connected component, while WISP B and C show a graph more sparse. 48
4.6 Distribution of public hotspots in Paris 49
4.7 Amount of data offloaded using the entire database. 51
4.8 Amount of data offloaded for the different periods. 51

5.1 OppLite framework forwarding algorithm for signaling offloading. 58
5.2 Architecture of OppLite framework. 58
5.3 Behavior of different values in Equation 5.6. 61
5.4 Variation of $U_r(X)$ 63

6.1 Distribution of inter-message generation time. 70

6.2 Properties of the traces used in our evaluation. The black squares show the number of connected components (left y-axis), while the lines show the average number of nodes in the connected components and in the dominating set (right y-axis). 72
6.3 Average number of nodes in relay and opportunistic mode based on threshold variation in OppLite. 74
6.4 Comparison of average relay nodes between Random and OppLite. 74
6.5 OpR: Offloaded messages based on the Relay Threshold in INFOCOM dataset 75
6.6 OpR: Offloaded messages based on the Relay Threshold variation in ROLLERNET dataset 75
6.7 Cache-and-Forward data offloading: Relay Threshold impact in INFOCOM dataset 78
6.8 Cache-and-Forward data offloading: Random vs OppLite in INFOCOM dataset. 78
6.9 Cache-and-Forward INFOCOM: Average messages forwarded by relay nodes. 79
6.10 Cache-and-Forward data offloading: Relay Threshold impact in ROLLERNET dataset 79
6.11 Cache-and-Forward data offloading: Random vs OppLite in ROLLERNET dataset. 80
6.12 Cache-and-Forward ROLLERNET: Average messages forwarded by each relay node. 80
6.13 Relay-as-Cache: Hit Ratio performance. Left column shows results for INFOCOM trace while right column shows hit ratio for Rollernet. This figure aggregates all sizes of the content pool. 84
6.14 Relay-as-Cache INFOCOM dataset: Data offloading. 85
6.15 Relay-as-Cache ROLLERNET dataset: Data offloading. 86

7.1 An example of transmission in an ad hoc network. Node B considers the cost to receive and to relay the message x. 90
7.2 Non-forwarding selfish behavior in RandomWay mobility model. 96
7.3 Dropping message selfish behavior in RandomWay mobility model. 97
7.4 Non-forwarding selfish behavior in swim mobility model. 98
7.5 Dropping message selfish behavior in swim mobility model. 99
7.6 Actors and their behaviors. Opportunistic offloading relies on the assumption that a set of users will cooperate with others. 100
7.7 Reward received by the relay when forwarding data of size $|k| = 100$. . . . 103
7.8 Distributed and centralized approaches to integrate OppLite and MINEIRO 104

List of Tables

1.1 Rate Plan for major WISPs in US, Brazil and France. 3

2.1 Aggregators Utility function for two scenarios. 22

3.1 Opportunistic mobile offloading proposals 31
3.2 Incentive mechanisms for Opportunistic mobile offloading 38

4.1 Summary of access point and 3G information. 44
4.2 Summary of the properties of the AP graph from each WISP. 47
4.3 Percentage of time connected in each type of cellular Network 49
4.4 Number of users and *check-ins* in each period 50

5.1 OppLite Parameters. 57
5.2 Best fit Parameters for power level in off-the-shelf devices [Huang et al., 2012]. 60
5.3 Criteria values. 63

6.1 Criteria Parameters 69
6.2 Average content requests for each delay tolerance. 81
6.3 Cacheability of the network in all analyzed scenarios. 83

7.1 Payoff's Matrices. 93
7.2 Benefit-cost for each node behavior in opportunistic offloading. 100

List of Acronyms

AP Access Point

CaF Cache and Forwarding application

D2D Device-to-Device communication

DTN Disruption/Delay Tolerant Networking

MANET Mobile Ad-hoc Networking

MWSNs Mobile Wireless Sensor Networks

OppNet Opportunistic Networking

OpR Opportunistic Relaying application

PoSNet Pocket Switched Networking

PSN Participatory Sensing Networking

RaC Relay as Cache application

LTE Long Term Evolution

VANETs Vehicular Ad Hoc Networks

WISP Wireless Internet Service Providers

Contents

Acknowledgments **xi**

Abstract **xv**

Résumé **xvii**

List of Figures **xxvii**

List of Tables **xxix**

List of Acronyms **xxxi**

1 Introduction **1**

1.1 Motivation 1
1.2 Definitions 3
1.3 Problem 4
1.4 Use Cases 7
1.5 Contributions 8
1.6 Document Organization 11

2 Fundamentals **13**

2.1 Mobile Broadband Internet 13
2.2 Opportunistic Networking 14
2.2.1 Types of Opportunistic Networks 16
2.3 Game Theory 19
2.3.1 Basics 19
2.4 Utility Theory 20
2.4.1 Definition 20
2.4.2 Utility Function: Single and Multi Criteria 21
2.5 Conclusion 22

3 Related Work **25**
3.1 Methodology . . . 25
3.2 Femtocell Offloading . . . 27
3.3 WiFi availability and Offloading . . . 28
3.4 Opportunistic Mobile Data Offloading . . . 29
3.4.1 Selecting Best Relay Candidates . . . 29
3.4.2 3GPP Device-to-Device Proximity Services . . . 31
3.5 Selfishness, Cooperation and Incentive Mechanisms . . . 33
3.5.1 Classes of Incentive Mechanisms . . . 35
3.6 Conclusions . . . 39

4 On the Feasibility of WiFi Offloading **41**
4.1 Connectivity Categorization . . . 42
4.1.1 Public WiFi Hotspot . . . 42
4.1.2 WISP and Private Hotspots . . . 42
4.1.3 WiFi Connectivity . . . 44
4.1.4 3G Connectivity . . . 48
4.1.5 Lessons Learned and Discussion . . . 48
4.2 Offloading User Generated Data . . . 49
4.2.1 User Localization Database . . . 50
4.2.2 Evaluation . . . 50
4.2.3 Towards WiFi Offloading . . . 52
4.3 Conclusion . . . 53

5 OppLite: An Opportunistic Mobile Data Offloading Framework **55**
5.1 Motivation . . . 55
5.2 OppLite Framework . . . 56
5.2.1 OppLite Framework . . . 56
5.2.2 Opportunistic criteria . . . 58
5.2.3 User Profile . . . 60
5.2.4 Utility Function for Single Criterion . . . 60
5.2.5 Multi-Criteria Aggregation Function . . . 61
5.2.6 Decision Algorithm . . . 62
5.3 Conclusions . . . 63

6 OppLite Evaluation **65**
6.1 Application Scenarios . . . 65
6.2 Methodology . . . 67

6.2.1 Simulations . . . 68
6.2.2 Parameters . . . 69
6.2.3 Traffic Model . . . 69
6.2.4 Content Request Pattern . . . 70
6.2.5 Traces . . . 70
6.3 Trace Analysis . . . 71
6.4 Simulation Results . . . 73
6.4.1 Amount of Relays and Opportunistic Nodes . . . 73
6.4.2 Opportunistic Relaying (OpR) Evaluation . . . 74
6.4.3 Cache-and-Forward (CaF) Evaluation . . . 77
6.4.4 Relay-as-Cache (RaC) Evaluation . . . 81
6.5 Conclusions . . . 87

7 Towards Incentive Mechanisms for Opportunistic Mobile Data Offloading 89
7.1 Incentive Mechanism for Opportunistic Forwarding . . . 89
7.1.1 MINEIRO - Reciprocity based Incentive Mechanism . . . 90
7.1.2 MINEIRO as a Bayesian Game . . . 92
7.1.3 MINEIRO Evaluation . . . 94
7.2 Engaging Cooperation in Opportunistic Offloading . . . 97
7.2.1 Data Reward to Engage Users' Cooperation . . . 99
7.2.2 User Centric Incentive Mechanisms . . . 103
7.3 Conclusions . . . 106

8 Conclusions and Future Work 107
8.1 Conclusions . . . 107
8.2 Future Work and Research Perspective . . . 109
8.2.1 WiFi Infrastructured Offloading . . . 110
8.2.2 Opportunistic Mobile Offloading . . . 110
8.2.3 Incentive Mechanisms Evaluation . . . 112

Bibliography 113

Chapter 1

Introduction

The evolution on Internet technologies has enabled new applications, such as VoIP and video streaming. In addition to that, recently, smartphones, tablets and other portable devices with communication capabilities have become very popular. These devices often have powerful storage and processing capabilities, as well as offer support to modern communication technologies such WiFi and LTE networks, also known as 4G. At the same time, these new devices and applications demand more bandwidth, which raise new challenges to the *Wireless Internet Service Providers* (WISPs).

Indeed, the number of mobile Internet users is increasing in the entire world. At the same time, the bandwidth demand increases exponentially each year [Cisco, 2015]. WISPs have been facing bottlenecks in the bandwidth available for users and to support several simultaneous users over a single cellular cell, due to spectrum constraints.

The purpose of this study is to investigate solutions to alleviate the traffic load on cellular networks. Mobile data offloading occurs when data traffic is moved off from the cellular network to other communication technologies [Han et al., 2012].

Industry and academia are attempting to offload data traffic from 3G networks through the use of femtocells [Haldar et al., 2013], public WiFi [Lee et al., 2010] and, more recently, device-to-device opportunistic networks [Han et al., 2012][3GPPP, 2013]. In this work, we focus on the already deployed WiFi access points in the cities and opportunistic device-to-device networking to offload the cellular network.

1.1 Motivation

Smartphones, tablets and other gadgets with communication capabilities have become very popular nowadays. In fact, the traffic generated by mobile Internet was about 885 Petabytes (10^{15} bytes) per month in 2012 [Cisco, 2013] while in 2014, mobile data

traffic grew up to 2.5 Exabytes (10^{18} bytes) per month [Cisco, 2015]. With the advent of faster mobile networks, the average data traffic per mobile device is expected to be more than 4 GB per month in 2019 [Cisco, 2015]. It is clear that WISPs will face significant challenges in the attempt to follow this increase in bandwidth consumption and radio resources.

In scenarios such as football matches, big concerts or other places with high concentration of mobile clients, WISPs antennas may be overloaded, which causes an interruption of communication. Increasing the number of 3G antennas does not solve the problem, since the main issue is the congested spectrum in metropolitan areas, where only signaling already consumes a huge amount of network resources [Balasubramanian et al., 2009]. Besides the data traffic caused by web applications and video over demand, several mobile applications, e.g. instant messengers, send continuous short messages to provide always-on connectivity, which causes signaling traffic, or so-called signaling storm [Choi et al., 2014]. In these cases, customers could migrate to other means of communication instead of using cellular networks.

In fact, a report from CISCO networks estimates that 46% of all mobile traffic will be offloaded through WiFi in 2017 [Cisco, 2013]. Since WiFi availability is larger than 3G or 4G networks, and some WISPs offer free WiFi hotspots for their customers throughout the city, it is expected that users are willing to migrate to those networks.

Furthermore, a work group in 3GPP tries to standardize the Device-to-Device (D2D) communication paradigm using the same spectrum of LTE networks [3GPPP, 2013]. Thus, end-users devices can help operators to improve cell coverage, spectrum efficiency and to offload signaling and data traffic when acting as *relays* for other users, which will carry the data to less congested antennas [Hui et al., 2005]. However, this method depends on the user's willingness to share their already scarce resources.

In addition to the technical aspects, WISPs shall consider carefully their business model to attend the demand. The common billing model used to offer mobile Internet is based on a flat rate or a limited data plan as shown in the Table 1.1. This table shows the most expensive data plans from WISPs at United States, Brazil and France[1]. Considering this plan pricing, the gross income of the WISPs is fixed by the number of users.

WISPs need to raise their prices (or number of customers) or to reduce the costs to increase their profit. Thus, we argue that opportunistic mobile data offloading can, at least, decrease costs to provide mobile broadband Internet for customers, since it avoids infrastructure deployment.

[1]These plans pricing were quoted on 1st June 2015.

Table 1.1. Rate Plan for major WISPs in US, Brazil and France.

		Monthly Price (USD)	Data Allowance
Verizon	US	75	2GB
At &T		425	50GB
Sprint		60	Unlimited
T-Mobile		80	Unlimited
Vivo	BR	138	8GB
Tim		63	50GB
Claro		76	7GB
Oi		28	2GB
Orange	FR	76	10GB
SFR		32	5GB
Free		22	20GB

1.2 Definitions

In this section, we give brief definitions of some important terms. Our aim is to avoid ambiguity or misunderstanding with similar terms used along the text.

User, device and node The difference between these terms is subtle. We refer to *user* when talking about an end user behavior or action; *device* is used when considering the communication capabilities and technological aspects of the mobile device. Finally, we use *node* to refer to the topological aspects of the graph formed by the network.

3G, 4G and 5G Networks The third generation (3G) of mobile broadband Internet provides data rates of 2 Mbps, while LTE and WiMAX networks, known as the fourth generation (4G) networks, improved the data rate up to 100Mbps with high mobility. The 5th generation mobile network, *5G network*, aims to offer Gigabits per second data rate for several users simultaneously and to reduce the latency significantly when compared with 4G [Qiao et al., 2015]. Nowadays, 3G networks are widely adopted in the market and have better coverage, thus, we used the term *3G network* along the text when referring to devices connected to the cellular infrastructure.

Opportunistic, Pocket Switched Network and Device-to-device (D2D) communication Mobile networks where data is forwarded opportunistically are known in the

literature as *opportunistic networks* [Fall, 2003]. These networks are based on direct communication between devices, which forward messages from one device to another when a transmission opportunity (contact) takes place.

Meanwhile, networks formed by mobile devices carried by people are called *Pocket Switched Networks* (PoSNets) (a subset of opportunistic networks) [Hui et al., 2005].

Recently, PoSNets are being referred as *Device-to-Device communication* (D2D) by the 3^{rd} Generation Partnership Project (3GPP) within the 5G network context [3GPPP, 2013]. Two important aspects characterize PoSNet and D2D communication: *i*) it is formed by devices with high storage and processing capacity - however, with limited energy and usually limited bandwidth; and *ii*) the mobility follows the human mobility pattern. These terms are used interchangeably along the text.

1.3 Problem

The exponential growth of devices using mobile broadband Internet through the cellular infrastructure has become a challenge for industry and academia. This challenge ranges from the capacity of simultaneous devices connected in a single cell, due to overloaded backhaul, to the exponential increase of data traffic in WISPs networks. The main problem addressed by this thesis concerns how to relieve the data traffic or the number of simultaneous users over the cellular infrastructure while maximizing the utilization of WISPs resources.

The solutions proposed by the industry and academia are based on (1) the use of WiFi infrastructure, allowing users to migrate from 3G/4G networks to WiFi, and (2) recently, on device-to-device opportunistic mobile data offloading, where selected devices act as relays to devices disconnected from the 3G network. In the latter, the main problem is how to select relays that maximize the data offload under certain delay constraints.

Narrowing the problem, we focused on the opportunistic mobile data offloading approach, addressing the following questions:

i) Given a set of devices connected to the cellular network and considering the willingness of their owners, how to select and ensure the minimum subset of devices to act as relays for a second subset of devices that will leave the cellular infrastructure and request or forward its messages through the relays?

ii) How to engage users' willingness to become part of these subsets?

Formally, we consider our network as a set $\mathcal{M}$ of user nodes (representing mobile devices) and a set $\mathcal{I}$ of sink nodes (representing the WISP infrastructure, such as 3G). We define the graph $\mathcal{G}_t = (\mathcal{V}, E_t)$ the topology of the network, where $\mathcal{V} = \mathcal{M} \cup \mathcal{I}$ is a set of vertices and E_t is a set of edges representing the links between pairs of nodes in time t.

Let $\mathcal{S} \subset \mathcal{M}$ be the subset of nodes in $\mathcal{M}$ acting as relays. Additionally, let $\mathcal{O} \subset \mathcal{M}$ the nodes without edge with any node from $\mathcal{I}$, representing nodes that decide to disconnect from the infrastructure network.

We consider two types of communication links: *infrastructure communication* and *opportunistic communication.*

We define *infrastructure communication* as the communication mode in which a node m sends data through a link $e_t = (m, i)$, where $m \in \mathcal{M}$ and $i \in \mathcal{I}$.

In *opportunistic communication*, data is forwarded from a node $m \in \mathcal{O}$, and thus to $n \in \mathcal{S}$ through a link $e_t = (m, n)$.

Figure 1.1 shows the described network model. An opportunistic node communicates whenever it has contact with a relay node.

Now, we can define formally question 1 above as following:

Definition 1.3.1. ***Opportunistic Mobile Offloading problem*** *- select a minimum subset of nodes $\mathcal{S}$ that guarantees message delivery under a delay constraint and maximize the number of nodes in the subset $\mathcal{O}$.*

Furthermore, we consider the following restrictions to our problem:

1. Nodes should become relays without any changes on the WISP side.
2. Messages should be delivered under a delay constraint.
3. A node becomes a relay based on the user willingness to offer his/her resources to other nodes.

Concerning the complexity of the problem, Theorem 1.3.1 states that the Opportunistic Mobile Offloading problem is NP-hard:

Theorem 1.3.1. *Selecting a minimum subset $\mathcal{S}$ of devices that attends all devices in the subset $\mathcal{O}$ is* NP*-Hard.*

Proof. Consider the definition below of the minimum dominating set problem, a well-known NP-hard problem.

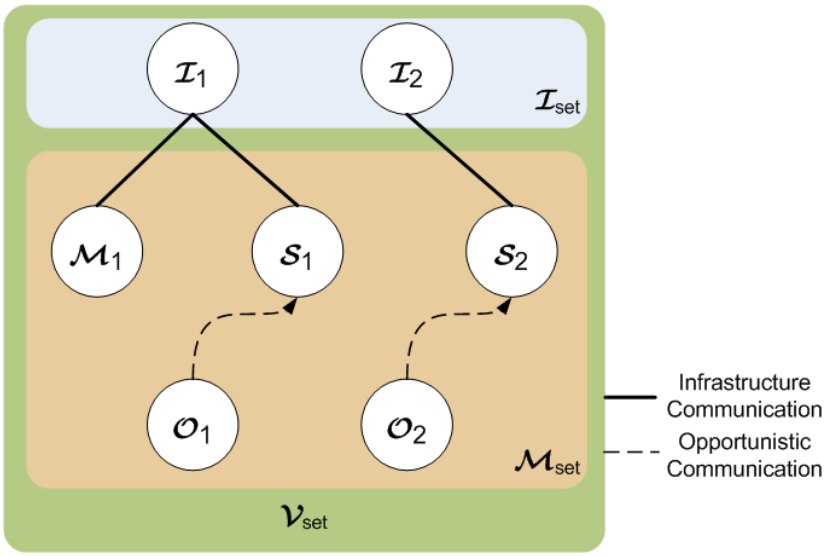


Figure 1.1. The network model.

Definition 1.3.2. ***Minimum Dominating Set problem*** - *The minimum dominating set of a graph $G = (V, A)$ consists of the minimum subset $S \cup V$, where nodes that do not belong to S are incident to at least one node in S.*

Our problem can be modeled as an instance of the minimum dominating set problem.

Let $\mathcal{G}(t) = (\mathcal{M}, E)$ be a snapshot of the graph formed by the network in the instant t. The maximum number of nodes in $\mathcal{O}$ arises only if $\mathcal{S} \subset \mathcal{M}$ is equal to $S \in V$, that is, the minimum dominating set is equal to the relay set. In this case, it is guaranteed that all nodes in $\mathcal{O}$ are incident to a node in $\mathcal{S}$ in time t. Thus, the nodes in $\mathcal{O}$ can send a message to $\mathcal{I}$ through a node in $\mathcal{S}$.

□

However, we argue that just finding the minimum subset may be not feasible for opportunistic Mobile Offloading, based on the following assumptions:

1. Finding the minimum dominating set requires a global view of the topology. We assume that devices are only aware of their neighborhood.

2. if the minimum subset is always chosen, the selected devices may have their batteries drained faster. In this case, the infrastructure would be offloaded, but this would overload the devices.

3. Since users can act selfishly or altruistically towards other users, the willingness of the selected nodes to collaborate with the network should be considered or incentives should be provided.

1.4 Use Cases

We considered two scenarios where mobile offloading improves network resource utilization:

Mobile Signaling Offloading: Relays forward all requests from opportunistic nodes. The gain is twofold: i) Opportunistic nodes leave the infrastructure network, thus, reducing signaling traffic [Choi et al., 2014]. ii) A relay connected to a high speed network can improve performance of opportunistic nodes.

Figure 1.2 illustrates this case. Alice has a device that supports only 3G network connection with 2Mbps data rate, while Bob is under 4G coverage with 100Mbps data rate connection. Bob acts as relay offering his connection to Alice. Since Bob is connected to a high speed network and the connection between them is faster than 3G, Alice increases her data rate and offloads the 3G network.

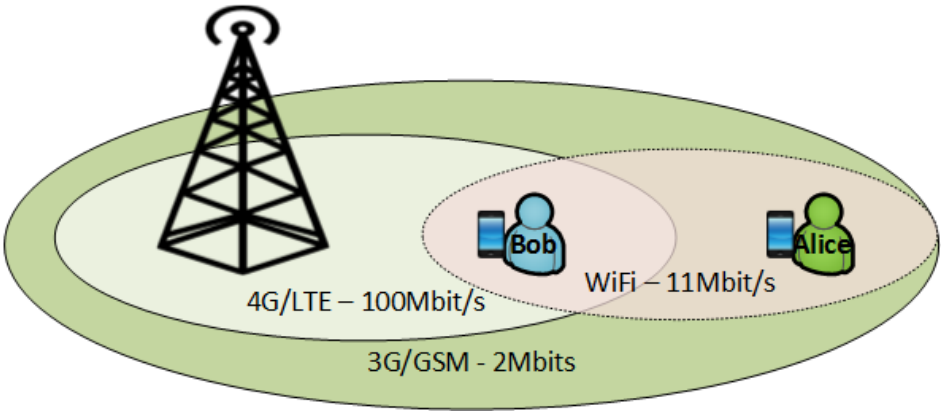


Figure 1.2. A scenario of opportunistic offloading: Bob acts as relay, other devices can leave the infrastructure network and to use the relay node opportunistically.

Mobile Data Offloading: An opportunistic node searches for a content in the relay nodes. The relay owns the content in its buffer with a certain probability. This probability depends on the popularity of the items in a certain region, e.g. several users requesting a goal replay in a football stadium.

Figure 1.3 illustrates this case. Bob acts as a relay and caches the content in his buffer. Other nodes in the range from Bob can request the content from Bob before requesting the content from the cellular network. This, alleviates the data traffic on

the WISP. Although search for content can spend some energy (WiFi or Bluetooth), this energy consumption is lesser than request and receive content through 3G or 4G network.

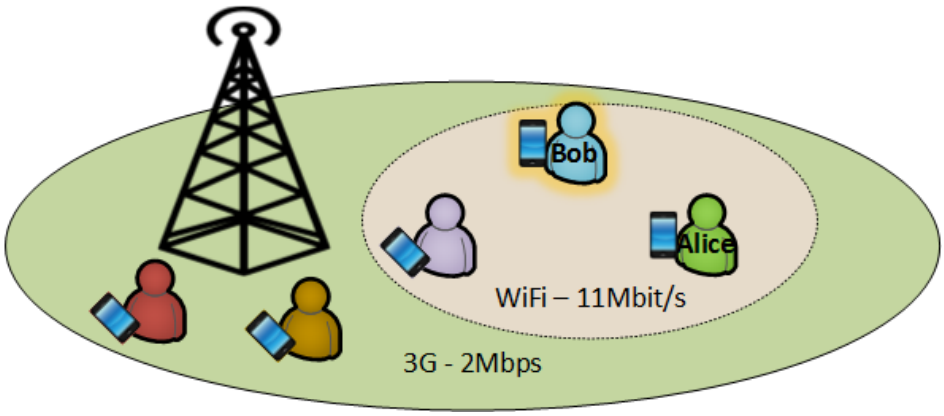


Figure 1.3. A scenario of opportunistic mobile data offloading: Bob acts as relay and other devices request the content opportunistically from the relay node.

1.5 Contributions

The contributions of this work are summarized below:

- We conducted a case study to discuss the feasibility of WiFi offloading in a metropolitan area. We observed that WiFi coverage and deployed WISPs access points can offload part of the data traffic. However, restrictions such as the time required to associate and authenticate in WiFi networks may restrain customers to widely adopt this solution.
- We proposed a framework, called OppLite, to offload the 3G networks using the devices of customers. Instead of using graph heuristics to select devices to act as relays, we applied utility theory to make this decision. Our solution considers the user willingness of becoming a relay and does not require changes in the WISP infrastructure.
- We show that in dense scenarios, opportunistic networking helps to relieve the number of users linked to the 3G network, offloading the data through relays. Furthermore, the node selection algorithm proposed improve data offloading through cache mechanisms when compared to Random solutions.
- We proposed an incentive mechanism called MINEIRO - *Message-based INcentive mechanism for End-user Improvement of Routing Opportunities*, which builds

a reputation rank based on the source of messages received by the forwarding nodes. Furthermore, we proposed two incentive approaches to engage cooperation in opportunistic offloading: *i)* Reward-based, WISPs reward cooperative users raising their data allowance. *ii)* An integration between MINEIRO and OppLite.

The contributions of this Ph. D Thesis were published in the conferences and journals below, where the first two are preliminary works on opportunistic networking:

1. **Electing Clusterheads in Delay Tolerant Networks**. Vinícius F. S. Mota, Daniel F. Macedo, José M. S. Nogueira. XXXI Brazilian Computer Science Conference, in proceedings of Brazilian Symposium on Ubiquitous and pervasive Computer Networks. 2011. In Portuguese.

2. **An Hierarchical Routing Protocol for Opportunistic Emergency Networks**. Vinícius F. S. Mota, Daniel F. Macedo, José M. S. Nogueira. 7th Latin America Networking Conference. 2012.

3. **On the Feasibility of WiFi Offloading in Urban Areas: The Paris Case Study**. Vinícius F. S. Mota, Daniel F. Macedo, Yacine Doudane-Ghamri, José M. S. Nogueira. IFIP/IEEE Wireless Days 2013.

4. **Protocols, Mobility Models and Tools in Opportunistic Networks: A Survey**. Vinícius F. S. Mota, Felipe D. Cunha, Daniel F. Macedo, José M. S. Nogueira, Antonio A.F. Loureiro. Computer Communications, 48, 5-19. Elsevier. 2014

5. **Managing the Decision-Making Process for Opportunistic Mobile Data Offloading**. Vinícius F. S. Mota, Daniel F. Macedo, Yacine Doudane-Ghamri, José M. S. Nogueira. IFIP/IEEE NOMS 2014.

6. **MINEIRO: Um Mecanismo de Incentivo para Aplicações em Redes Oportunísticas**. Vinícius F. S. Mota, Daniel F. Macedo, Yacine Doudane-Ghamri, José M. S. Nogueira. In proceedings of XXXIII Brazilian Networking and Distributed Systems Symposium (SBRC), 2015. In portuguese.

7. **A Message-Based Incentive Mechanism for Opportunistic Networking Applications**. Vinícius F. S. Mota, Daniel F. Macedo, Yacine Doudane-Ghamri, José M. S. Nogueira. In proceedings of 20th IEEE Symposium on Computers and Communication (ISCC), 2015.

Furthermore, I co-authored the following works as part of my interest in ubiquitous computing, wireless sensor networking and incentive mechanisms applied to mobile network.

1. **Overview of Ubicomp Research Based on Scientific Publications**. Thiago H. Silva, Clayson S. F. S. Celes, Vinícius F. S. Mota, Antonio A.F. Loureiro. XXXII Brazilian Computer Science Conference, in proceedings of Brazilian Symposium on Ubiquitous and pervasive Computer. 2012.

2. **A picture of actual Ubicomp research exploring publications from important events in the field**. Thiago H. Silva, Clayson S. F. S. Celes, Vinícius F. S. Mota, Antonio A.F. Loureiro. Journal of Applied Computing Research. 2012.

3. **Uma Avaliação de Abordagens de Distribuição para Gerenciamento de Redes Tolerantes a Atrasos e Desconexões**. Ewerton Salvador, Vinícius Mota, Virgil Almeida, Daniel Fernandes Macedo, José M. S. Nogueira, Jéferson Nobre, Pedro Arthur Duarte, Lisandro Zambenedetti Granville. In proceedings of Fault Tolerance Workshop, 2014.

4. **Real-Time Monitoring of Transmission Lines Using Wireless Sensor Networks**. Jesse L. Leoni, José M. S. Nogueira, Mario F. M. Campos, Daniel F. Macedo, Ewerton M. Salvador, Vinícius F. S. Mota, Daniel B. Resende, Vinícius F. Silva, Luiz H. A. Correia, Luiz F. M. Vieira, Mathias F. Kriebel. IEEE PES Transmission and Distribution Conference and Exposition, 2014.

5. **Redes de Sensoriamento Participativo: Desafios e Oportunidades**. Thiago H. Silva, Pedro O. S. Vaz de Melo, João B. B. Neto, Anna I. J. T. Ribeiro, Clayson S. F. de S. Celes, Vinícius F. S. Mota, Felipe D. da Cunha, Ana P. G. Ferreira, Kássio L. da S. Machado, Raquel A. de F. Mini, Jussara M. Almeida e Antonio A. F. Loureiro. Short course at XXXIII Brazilian Networking and Distributed Systems Symposium (SBRC), 2015. In portuguese.

6. **Users in the Urban Sensing Process: Challenges and Research Opportunities**. Thiago H Silva, Felipe D da Cunha, Anna I J T Ribeiro, João B B Neto, Clayson S F de S Celes, Vinícius F S Mota, Ana P G Ferreira, Pedro O S Vaz de Melo, Jussara M Almeida, Antonio A F Loureiro. Book chapter at Pervasive Computing: Next Generation Platforms for Intelligent Data Collection Morgan Kaufmann / Elsevier, in Book Series "Intelligent Data-Centric Systems", 2015. To appear.

1.6 Document Organization

The rest of this document is organized as follows. Chapter 2 presents a background of the topics addressed throughout the text. It describes technologies for mobile broadband Internet, concepts regarding opportunistic networking and introduces basic concepts of game theory and utility theory. Chapter 3 overviews the related works regarding opportunistic networking, Femtocell, WiFi and opportunistic approaches for mobile data offloading and incentive mechanisms to engage users' cooperation in opportunistic networks.

Chapter 4 analyzes the cellular network and WiFi deployment through a case study in Paris - France and discusses the feasibility to offload cellular network through the already deployed WiFi network.

Chapter 5 describes the multi-criteria decision framework, called OppLite, which defines when nodes become relay or when nodes use opportunistic communication, based only on local information. The performance and evaluation of OppLite are presented in Chapter 6 . This chapter also characterizes the traces used in the simulations to evaluate OppLite.

Chapter 7 proposes a reciprocity based incentive mechanism, called MINEIRO, to engage users to forward messages to other users in opportunistic networking. Furthermore, it discusses two approaches as incentive mechanisms for opportunistic offloading: i) data reward-based mechanism, where WISPs can manage the award value for cooperative users and ii) OppLite integrated with MINEIRO, which ranks cooperative and non-cooperative users.

Finally, Chapter 8 presents the conclusions and future works.

Chapter 2

Fundamentals

This chapter presents the background required for the understanding of this doctoral thesis. Section 2.1 describes technologies for mobile broadband Internet. Section 2.2 presents the concepts regarding opportunistic networking. Basic concepts about game theory are briefly introduced in Section 2.3. Finally, Section 2.4 describes utility theory.

2.1 Mobile Broadband Internet

With the advance of smart devices and multimedia applications, users demand even more bandwidth from their cellular operators. Technologies to provide mobile broadband Internet has been evolving in the last years. The main advance relates to the data rate offered by the network, which has evolved from few kilobits to gigabits in the past years.

The first generation network (1G) introduced the concept of centralized cellular architecture and applies Frequency Division Multiple Access (FDMA) to separate users in the frequency domain. However, 1G provides only voice services. Wireless data communication through mobile phones was introduced by the Global System for Mobile Communications (GSM) standard, the second generation (2G) of mobile networks. 2G networks use Time Division Multiple Access (TDMA) and provide data rates of 9.6 Kbps. General Packet Radio Service (GPRS) and Enhanced Data rates for GSM Evolution (EDGE), known as generation 2.5, enhanced the data rate of GSM networks up to 237 Kbps.

The third generation network (3G), defined by International Mobile Telecommunication (ITU), offers a peak data rate of at least 200 Kbps and up to 84Mbps. Nowadays, 3G networks are already widely deployed in the market. However, 3G networks are being replaced by the fourth generation, also called 4G, which uses two

new standards: Long Term Evolution (LTE) [Sesia et al., 2009] and Worldwide Interoperability for Microwave Access (WiMax) [Vaughan-Nichols, 2004]. Both LTE and WiMax are based on Orthogonal Frequency Division Multiple Access (OFDMA). The standards define peak rates of 100 Mbps for high mobility communication and 1 Gbps in low mobility scenarios. Nowadays, LTE networks dominate the market of 4G cellular data services.

The key difference between 3G and 4G networks is that OFDMA increases the flexibility of resource allocation by increasing the quantity of time and frequency slots [Ghosh et al., 2010]. For simplicity, we will refer to 3G and 4G networks as 3G, since 3G is largely deployed by the WISPs and has broader support by commercial off-the-shelf devices.

Academia and industry are working towards the fifth generation of mobile network standards (5G), which aims to improve scalability in terms of cost, energy and resource efficiency [Osseiran et al., 2013]. 5G is a work in progress and it is planned to be rolled out to the general public in 2020 or further.

As part of the 5G network, a work in progress group in 3rd Generation Partnership Project (3GPP), formed by several IT companies, attempts to extend the coverage of LTE antennas through Device-to-Device communication (D2D). The project, called *Study on LTE Device to Device Proximity Services* (ProSe), has identified the use cases and requirements, and now the physical layer is being refining to allow Device-to-Antenna and Device-to-Device communications on the LTE radio [3GPPP, 2013]. This project can be seen as a promising solution for short-range communication in the mobile opportunistic data offloading problem.

2.2 Opportunistic Networking

Opportunistic networks (OppNets) are networks that, unlike classic networks, are prone to frequent disconnections and high communication delays. In some scenarios, it may be the case that node disconnection is the most frequent state since nodes may only communicate when a link is established (in OppNet jargon, this is called a *contact*).

The frequent disconnections preclude the use of classic message forwarding paradigms, since these paradigms are based on the establishment of an instantaneous end-to-end path from source to destination. As a consequence, OppNets employ the *store-carry-forward* paradigm, where messages are stored in intermediate nodes until a suitable forwarding opportunity occurs. Each node selects a set of messages to be forwarded using the recently established link, using some sort of priority scheme [Zhang,

2006]. The process of storing a message for later transmission is also known in the literature as *custody*.

The second key aspect of OppNets is the typically long end-to-end delay. In interplanetary networks, this delay is due to the distance between source and destination, while in vehicular networks or pocket switched networks this is due to the long disconnection times. As a consequence, OppNet messages tend to be *self-contained*, as connection-oriented protocols or interactive protocols tend to perform poorly under long delays (due to the high bandwidth-delay problem [Katabi et al., 2002]).

Opportunistic networking is based on the same principles of Disruption/Delay Tolerant Networking (DTN) [Burleigh et al., 2003]. The term DTN was used to referring communication between satellites in interplanetary networks. Since the term OppNets was coined after DTNs, it is important to clarify the difference between both concepts, which are frequently mistaken as the same thing. In our view, DTNs are a special case of opportunistic networks. DTNs were developed for the interconnection of networks (i.e., an inter-network protocol for Internets), where the interconnection among those networks suffers from long disconnections and interruptions. DTNs operate over the TCP/IP protocol stack, serving as a "gateway" for interconnecting networks over delay and disruption-constrained links.

OppNets, meanwhile, are a broader concept, since they support the disconnection and interruption of communication among networks, as well as among nodes within the same network. OppNets can use DTN or TCP/IP protocol stack or any protocol whatsoever. Further, OppNets are characterized by the use of the *store-carry-forward* paradigm, where messages are stored in secondary memory (e.g., hard drives and flash cards), and those messages are forwarded whenever a communication link is established. Although the protocols proposed in the DTN RFCs may be employed in other scenarios, it may be too costly to implement the DTN protocol stack in other networking applications (e.g., sensor nodes installed in wild animals [Juang et al., 2002]), and as such we consider only the concept of custody and self-contained messages from the DTN terminology.

OppNets are also found in the literature under different names, such as *Challenged Networks* or *Intermittently Connected Networks*, referring to scenarios in which it is not possible to guarantee an end-to-end path between nodes.

Due to the frequent confusion among the DTN and OppNet concepts, as well as the large range of applications and restrictions found in OppNets, we propose a taxonomy to clarify the similarities and differences between the concepts and applications below. More details can be found in [Mota et al., 2014].

2.2.1 Types of Opportunistic Networks

We propose the taxonomy shown in Figure 2.1. Opportunistic Networks are divided into Challenged Networks, in which contacts are opportunistic, and Delay Tolerant Networks, in which contacts are predictable. These networks are detailed below.

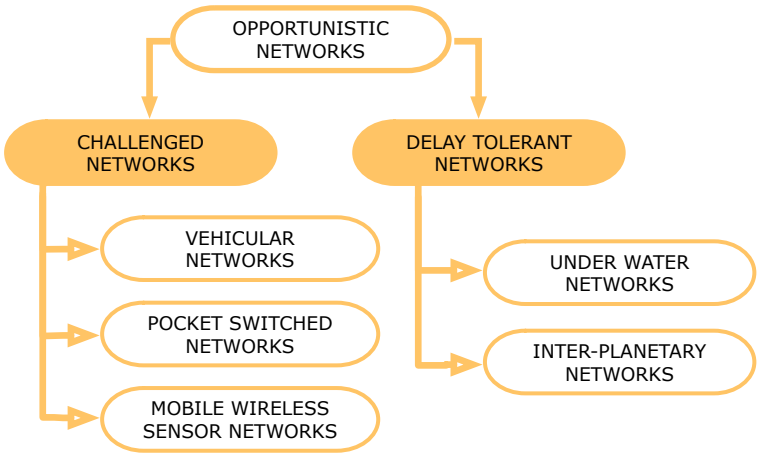


Figure 2.1. Taxonomy of Opportunistic Networks.

2.2.1.1 Delay Tolerant Networks

We classify DTNs as the scenarios that strictly follow the Bundle Protocol implementation developed by the IRTF DTN Research Group [DTNRG, 2013]. Figure 2.2 shows the protocol stack of a DTN node. The Bundle implements the *store-carry-forwarding* paradigm, implementing hop-by-hop reliability and security, instead of end-to-end as in the TCP/IP protocol stack. The DTN reference implementation is based on Unix operating system and follows the standards described in RFC 4838. This DTN protocol implementation also supports reliable communication among two DTN nodes [Ramadas et al., 2008]. Two subclasses of networks belong to DTNs:

- **Underwater Networks:** Enable applications such as oceanographic data collection, pollution monitoring, offshore exploration, disaster prevention, assisted navigation and tactical surveillance [Partan et al., 2006; Akyildiz et al., 2005]. As in terrestrial networks, the energy is a limited resource in this type of network, since the nodes have a limited battery and the transmission cost is high. The high error rates of underwater links as well the sparsity of nodes, which also could have depleted their energy resources, make use of DTN paradigm as a suitable solution.

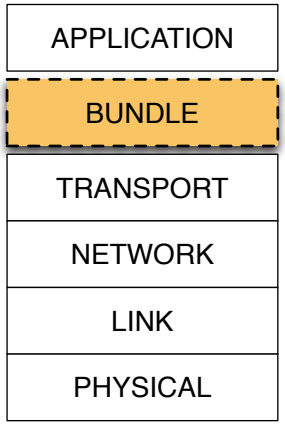


Figure 2.2. Protocol stack of a DTN node.

- **Inter-Planetary Networks:** As mentioned before, the interplanetary network was the primary research focus in DTN, since it may suffer from very large and variable propagation delays, low data rates, time-disjoint periods of reception and transmission, intermittent connectivity but predictable contacts, high link error rates, and lack of fixed communication infrastructure [Wang et al., 2009]. The two main factors that influence the operation of IPNs are the long distances between nodes, as well as the movement of planets and satellites [Akyildiz et al., 2003]. The former generates long delays and high error rates, while the latter creates frequent disconnections. A particular factor in IPNs that differ them from other types of opportunistic networks is the predictability of contacts, since it is possible to calculate when the next contact will happen using the equations for the trajectories of celestial bodies and satellites.

2.2.1.2 Challenged Networks

Challenged networks violate basic assumptions of the Internet architecture: the existence of an end-to-end path or the existence of a bounded round trip time between source and destination [Fall, 2003]. Moreover, *challenged networks* are characterized by one or more of the following characteristics: high latency, frequent disconnections, long queuing times, limited longevity and limited resources [Fall, 2003].

Challenged networks occur frequently in mobile networks, due to high node mobility, sparse density or unreliable links. In fact, how to determine whether a network is a Mobile Ad hoc Network (MANET), where the traditional Internet architecture applies, or whether it is a challenged network is unclear. Several routing protocols for MANETs were proposed, for instance AODV [Perkins and Royer, 1999], OLSR [Clausen and

Jacquet, 2003] and DSDV [Perkins and Bhagwat, 1994]. However, these protocols perform poorly in challenged networks, which require novel routing protocols.

Manfredi et al. proposed a framework to evaluate when to use traditional routing strategies, opportunistic routing or simply message flooding based on three measures: the average number of flows in the network, the probability of existing an arbitrary route, and the unpredictability of the network topology [Manfredi et al., 2011]. After analyzing several traces, they found that real networks may sometimes behave as an opportunistic network and sometimes as a MANET, and parts of the network may display different behaviors.

Since challenged networks encompass several applications, we highlight below the three most popular scenarios in OppNet context:

- **Vehicular Ad Hoc Networks (VANETs):** VANETs differ from MANETs in the following characteristics [Martinez et al., 2011]:

  - Predictable mobility: vehicles have to stay on the road and follow directions and speed limits;
  - High mobility: the network topology changes rapidly because of the vehicle speed;
  - Variable topology in time and space: traffic jams and location (urban or rural) influence the network topology;
  - Large scale: all vehicles are potential nodes of the network;
  - Partitioned networks: the communication range is limited, creating partitions;
  - No energy limitations.

  Further, the duration of contacts between cars is small even at low speeds and traditional TCP is unreliable in such scenarios [Gil-Castineira et al., 2008].

- **Mobile Wireless Sensor Networks (MWSNs):** Wireless Sensor Networks (WSNs) can be used to monitor remote locations such as rainforests and volcanoes. Such networks employ low power sensor nodes due to size, price and battery constraints, and as result, the network lifetime is usually short [Akyildiz et al., 2002]. The concept of WSN was extended to Mobile WSN (MWSN), where sensors are used to monitor wildlife. For instance, in the ZebraNet project [Juang et al., 2002] zebras use a collar with a sensor and data is transferred opportunistically whenever zebras pass near fixed access points. Examples of other

projects, which use opportunistic communication for animal tracking, are the SWIM project for whale tracking [Small and Haas, 2003], and TurtleNet project for turtle tracking [Turlenet, 2013].

- **Pocket Switched Networks (PoSNets):** It is a type of network formed by people carrying portable devices (e.g. smart phones and tablets) [Hui et al., 2005]. In such environments, mobile nodes are sparsely distributed and networks are often partitioned due to geographical separation or node movement. Knowing the human mobility pattern is one of the key factors to develop efficient protocols in this kind of network. PoSNets can enable social-based and gaming applications, and further, help the infra-structured network, with the fixed nodes providing services (e.g., Internet), and mobile users interacting in an opportunistic fashion. This is the type of network being studied in this thesis.

2.3 Game Theory

Game theory aims to address situations which the outcome of a participant of the game is affected by its decision, as well as the decisions made by other participants that are interacting with it [Easley and Kleinberg, 2010]. Game theory has been used over the years in the context of psychology, economy and computer science.

2.3.1 Basics

The participants of the game are called *players*, while their decisions are known as *strategies*. A player receives a *payoff* based on the strategy chosen by it and the strategies chosen by the other players. It is reasonable to assume that each player will choose a strategy that maximizes his or her payoff, taking into consideration what strategy the other players can use. Formally, a game consists of a set of players, where each player i selects a strategy $s_i \in S$. The set of strategies selected by all players is denoted by $P = (s_1, ..., s_n)$, where s_i is the payoff given by the strategy selected by player i.

Two fundamental concepts in game theory are *Best Response* and *Nash Equilibrium*. *Best response* is the strategy chosen by a player that maximizes his/her payoff, independently of the strategy chosen by other players. The *Nash equilibrium* is achieved when all players choose strategies that are their best responses. This definition of equilibrium was coined by John Nash in 1951, which has proven the following statement:

> *A finite non-cooperative game always has at least one equilibrium point* [Nash, 1951].

The idea behind Nash equilibrium is based on the fact that a player acts independently and chooses a strategy that maximizes his/her payoff regardless of the strategy chosen by others, and other players will do the same. Thus, players have no reason to deviate to an alternative strategy.

A *Bayesian game* (or incomplete information game) occurs when players have private information, and they are uncertain about the preferences and intentions of others. Bayesian games also consider the *type* of each player, which defines the behavior of each player. Furthermore, a player can learn the behavior of other players during the game.

For instance, consider routing in opportunistic networking as an example of a Bayesian game. In this game, nodes represent the players; to be selfish or to be cooperative are the set of types each player can assume, and the strategies are forwarding or not-forwarding the messages. In the beginning of the game, a node (a player) knows its behavior (type), selfish or cooperative, but has no knowledge about the behavior of the others. However, as time passes, nodes can estimate if another node is selfish based on its past behavior.

A strategy profile defines a type of a player, its strategy and the strategies chosen by other players. In a two-player game, a strategy profile is denoted as *(Type of player 1:strategy chosen by player 1, strategy chosen by player 2)*. The *Bayesian equilibrium* is a strategy profile that maximizes the payoffs of all players given the type and the strategy chosen by one player and his/her beliefs about the type and strategy chosen by the other players. In other words, a strategy profile is a Bayesian Nash equilibrium if only it corresponds to the best response for all players.

2.4 Utility Theory

In this thesis, we propose a multi-criteria decision model based on *Utility Theory* to decide when and who should switch from infrastructure to opportunistic communication.

2.4.1 Definition

Utility theory quantifies the set of preferences of a customer into a numeric scale [von Neumann and Morgenstern, 1953]. Since its proposal in game theory, it has been used

in economics and other fields where decisions are taken based on a set of quantifiable preferences. The utility of a good or a service (named x) can be denoted as a mathematical function $u(x)$.

A utility function $u(x)$ quantifies the preference for a given criterion x. A utility function which quantifies only one criterion is called *single-criterion* decision utility, otherwise it is called *multi-criteria* decision utility. Furthermore, the utility of an *upward* criterion x increases with x , while for a *downward* criterion the utility decreases with an increase of x. The key aspect of utility theory is to choose the adequate utility function $u(x)$ for each criterion x, where $0 \leq x < \infty$. Assuming the decision maker is rational, it always chooses the criterion value that maximizes its payoff.

In network selection, utility theory aids in mobile device migration among access points [Ormond et al., 2006] and among radio technologies, e.g: roaming between 3G and WiFi, based on the properties of the available networks and links [Nguyen-Vuong et al., 2008; Abid et al., 2012].

2.4.2 Utility Function: Single and Multi Criteria

Several mathematical functions were proposed to quantify a criterion, including linear, exponential and sigmoid functions [Nguyen-Vuong et al., 2008; Abid et al., 2012]. In [Nguyen-Vuong et al., 2008] the authors analyzed several of these functions and showed that only sigmoid functions fit all the requirements for network selection, that is, they select the best communication interface based on several criteria.

A sigmoid function is smooth, bounded and strictly increase/decrease output values [Han and Moraga, 1995]. Thus, a sigmoid equation must satisfy the following properties:

1. $u(x) \in [0, 1]$
2. $u(x_t) = 0.5$ for a given x_t

In multi-criteria utility theory, an aggregator function $U(.)$ combines the utility function of each criterion $u(x)$. In the multi-criteria decision problem, it is necessary to compute the aggregate utility of all criteria. Moreover, the complexity increases with the number of criteria since two or more criteria can be conflicting. Moreover, satisfying all the criteria at the same time is a NP-hard problem [Gazis et al., 2003].

Several aggregators were proposed in the literature, however the most common are the additive multi-criteria utility function [Jacquet-Lagreze and Siskos, 1982] and the weighted product model [Triantaphyllou and Mann, 1989], also known as multiplicative multi-criteria utility [Nguyen-Vuong et al., 2008]. We describe below each one:

1. *Additive Multi Criteria Utility function (AMC)*: Function described in [Jacquet-Lagreze and Siskos, 1982] and used as aggregator to network selection problem in [Adamopoulou et al., 2006]. This function is a weighted sum of the utility of each criterion, as shown in Equation (2.1):

$$AMC(x_1 \dots x_n) = \sum_{i=1}^{n} w_i u(x_i) \tag{2.1}$$

   where w_i is the preference for the criterion x_i and $\sum_{i=1}^{n} w_i = 1$. In this case, users can assign a weight for each criterion.

2. *Weighted Product Model Utility function (WPM)*: the final utility is a weighted product of each criterion, as shown in Equation (2.2):

$$WPM(x_1 \dots x_n) = \prod_{i=1}^{n} [u(x_i)]^{w_i} \tag{2.2}$$

Table 2.1 illustrates an example for these two aggregation functions. Let suppose we need an aggregated utility function based on three criteria. Consider two scenarios with different values for the single utility for each criterion: $u(1)$, $u(2)$ and $u(3)$. In *Scenario 1*, all single utilities are equal to 0.5, then both aggregators have the same value. However, in *Scenario 2*, $u(3)$ is zero and AMC provides a high aggregated utility value (dismissing a criterion with value zero). In WPM, a criterion with zero value causes the aggregated value to be equal zero. This is known as the problem of zero limits.

Utility Criterion	Weight (wi)	Scenario 1	Scenario 2
$u(1)$	0.333	0.5	0.7
$u(2)$	0.333	0.5	0.9
$u(3)$	0.333	0.5	0.0
AMC		0.5	0.56
WMP		0.5	0

Table 2.1. Aggregators Utility function for two scenarios.

2.5 Conclusion

This chapter introduced the concepts used in this thesis to face the mobile offloading problem. We presented the mobile broadband Internet technologies. Next, we intro-

duced the concepts of opportunistic network, which gives support to disconnections when devices are out of range of each other. Further, basic game theory was introduced, which will be used to model user behavior. Finally, we presented utility theory, which will be employed in our solution to assist devices switch from 3G network to opportunistic communication.

Chapter 3

Related Work

The struggle of WISPs due to increasing number of users and bandwidth requirements in the last years has attracted the attention of both academia and industry to mobile traffic offloading. This chapter presents a review of the state of the art and related works in the literature.

The rest of this chapter is organized as follows: Section 3.1 describes the methodology used to shed light over the state of the art. Related work addressing cellular offloading are classified in three distinct ways to offload the cellular network: Femtocell, WiFi and opportunistic offloading, which are discussed in sections 3.2, 3.3 and 3.4, respectively. Incentive mechanisms to engage users cooperation on opportunistic networks are discussed in Section 3.5. Finally, Section 3.6 concludes the chapter.

3.1 Methodology

The state of the art was obtained through the major online search platforms for scientific literature[1], which return as results publications from several scientific publishers (e.g. ACM, IEEE and Elsevier). The results were ranked by year, and high priority was assigned for the latest published works and by their quality[2]. In the cases where the same paper had been published in conference proceedings and periodicals (as an extended version), the periodical was referred.

As mentioned in Chapter 2, opportunistic cellular offloading emerged as an application scenario for opportunistic networking. Thus, first we conducted a deep literature review on opportunistic networking. Next, we narrowed our research to focus on cellu-

[1] http://scholar.google.com and http://www.sciencedirect.com

[2] A Brazilian educational governmental committee defines a metric, called Qualis, to measure quality of journals and conferences.

lar offloading proposals. Finally, we reviewed the state-of-the-art on cooperation and incentive mechanisms to engage users to adopt opportunistic offloading. The systematic research for each step is described below:

i) Opportunistic networking: Works published from 2000 up to 2015 matching the following query string:

(OPPORTUNISTIC NETWORK) OR (DELAY TOLERANT NETWORK))

As result, the forwarding algorithms, real applications developed and commonly used mobility models, real contact traces and simulators were surveyed in [Mota et al., 2014].

ii) Mobile Internet cellular network offloading: the state-of-the-art in cellular offloading was obtained with the research query:

(((MOBILE BROADBAND INTERNET) OR 3G OR 4G OR 5G)
AND
(CONGESTION OR OFFLOADING)

Three approaches to offload the cellular architecture were observed: Femtocell-based, WiFi Based and Opportunistic (or Device-to-Device) Offloading. For the sake of clarity, we discuss each approach separately in the next three sections.

iii) Incentive Mechanisms: The query string to obtain related works on incentive mechanisms was:

(COOPERATION OR (INCENTIVE MECHANISM))

Since incentive and cooperation are well studied in other research areas, such as economics and sociology, this research returned thousands of papers, journals and books. High priority was given for the most cited published works. Then, this research was narrowed to consider also the keywords:

(COOPERATION OR (INCENTIVE MECHANISM))
AND
(P2P OR AD HOC OR OPPORTUNISTIC OR DEVICE-TO-DEVICE OR
(CELLULAR OFFLOADING))

The research for related work using the queries above was performed until June 2015.

3.2 Femtocell Offloading

Femtocells are small low-power base stations connected directly to the wired backhaul connection [Chandrasekhar et al., 2008]. Modern femtocells have autonomous auto-configuration capabilities, allowing them to be deployed in a plug-and-play manner by the end user [Chandrasekhar et al., 2008]. In this sense, they are similar to WiFi access points, but using the cellular licensed spectrum.

Femtocells were originally applied to improve the cellular indoor coverage. However, with the increase of mobile Internet demand, Femtocells became a solution to offload data traffic from cellular infrastructure. Thus, WISPs and their partners might deploy femtocells in order to create smaller cells in cities, moving off the traffic from their main infrastructure.

In [Mukherjee, 2011] the authors make an analytical formulation of the offloading capacity of femtocells in WCDMA employing statistical models. They showed, under certain realistic parameters of radio propagation, that 25% of the users distant more than 3 km of a cellular network may switch to a femtocell network if available. However, they concluded that if femtocell antennas do not decrease their transmit power when close to a cellular infrastructure, femtocells will generate high amounts of interference with the cellular network.

A known issue with large deployment of femtocells is the resulting interference [Haldar et al., 2013]. Furthermore, [Andrews et al., 2012] surveyed the following technical challenges in femtocell offloading:

- Cell Association. Autonomously assign a user to an antenna considering the different cell sizes.
- Mobility and handover. Femtocells have smaller coverage. Thus, it requires seamless handover. In traditional cellular networks, each device has a constant IP Internet address with data routed through a fixed gateway. Different femtocells may represent different networks, thus requiring new IP addresses to the device. 3GPP standardizes specific procedures for vertical handover between femtocells and non-cellular technologies.
- Self-organizing network. Since femtocells are deployed by customers or enterprises, it is expected than the number of femtocells overcomes cellular antennas in orders of magnitude. Thus, maintenance is not scalable. Due to this reason, features such as automatic registration and authentication, neighbor discovery and network optimization have been defined by the 3GPP standards to femtocells.

3.3 WiFi availability and Offloading

IEEE 802.11 was originally designed with data rate of just 1 Mbps. However, the amendments 802.11a/b/g/n, broadly adopted in the market, increase the data rate up to 600 Mbps. Furthermore, IEEE 802.11ac aims to achieve 1 Gbps using multiple antennas, and the target data rate of IEEE 802.11ad, as proposed by the Wireless Gigabit Alliance, is about 7 Gbps. Nowadays, the deployment of IEEE 802.11 access points scattered around several cities leads the major effort to offload overloaded cellular infrastructures.

WiFi coverage in metropolitan areas was characterized in [Bychkovsky et al., 2006]. Bychkovsky*et al.* conducted an experiment to evaluate the feasibility of using WiFi access points around Boston metropolitan area, driving vehicles at regular speeds in the city, during July 2005 and July 2006. They showed an average duration of link layer connectivity of 24 seconds, while only 3.2% of access points provided end-to-end communication, which means that applications using only open WiFi connection should be delay-tolerant.

A comparison between WiFi and 3G networks appears in [Gass and Diot, 2010] and [Chen et al., 2012]. In [Gass and Diot, 2010], the authors show that since the contact time 3G networks is greater than on WiFi networks, when the client is moving, the amount of transferred data (download) is larger in a 3G network. However, since the upload data rate is of the order of kilobits per second in 3G networks, WiFi networks outperform 3G network on data upload.

In [Chen et al., 2012] the authors characterize throughput, loss rate and round trip time of the three major WISPs in the United States, and show that in some cases 4G/LTE networks outperform WiFi networks. They propose a multi-path version of TCP to increase the throughput of mobile clients using 4G and WiFi network simultaneously.

Detection of available WiFi Hotspots, as well as a handoff mechanism, were proposed in [Balasubramanian et al., 2010] and [Lee et al., 2010]. In 2010, Balasubramanian *et al.* measured 3G availability in Amherst, Seattle and San Francisco. 3G signal covered 87% of the cities, while open WiFi access points covered only 11% [Balasubramanian et al., 2010]. The authors also propose *Wifler*, a framework to switch between WiFi and 3G quickly, based on the characteristics of traffic, delay-tolerance and application-specified QoS metric. For instance, VoIP application has zero delay-tolerance, while web apps may tolerate 20 seconds of delay. Lee *et al.* recruited one hundred iPhone users to collect WiFi connectivity statistics in Seoul, they observed that users are most of the time under WiFi coverage, although it is unclear whether

users were able to connect to WiFi networks or not. In both works, it was observed that delay-tolerant applications could offload data from 3G networks without quality loss to the user.

Regarding the use of femtocell or WiFi networks to offload the traffic, as noted by [Andrews et al., 2012], providers are likely to adopt both solutions. WiFi and femtocells are complementary solutions to offload data traffic. Nevertheless, performance of WiFi network degrades while the density of access points increases. This degradation occurs since the 802.11 standards do not coordinate use of the spectrum between access points. Meanwhile, femtocells can deal better with this issue.

3.4 Opportunistic Mobile Data Offloading

In opportunistic offloading, mobile user's devices can also play the role of a *relay* between other devices and infrastructure. Since opportunistic offloading is a recent research subject, varied nomenclature for the same offloading approach appears in the literature, such as D2D offloading and content floating offloading. In all cases, it is assumed that the network may not be connected all the time. Thus, we opted for the term *mobile opportunistic offloading*, which better depicts this network feature.

We considered two opportunistic offloading approaches: *i)* Opportunistic nodes forwarding their data through relays thus reducing the number of users in a cellular infrastructure. *ii*) Special nodes (acting as relay) keeping content in their buffer and opportunistic nodes getting the content through these relays, thus reducing data traffic in a cellular infrastructure. In both approaches, a selection algorithm must select nodes as relay and nodes as opportunistic, which may try to use these relays instead of using infrastructure.

3.4.1 Selecting Best Relay Candidates

The efficient dissemination of a message among nodes in an opportunistic network was studied [Khabbaz et al., 2011]. In this case, the main goal was to select the best candidate to forward a message from source to destination. Opportunistic Mobile data offloading aims to select a subset of k nodes to act as a *relay* between other nodes and the infrastructure.

In [Doppler et al., 2009], the authors studied the feasibility of opportunistic offloading in a local area cellular network, which they called Device to Device (D2D) communication. The authors focused on changes required in the cellular infrastructure to deploy D2D communication. Furthermore, devices are chosen as relay randomly

among all devices. They showed that use D2D communication instead of cellular infrastructure increases the overall throughput up to 65%. Doppler et al. extended their mechanism by choosing device communication mode, opportunistic or infrastructure, based on link quality between the devices, interference and quality of the cellular link [Doppler et al., 2010]. However, both mechanisms disregard user willingness to become a relay.

In Barbera et al., the authors proposed a mechanism called *VIP*, in which the most important nodes of the network are chosen to act as relays carrying data for other network participants. The importance of a node is measured through social network metrics - such as Betweenness centrality, Degree Centrality, Closeness centrality and page rank [Barbera et al., 2011]. After ranked, the nodes are promoted to a status of *VIP nodes* based on a global strategy and a community-based strategy. The authors showed their strategies select a small number of VIP nodes covering up to 90% of the network. However, the main constraint in this mechanism is the requirement of knowledge and time to observe the topology graph of the network and to calculate the metrics.

Han et al. proposed offloading 3G networks by selecting a subset of k nodes to receive a defined content. Thus, the other nodes, called *infected users*, could get the content directly from this subset, reducing the data traffic in 3G infrastructure. The goal is to select the minimum k nodes, which the 3G network would push the content to maximize the number of *infected users*. The authors modeled this problem as the *target-set problem* [Han et al., 2012]. They showed that a *greedy* approach can achieve an approximation ratio of $1 - 1/e$ of the optimum. However, this approach requires knowledge of the future mobility of users. To overcome this issue, the authors considered mobility history of the user to predict future content delivery. Nonetheless, WISPs are responsible for selecting the nodes and the content *pushed* to the nodes, requiring changes in its infrastructure.

Similarly, in [Whitbeck et al., 2012] was proposed a framework, called *push-and-track*, with several strategies to decide *when*, and to *whom* the content should be injected in a network. The authors considered a periodic flooding scenario, where content created at time t must be delivered to all nodes within a period T. The framework needs to decide how many copies of content will be produced and to whom send these copies. The question is how many copies will be initially pushed to nodes and when it will be mandatory to re-inject new copies. The authors used two strategies: slow start, when few copies of a content are pushed to the network; and fast start, a large number of copies is introduced instead. The authors show that, when there is delay tolerance, both strategies are equivalent.

Whitbeck et al. considered four strategies to decide to whom they would deliver the content: i) *random*, pushing content to random nodes; ii) *Entry time*, pushing content based on the time the node entered an area; iii) *GPS-Based*, pushing content to a node within the highest density area or to nodes with highest potential to infect other nodes. The authors compared Push-and-track against a dominating set oracle, and observed that besides the sophisticated proposed strategies, *random whom-strategy* performs better in most of the analyzed cases. We used this observation to propose a pseudo-random solution, although considering a set of criteria observed by the device.

MobiCache aims to offload cellular networks by choosing the relay nodes in an area of interest [Zhang et al., 2015]. The WISP in turn injects the content in nodes with a greater probability to achieve the area of interest. The authors consider a geographical routing to achieve this, and all nodes inside the area determined by the WISP receive the content.

Table 3.1 summarizes the algorithms proposed for relay selection in the literature.

Table 3.1. Opportunistic mobile offloading proposals

Name	**Base Algorithm to Select Relays**	Reference
Doppler, 2009	Random-based	[Doppler et al., 2009]
Doppler, 2010	Quality of signal and interference	[Doppler et al., 2010]
VIP	Social networks metrics	[Barbera et al., 2011]
Ho, Ban proposal	Target set problem	[Han et al., 2012]
Push and track	Random and GPS-based	[Whitbeck et al., 2012]
MobiCache	GPS-based	[Zhang et al., 2015]

Our proposal differs from the state of the art because nodes (instead of the infrastructure) elect themselves as relays, based on a multi-criteria utility function, which takes the node context into account (e.g. number of neighbors and battery life). Unlike previous work our proposal does not require modifications in the infrastructure.

Furthermore, use graph or location metrics can bring disadvantages, such as draining battery faster and disregarding user willingness to become a relay.

3.4.2 3GPP Device-to-Device Proximity Services

The related works in the previous subsection assume a short-range ad hoc communication between close devices. Bluetooth and WiFi are the most common technologies responsible for these types of communications. In practice, both have limitations to perform ad hoc style communication. For instance, Bluetooth requires paired devices, and off-the-shelf WiFi does not support neighborhood discovery.

The 3GPP consortium targets the reuse of LTE radio interface for short communication between devices using the specifications of the Device-to-Device Proximity Services (D2D ProSe) [3GPPP, 2013]. The three main market drivers for the development of D2D proximity are: Public safety networks in infrastructure-less situation; Advertisement on devices that are near or inside a commercial place; and Network offloading. The latter being our research interest.

The key difference between D2D ProSe and solutions proposed to mobile ad hoc network relies on an assistance from the network infrastructure, where infrastructure can participate of session's setup, connecting and creating routes between devices. D2D proximity service can be defined as the communication between two devices close to each other using LTE radio interface, whose data messages are routed directly or via local infrastructure network. Moreover, D2D concerns one-hop communication between two devices.

In Doppler et al., the authors proposed a framework for device-to-device communication in which the cellular infrastructure controls the session initiation and setup among devices. The authors show, through simulations, that device-to-device communication can increase the throughput in a cell area [Doppler et al., 2009].

However, since D2D proximity services are still under specification, deployed cellular infrastructure has no support for device-to-device communication, and this support may still take some time to be rolled out in production networks. In [Tsolkas et al., 2013], the authors proposed a framework to manage device-to-device connections with minor changes in the core of the cellular infrastructure.

In Asadi and Mancuso, the authors used WiFi radios to form clusters among devices, only the cluster head communicates with the infrastructure, using LTE. They proposed the use of WiFi Direct, which allows neighborhood search, discovery and direct communication via WiFi interface. Hence, it minimizes changes in the cellular infrastructure [Asadi and Mancuso, 2013].

The improvements on the network capacity and spatial spectrum reuse were shown in [Min et al., 2011] and in [Yu et al., 2011]. An extended overview of proximity services using LTE radio proposed by the 3GPP can be found in [Qualcomm, 2012] and [Lin et al., 2013].

Research conducted by the 3GPP working group and the academia are complementary, since the main focus differs. 3GPP focuses on how to discover neighbors, connect and transfer data, and standardizing device-to-device communication. Whereas academia focuses on routing, content distribution, privacy, security, etc.

3.5 Selfishness, Cooperation and Incentive Mechanisms

Selfishness, altruistic and cooperative behavior of human beings were extensively studied in philosophy, psychology, economics and recently, in the context of computer science [Miao et al., 2013]. Selfishness can be defined as the act of benefiting itself instead of another. On the other hand, altruism favors others instead of itself [Levine, 1998]. Cooperation occurs when an individual devotes an effort that implies a cost in some collective activity expecting some benefit. Unlike altruism, in cooperation, the individual expects some benefit greater than the costs [Bowles and Gintis, 2003]. Incentive mechanisms aim to engage users to cooperate with others.

In computer networks, selfish nodes can be defined as nodes unwilling to forward packets. Selfishness has been perceived as a bottleneck for network applications that depend on user cooperation. Indeed, several studies have been carried out to shed light on the impact of selfish nodes within different networking contexts. For instance, in Gnutella, a popular Peer-to-Peer file sharing years ago, 70% of the users shared no files and 50% of the file chunks were uploaded by the 1% most active users [Adar and Huberman, 2000].

In mobile ad hoc networks (MANETs), Marti el al. showed that in a network with 40% of selfish nodes, the average throughput degrades up to 32% [Marti et al., 2000]. The authors propose a mechanism to detect and avoid selfish nodes using watchdogs. Crowcroft *et al.* considered the costs to relay messages and modeled an incentive mechanism for ad hoc networks based on bandwidth and power usage of each node, where each node is restricted to generate an amount of traffic directly related to its credit balance [Crowcroft et al., 2004].

The first study on the impact of selfishness in opportunistic network appeared in [Panagakis et al., 2007]. The authors showed that the delivery ratio decreases linearly with the amount of selfish nodes.

In *Participatory Sensing Networks* (PSNs), users are required to sense data, process it and transmit it to PSN servers. In PSNs, the user willingness and cooperation among participants reflects directly on quality and quantity of sensed data, and hence improving services offered by the network [Lee and Hoh, 2010].

However, PSN applications may consume resources of devices. Indeed, users may deem too costly to contribute to the network. These costs can be in terms of energy or transmission data in case of use the allowance data from its operator.

In opportunistic mobile offloading, all those costs also may affect user's willing-

ness to cooperate. However, incentive mechanisms for opportunistic offloading are a broader problem than the previously context mentioned, since it may involve the cellular operator. The natural solution to engage users to help alleviate traffic congestion is reward mechanisms. However, there is a tradeoff between the value paid for cooperative users and WISP revenue.

In all cases discussed above, cooperation in the context of wireless networking depends on the relationship between benefits (B) and costs (C) to participate on it [Fitzek et al., 2013]. Fitzek et al. claim cooperation occurs whenever user believes that the benefit is higher than the cost of collaborating. This benefit may be a reward, improvement of quality of experience, e.g. increase throughput, or even by altruistic feeling of contributing to the network.

There are also situations in which the benefit for cooperative behavior is unclear. In these situations, incentive mechanisms act as a "driving force" to encourage user cooperation. In this section, we discuss costs and benefits and how incentive mechanisms could engage user cooperation in the context of wireless networking.

Accordingly to the *Benefit-Cost* rule, cooperation arises when the benefit is greater than the cost. The following types of cooperation satisfy this rule [Fitzek et al., 2013]:

1. **Altruistic** $rB > C$.- Altruism based on Hamilton's rule. It takes into account a relationship factor r between who bears the cost and the beneficiary. An example of Hamilton's altruism is when someone cooperates with friends or family members. Thus, even if real benefit is small, the relationship factor increases the benefit.

2. **Forced** - $B < C$. Cooperation is mandatory or there is no benefit ($B = 0$). Forced cooperation applies when the network consists of a legacy system or devices must cooperate to participate on network. For instance, in wireless sensor networks (WSNs) nodes forward third party packets without requiring benefits. Another example, WISPs could deploy opportunistic offloading without users' consent. In this case, there are no benefits for users.

3. **Technical** - $B > C$. There is a clear benefit to cooperate with the network. For instance, a user cooperates with participatory traffic information systems in order to obtain up to date information about real-time traffic and road info.

4. **Social** - $B' > C$ e $B > 0$. A direct benefit B can be lesser than cost. However, a new benefit B', based on social network context, is offered for those who coop-

erate. As an example, a user can cooperate and offload or share mobile Internet with friends, which in turn reward his/her social network with a positive rating.

In recent years, academy and industry have proposed dozens of incentive mechanisms which overcome the benefit for cooperative users. These mechanisms either punish selfish users, classify users by their reputation, make cooperation more pleasant or even employ direct payment for cooperative users. We discuss this classification in the next section.

3.5.1 Classes of Incentive Mechanisms

Incentive mechanisms attempt to offer benefits that outweigh costs for each network participant. Miao *et al.* classified incentive strategies to engage relay nodes forward messages in opportunistic networking as *barter-based* (also known as *tit-for-tat*), *reward-based* and *reputation-based* [Miao et al., 2013]. Recently, Gao *et al.* surveyed thirty theoretical incentive mechanisms, eight experimental studies and nine implementations of incentive mechanisms applied to Participatory Sensing Network in the last ten years [Gao et al., 2015]. The authors classified the mechanisms as extrinsic (monetary incentive) and intrinsic (non-monetary). Figure 3.1 summarizes those incentive mechanisms classes.

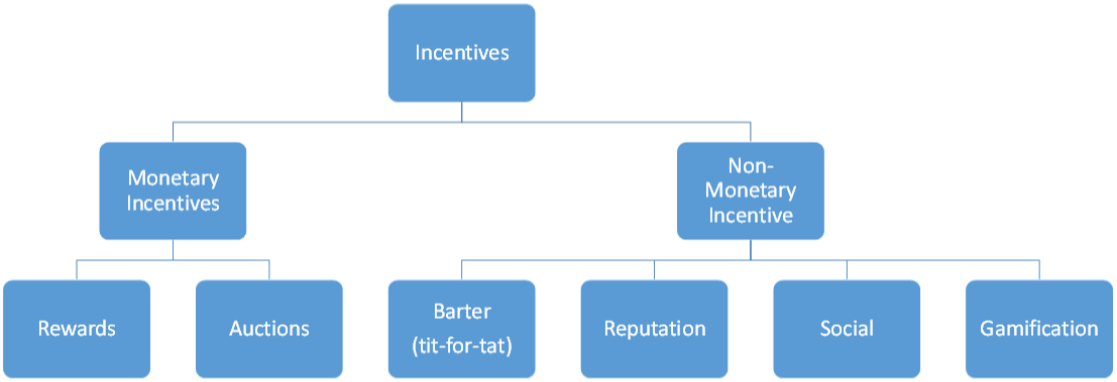


Figure 3.1. A classification of incentive mechanisms.

Monetary-based incentive mechanisms offer a direct benefit or income, where users may verify whether this benefit satisfies or not their expectations. On cooperative relaying, payment based on the amount of data forwarded from one device faces the problem of how to maximize the income for cooperative users, while minimizing the costs of who is paying.

On the other hand, non-monetary mechanisms avoid costs with direct payment, while the motivation for engagement is subtle. In fact, reciprocity drives motivation. For instance, "Bob" cooperates to improve others performance expecting cooperation of other users to improve his performance. Building ranks, reputation tables, barter based and gamification are techniques used to measure trust and reciprocity of the users. These mechanisms differ on how they quantify users' contribution to a network.

3.5.1.1 Non-Monetary Incentive Mechanisms

Barter, reputation, social and gamification based represent non-monetary incentive mechanisms.

In *barter-based* strategies, nodes exchange the same amount of messages, thus balancing the use of resources among nodes [Buttyán et al., 2010]. However, such strategies impose a strong limit on the number of messages exchanged between the nodes.

In *Reputation-based* strategies, incentive mechanisms attempt to detect and avoid the selfish nodes by ranking them according to some metrics. For instance, in *Ironman*, a positive reputation is given by the nodes that participated in the path of a message from source to destination [Bigwood and Henderson, 2011]. A similar mechanism is *MobiGame* [Wei et al., 2011], which works as following: node A creates a message x to a node C, then A forwards x through B and B forwards it to C. Node C sends a *relay evidence* to B after receiving x. Finally, when node A meets node B, which is carrying the *relay evidence*, node A increases the reputation of B.

In *Social-based* strategies, the social structure, based on social networks as Facebook and twitter, is exploited to design efficient mechanisms. For instance, users can offload traffic for other users based on their social network. The cooperative communication can be by *trust*, as friendship or other relations, or by *reciprocity*, where users will cooperate with strangers expecting to receive cooperation in the near future. Chen et al. proposed an incentive mechanism based on social trust and social-reciprocity, to promote cooperative forwarding among devices [Chen et al., 2014].

In *Gamification-based* strategies, elements and design of games are used in non-related game contexts [Deterding et al., 2011]. Examples of such elements are scoring tables and classification; trophies or medals to reward users who perform a given task; and social graphs, see what "friends" are doing or how they progress in the task, so who performed more tasks has a higher ranking. The main goal is to use game elements with the purpose to perform tasks non related to the game [Werbach and Hunter, 2012]. For instance, these tasks may be to improve a skill, encourage fitness, or within

an opportunistic offloading, engage users to offer its device as a cache and relay in the network.

As example of gamification mechanism to encourage users, we can mention the *Waze* app for smartphones[3]. Waze provides various information of interest to users such as traffic alerts, fuel price of close gas stations, besides route calculation. In order to keep information up to date, Waze requires active participation of users, e.g. participants must manually report situations observed such as transit accident warnings. In Waze, game elements are represented by the use of avatars and scoring system. In this case, more cooperative users achieve special avatar or badges. As result, Waze achieves its goal of improving the volume of sensed data and quality of traffic information for all users.

3.5.1.2 Monetary Incentive Mechanisms

In *reward-based* strategies, the users receive endowments to incentivize cooperation. In the context of forwarding protocols, payments can be made either by each forwarded message or only when the message reaches the destination. There are mechanisms designed to allow participants negotiate with the central entity of the network the amount to be paid. For instance, in Participatory Sensing Networks, the participants negotiate the reward value for sensed data before sending them to the PSN's servers. In other mechanisms, a central entity (server) decides how much it will pay for data already sent by the participants of the network [Gao et al., 2015].

Yang et al. proposed two incentive mechanisms [Yang et al., 2012]: *MSensing Platform-Centric* and *MSensing Auction*. In *MSensing Platform-Centric*, PSN has a limited budget to spend with sensing tasks. PSN announces the reward for a task based and each participant receives a reward proportional to the time dedicated to the task. One problem with this model is that increase in the number of active participants decreases the reward received by each one.

In *MSensing Auction*, the PSN platform announces a set of tasks and each user chooses a subset by themselves. For each task that users select, they submit a tuple (task-bid) to the platform, where the "bid" is the value of the reward they want receive to perform the task. After receiving the offerings of users, the PSN platform selects a set of users as the winners of the auction, which will perform the tasks. A problem encountered in this type of mechanism is the explosion of the incentive costs [Lee and Hoh, 2010]. This cost explosion problem can derail the mechanism due the high cost expended by the platform. In addition, if the winner is always the user that offer the

[3]http://www.waze.com

lowest price, this user may be discouraged to continue sensing data for the PSN, due to the low values received.

Few studies addressed incentive mechanisms for opportunistic mobile data offloading. In VIP algorithm [Barbera et al., 2011], the authors just discussed that providers should pay or give gifts for users that help to alleviate the traffic. However, how much cooperative users should get paid is out of the scope.

A framework to incentivize users to tolerate delay and reward potential offloading nodes, called Win-Coupon, was proposed in [Zhuo et al., 2014]. Win-Coupon uses reverse auction mechanism to engage users to offload the cellular infrastructure. In a reverse auction, users send bids to the WISP informing how long they can wait for a request and how much they expect as a reward. WISP decides which bidders are winners of the auction based on the delay tolerance, income and offloading capacity of the bidders. The WISP infers the offloading capacity of each bidder based on historical parameters, such as previous requests and mobility pattern. The goal of Win-Coupon is to minimize the WISP incentive cost for a given offloading scenario.

MobiCache is another proposal based on reward to engage user cooperation [Zhang et al., 2015]. MobiCache rewards users that forward contents to others. Further, users that wait for content opportunistically instead of requesting it directly from the WISP pay a cheaper fee for the WISP.

The reward of users caching content to others is based on the value of energy savings to the operator. Zhang et al. argue operators shall pay for cooperative users the costs saved with the data offloading process. Therefore, MobiCache offers $6.67s \times 10^{-6}$ units of reward (e.g. cents) for each data of size s offloaded. Since users may consider such value too small, they have to forward (and offload) a high volume of data to receive a reward with some worth.

Table 3.2 summarizes incentive mechanisms focused on opportunistic offloading.

Table 3.2. Incentive mechanisms for Opportunistic mobile offloading

Name	**Incentive Mechanism**	Reference
VIP	Gifts or discounts	[Barbera et al., 2011]
Win-Coupon	Reverse-auction	[Zhuo et al., 2014]
MobiCache	Flat rate payment	[Zhang et al., 2015]

Reddy et al. realized small scale experiments to evaluate the effect of cooperation payments in participatory sensing networks [Reddy et al., 2010]. The authors concluded that incentives work better when micro payments are combined with other factors such as user altruism and competition among them. In addition, they showed that a fair payment for all participants kept them motivated for longer time than lower payments.

Indeed, payments can be counterproductive in some cases as shown in [Kamenica, 2012]. The author reviewed many studies in psychology and economics on the payment effect as mechanisms incentives. In many cases, a higher value or too low value proved counterproductive to induce him or her into collaborative behavior.

3.6 Conclusions

In this chapter, we presented the state-of-the-art techniques to offload cellular infrastructure and incentive mechanisms proposed to engage user cooperation. Mobile data offloading has gained attention from industry and academia in recent years. The study of availability of WiFi networks in cities allows WISPs estimate how much they can offload from their infrastructure through WiFi networks. Femtocells try to expand the 3G network coverage, although its adoption depends on users buying new equipment. Opportunistic Offloading attempts to extend coverage and relieves the signaling congestion on 3G networks by using end-users' devices.

Finally, we presented a background in cooperation and incentive mechanisms. Mechanisms to engage cooperation in human beings were extensively studied, and still, there is no silver bullet solution, which guarantee engagement under any scenario. In the computer networking context, several mechanisms were proposed for P2P, ad hoc networks, opportunistic networking and, recently for opportunistic offloading scenarios. These proposals range from the punishment of selfish users to payment for cooperative users. We discuss the impact of selfish user and propose a decentralized algorithm to engage users to cooperate in Chapter 7.

Chapter 4

On the Feasibility of WiFi Offloading

WiFi offloading appears as a solution to data traffic bottleneck in the WISPs infrastructure. Nowadays WISPs offer WiFi hotspots to their customers to engage them switch from 3G to WiFi communication. Besides WISPs hotspots, there are also public WiFi hotspot provided by the governments.

This chapter verifies the feasibility of WiFi offloading through the already installed hotspots in the cities during the user daily travels to work, home or leisure. We attempt to answer the following question: *Is it possible for WISPs to offload the data through the already deployed WiFi infrastructure?*

In order to achieve our goal, we implemented an Android Application and mapped the 3G and WiFi coverage through several bus routes in Paris in order to evaluate how users and WISPs can benefit from the existing WiFi infrastructure to offload mobile data. Paris was chosen based on the fact WiFi access point provided by a WISP to a customer allows other customers in the range to connect with it.

We carry out an analysis of the WiFi coverage, separating the public hotspots provided by the government, the access points from the WISPs and the private access points. Next we used a public database from a Location Sharing Service, in which users tell their localization, in order to realize the percentage of users that could have their data traffic offloaded by WiFi.

The rest of this chapter is organized as follows: We detail our experiment and analyze the results in Section 4.1. Section 4.2 discusses the feasibility of offloading with the deployed WiFi network. Finally, Section 4.3 presents the concluding remarks and the lessons learned with this experiment.

4.1 Connectivity Categorization

We used two databases: The free WiFi hotspots offered by the municipality of Paris, France and WiFi access points detected through several bus routes in the same city. The next sections detail our experiment and summarize the collected data.

4.1.1 Public WiFi Hotspot

The list of free WiFi Hotspots in Paris is available in the open data project[1]. The last update of this dataset is from 12 April 2012 and consists of 312 hotspots spread over the city.

We calculate the density of access points per km^2 (APs/km^2) by splitting the map in rectangular sections of 1km^2, and counting the number of access points on each section. The analyzed area is 12.84 x 10.34 km^2.

We observed that there are 2.18 APs/km^2 on average. Since the rectangles take into consideration areas outside the city, we calculate the density considering only regions with at least one access point. In this case, the average density of access points was 3.9 APs/km^2. Figure 4.1 shows the map of Paris with 200 x 200m grid cells. Each cell indicates access points density. Darker cells indicate regions with more hotspots available. Since there are few public access points, most of cells have 1 to 5 access points.

4.1.2 WISP and Private Hotspots

We implemented an application to Android device to collect and to log information about 3G network status and WiFi availability. We gathered data using two different off-the-shelf devices running Android 4.1: a Google nexus smartphone and a Samsung Galaxy Tab 10.1 *tablet*. In the smartphone, we used a SIM card with Internet plan to capture the 3G network information. Although the number of detected access points was slightly different in the devices due to the antenna gains, we averaged the results based on timestamps and positions.

The participants collected the data through several bus routes in Paris, totaling 12 hours of collection between 27 April and 29 May of 2013, and 82km of distance traveled.We gave preference to bus lines with tourist routes (close to sightseeings). In both devices, the application logs network information scanned every 3 seconds. We log the following information:

[1] http://opendata.paris.fr

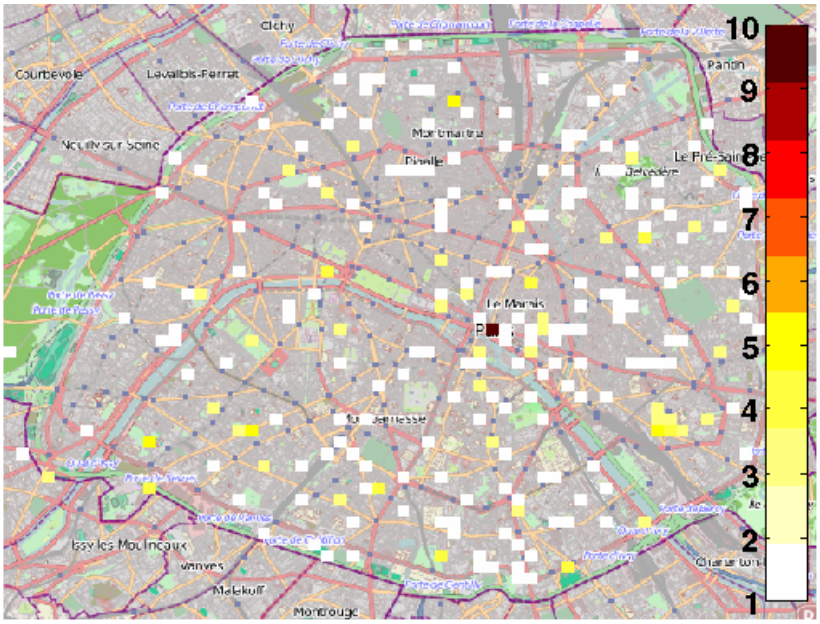


Figure 4.1. Concentration of APs in Paris (200 x 200m grid cells). Darker cells indicate more hotspots available in a region.

- Timestamp
- Latitude and longitude
- Speed
- Communication protocol in the 3G network
- Signal strength of the 3G network in the ASU format[2]
- For each detected access point:
  - MAC address
  - Network name
  - Received signal strength indicator (RSSI)
  - Security protocol
  - Frequency (channel)

[2]The Android Interface provides signal strength information in Arbitrary Strength Unit (ASU). In GSM networks, ASU is equivalent to the received signal strength indicator (RSSI), defined by $dBm \equiv 2 \times ASU - 113$, and ASU $\in [0, 31[$. In UMTS networks, ASU is equivalent to the received signal code power (RSCP), defined by $dBm \equiv ASU - 116$, ASU $\in [-5, 91)$.

Table 4.1. Summary of access point and 3G information.

Number of devices	2
Measurement duration	12h
Distance travelled	82.76km
Distinct Access Points	21640
Private APs	9650 (44.6%)
APs from WISP	11990 (55.4%)
APs from WISP *A*	6038 (27.9%)
APs from WISP *B*	3432 (15.85%)
APs from WISP *C*	2520 (11.6%)
All APs without authentication	27.7%
WISP APs without authentication	17.1%
Average APs detected per scan	24.7
Average distance an AP is in range (m)	52.35
Average time an AP stay in range (s)	13.5
Average WiFi RSSI signal (dBi)	-80.1
Average 3G signal (ASU)	18.5

During the collection, 21.649 APs were identified, 55.4% of which being from WISPs. The three major WISPs represent 27.9, 15.85, and 11.6%, respectively of the APs[3]. Table 4.1 summarizes our findings.

4.1.3 WiFi Connectivity

4.1.3.1 WiFi Coverage

Figure 4.2 shows the distribution of APs through bus routes in Paris. We observed an average of 27.5 access points per scanning, with a minimum of one and maximum of 66. This indicates that today's WiFi networks are accessible throughout the city, which does not mean free Internet access in those hotspots.

Similar to the public hotspots, we calculate the density of the private access point per km^2 using the same area. We observed that there are 151.58 APs/km^2. When considering only the cells in the grid where there is at least one access point, the density rises to 676 APs/km^2. This high number of access points could lead to radio interference. The WISP APs density (considering only cells with APs) are 189, 107 and 101 APs/km^2 for WISPs *A*, *B* and *C*, respectively.

To calculate the maximum distance that the devices sense an AP, we summed the distance between all consecutive points that an AP appears in the dataset. We

[3]We purposely omitted the name of the WISPs to avoid marketing issues.

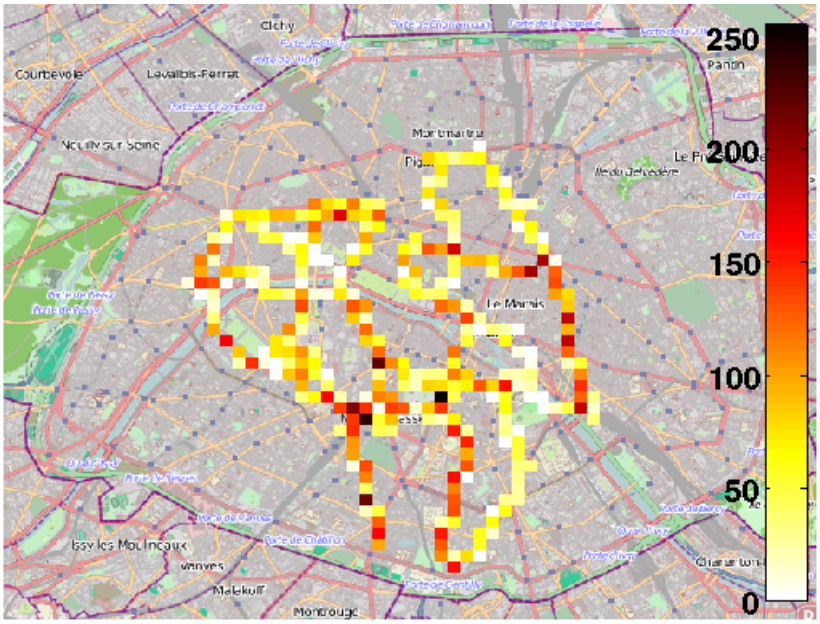


Figure 4.2. Distribution of access points through the measured bus routes.

calculate the Great Circle Distance (GCD) using the *haversine* formula [Sinnott, 1984] as follows. Let $\delta \in \Gamma$ be the set of all (*latitude, longitude*) in which an AP α appears. The distance D of an AP α is calculated by:

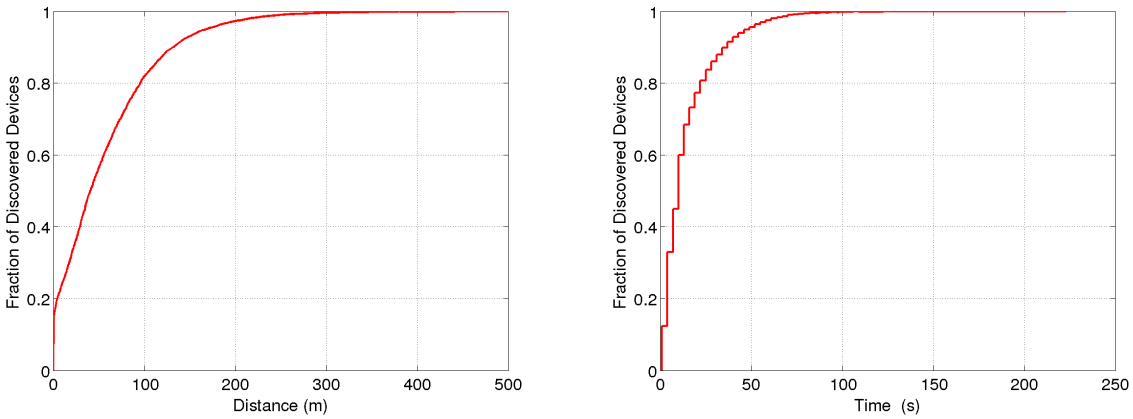
$$D_\alpha = \sum_{i=1}^{|\Gamma|-1} GCD(\delta_{i+1}, \delta_i)$$

Where δ_{i+1} is a scan that happens after δ_i.

In Figure 4.3a, we show the CDF of the distance that a device can move without losing contact with an AP. We observed that on average the mobile client is able to scan an access point for a distance of 60m and 90% of the APs are discoverable above 130m. These results indicate that for mobile clients moving at high-speeds, the maximum connection time should be small. In fact, in [Bychkovsky et al., 2006] the authors show that for speeds in excess of 60km/h a connection with an AP lasts less than 10s.

Since the duration and the distance that a device can discover an AP are small, we analyzed the time without connection with any AP from a given WISP. We found APs from WISPs A, B and C in 98, 98 and 92% of the scans operation. This finding implies that WISPs have coverage in almost the entire bus routes, and data could be offloaded with the existence of an efficient handover mechanism among APs. Looking only the time without connection, the average time (and standard deviation) until finding a new AP was 28.31 (49.55), 27.27 (32.01) and 38.62 (90.33) seconds for WISPs A, B and C,

respectively. The difference of the time without connection between WISPs B and C to WISP A is about 10s, even though WISP C has less than half of APs than A.



(a) CDF of the distance an access point was discovered by the devices

(b) CDF of the cumulative time one access point answered the scanning of the devices.

Figure 4.3. WiFi connectivity Properties.

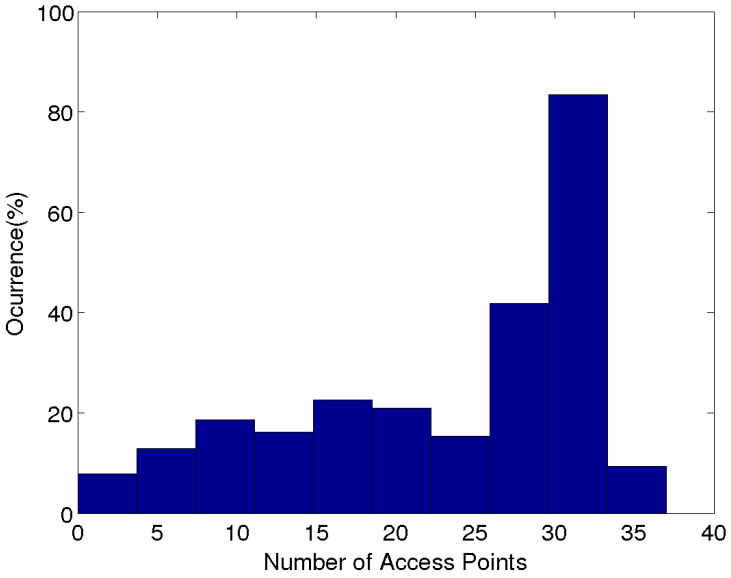


Figure 4.4. Histogram of access points found in each scan operation.

Figure 4.3b plots the CDF of the time that our devices were able to scan each access point. In 90% of the time, we observed that the devices can reach an AP during 40 seconds, thus the user will need to find another suitable AP in less than one minute. This result is expected since buses move at low speeds and there are bus stops, traffic lights and traffic jams in the way.

Table 4.2. Summary of the properties of the AP graph from each WISP.

WISP	**A**	**B**	**C**
No. of connected components	20	273	362
Avg. number of neighbors	22.7	7.65	5.8

Figure 4.4 shows the distribution of the number of access points scanned in each scan operation. The average number scanned APs was 23.19 APs, and we discovered up to 37 APs on one single scan operation. In our experiments, the average and the maximum speed of the bus was of 10.8 and 55km/h, respectively. The number of scanned APs did not vary significantly with speed.

We observed that 27% of the APs do not require link-layer authentication (e.g. WPA2-PSK). However, 57% of them are WISP APs, which means that they are exclusive for customers, and the authentication step will be performed within a proxy. The other APs could also use proxies for authentication, thus we cannot affirm that there is "ubiquitous open WiFi" in the city. The proportion of APs using PSK, EAP and WEP authentication was 45.17, 19.7 and 7.92%, respectively.

4.1.3.2 Graph Topology

Since we observed a dense topology, we analyzed the topology formed by the WISP APs to figure out how connected is the graph of each WISP.

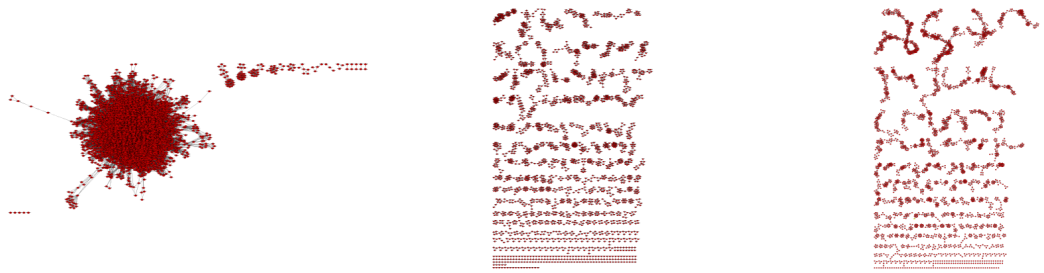
We define the graph $\mathcal{G}_w = (v, e)$ for a WISP w, where the vertices v are APs from WISP w. Two distinct vertices have an edge e, if and only if the APs were scanned at the same time at least once. Figure 4.5 represents the graph of each WISP. Figure 4.5a shows that WISP A has a dense graph with large connected components, while the other graphs are composed by smaller components. Table 4.2 summarizes our findings.

The number of connected components[4] was 20, 273 and 362 for WISP A, B and C , respectively. A low number of connected components represent a more connected network and, as a consequence, more potential to maintain a WiFi connection over time.

The average number of neighbors represents the number of available APs in a region, which indicates the choices a client has when attempting to offload data. More APs are beneficial since the terminals can choose APs with better signal strengths, and also because the load could be balanced among those APs. Although WISP A has larger connected components with a higher average number of neighbors, and WISP C has the lowest number of neighbors, in practice, the time to switch between connected

[4]A connected component is a sub-graph where any pair of vertices have a path between them.

components is similar among all the WISP as discussed in Section 4.1.3.1.



(a) WISP A (b) WISP B (c) WISP C

Figure 4.5. WiFi Graph for each WISP. WISP A shows a high connected component, while WISP B and C show a graph more sparse.

4.1.4 3G Connectivity

We observed that the 3G network was selected in more than 90% of the time, result consistent with the results obtained in [Balasubramanian et al., 2010] in the city of Amherst, Massachusetts in 2010.

In 56% of the scans the cellphone was connected to High-Speed Downlink Packet Access (HSDPA) network, and 36% of the time in High Speed Packet Access (HSPA). In only 5% of the time the connection was over High-Speed Packet Access (HSPA+), which improves throughput up to 168 Mbps, called the 3.5G network. Finally, in 3% of the time the data connection was provided by Enhanced Data rates for GSM Evolution (EDGE) networks, known as a pre-3G network. We note here that the type of network a mobile client will stay connected depends on the availability of the network and the mobile data plan. Table 4.3 summarizes the percentage of the time connected in each type of network (and the theoretical downlink data rate).

Figure 4.6 shows the evolution of number of WiFi antennas and 3G signal strength (0-30 at ASU) in one of the bus routes we scanned. We observed an island of connectivity in some regions, where other regions are sparser. For instance, in the first 1000 seconds there are peaks of 35 or more access points. On other hand, there are some periods with less than 5 access points.

4.1.5 Lessons Learned and Discussion

Due to legal issues, authentication is required to identify the user on the Internet in several countries [Hale, 2004], making the idea of an "open WiFi" network unfeasible. Several WISPs offer customers local ADSL Internet with a WiFi router that comes with

HSDPA (14 Mbit/s)	56%
HSPA (14 Mbit/s)	36%
HSPA+ (84 Mbit/s)	5%
EDGE (1 Mbit/s)	3%

Table 4.3. Percentage of time connected in each type of cellular Network

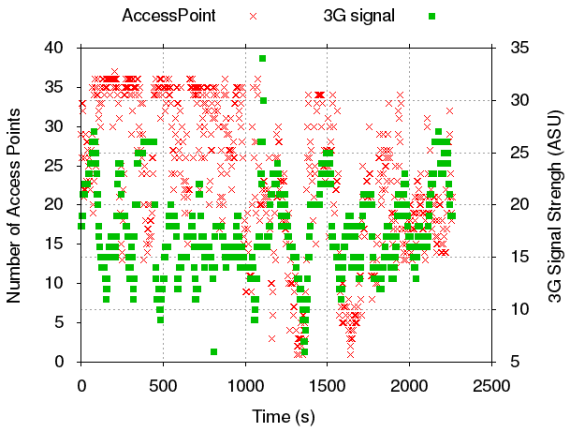


Figure 4.6. Distribution of public hotspots in Paris

two antennas, one for internal use of the client (e.g. home) and another which can be used by other customers of the WISP. In some cases, those antennas use EAP-SIM authentication.

Although our results do not include all possible bus routes, this work gives us insights about WiFi coverage in the city. For instance, if the WiFi density observed is the same in all part of the city, the city is almost entirely covered by WiFi APs. Therefore, WISPs can explore better these already deployed APs. In the next section, we discuss the practical problems to apply WiFi offloading.

For the research community, these results can provide a base when setting parameters in simulations for instance, in order to tune the WiFi range. We observed that, on average, a device can find a given AP moving up to 70 meters in low speeds on average. Moreover, the distribution of WISP APs can be used to model network topologies in urban areas.

4.2 Offloading User Generated Data

In order to identify how much data the deployed WiFi could offload, we compared the public Paris Hotspots and the access points scanned in our experiment with thousand

of user-generated data in a Location Sharing Service. In these services, users spontaneously report their position (latitude and longitude) and their impressions about a place.

Our main goal was to analyze if the WiFi deployed in the metropolitan area is able to offload the data traffic of mobile clients from the 3G network. We attempt to answer the following question: Is the user always close to a WiFi AP?

We assume that if a user is closer than a certain distance of a hotspot, the mobile client can use WiFi instead of 3G. We considered all antennas as public hotspots, which represent the maximum data offload based on our map. After, we observed the data offload capacity of WiFi access points from each WISP.

4.2.1 User Localization Database

Cheng *et al* made a fine-grained characterization of human mobility pattern after analyzing 22 million *check-ins*[5] of more than 22,000 users during September 2010 and January 2011 from several location sharing services, such as Foursquare [Cheng et al., 2011]. The authors characterized the mobility pattern of the users, observing that it follows the Levy flight mobility pattern [Cheng et al., 2011].

We used this database to interpolate the localization informed by the users and the positions of APs in Paris. We separated three sets from the database: *i*) the entire set of check-ins conducted in Paris. *ii*) the check-ins conducted in two weeks in august, which represent the holidays, and *iii*) the check-ins made in a week of December, including the Christmas day and new year's eve.

Interval	Nbr. of Users	Nbr. of Checkins
September/10 to January/11	3203	45696
August	367	2045
December	239	1045

Table 4.4. Number of users and *check-ins* in each period

4.2.2 Evaluation

In our evaluation, a device can offload through WiFi if he/she is closer than a distance ρ from a suitable AP. We compare each coordinate informed by the user with our

[5]We call *check-in* the act of a user publishing his/her GPS position in a public system.

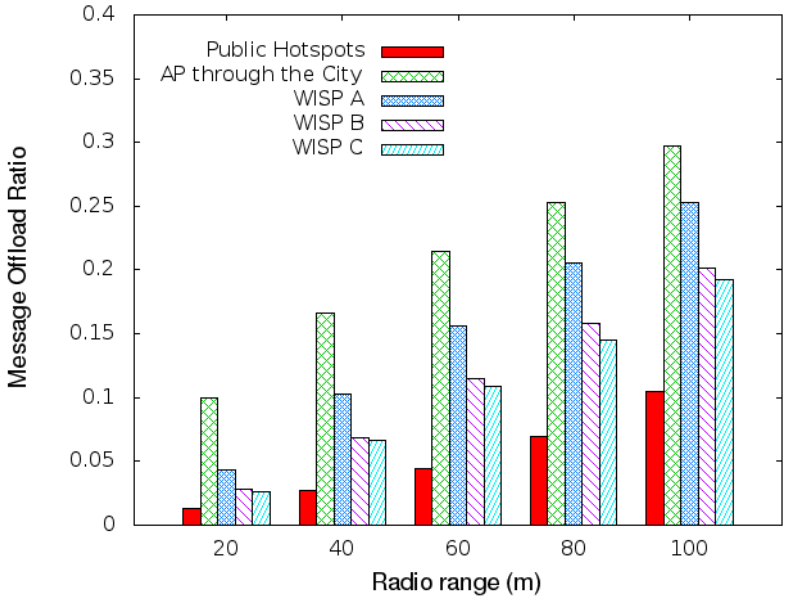
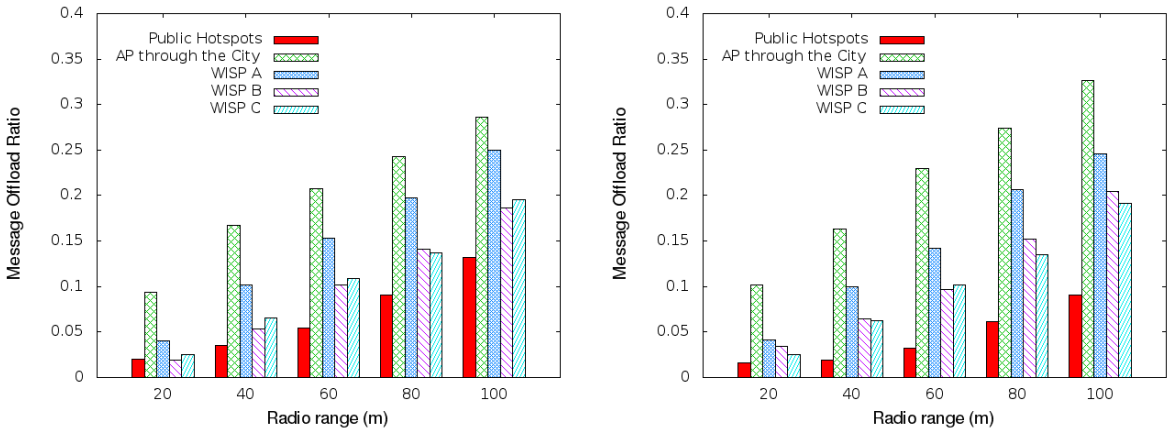


Figure 4.7. Amount of data offloaded using the entire database.



(a) One week in August

(b) One week in December

Figure 4.8. Amount of data offloaded for the different periods.

list of APs. The distance among the *check-in* and the APs was calculated using the *Haversine* formula [Sinnott, 1984].

We compared the checkin datasets within the following scenarios: The public hotspots; all WiFi AP collected through the bus routes; and the subset of APs from WISPs, named accordingly to Table 4.1.

Figures 4.7 and 4.8 show the results of the data offloaded in our scenarios, varying the radio range. Only the public hotspots from the government, which represent the most widely spread dataset of access points, could offload up to 10% of the total traffic

(Figure 4.7) with a range of 100m. In the best scenario, if all access points gathered were open and available for mobile clients, would be possible to offload from 10 up to 30% of the network traffic. Note here that we considered all *check-ins* in the region of Paris, while the access points collected represent 8% of the region (grid region in the Figure 4.2).

Figures 4.8a and 4.8b represent two distinct weeks in the year. In August, in the middle of the holidays, the data offloaded by the public hotspots rises from 10 to 13%. We observed a similar behavior on the traffic offloaded by all the access points gathered and the WISPs in all periods of the *check-in* database.

The offload capacity of the WISPs is shown in the Figure 4.7. Only the WISP A, which represents 28% of the APs, could offload 71% of the traffic offloaded by all APs, with a range of 60m and up to 86% when considering 100m of range. Since the distribution of APs is similar between the WISPs, they show similar results in all scenarios.

Based on the density of AP in the routes gathered and assuming that the rest of the city has the same WiFi AP density, WiFi could lead to a better result in the traffic offloading of the 3G network.

4.2.3 Towards WiFi Offloading

Although the offload of traffic to WiFi is expected to be over 46% of the traffic in 2017 [Cisco, 2013], there are some struggles in the feasibility of mobile WiFi offload in off-the-shelf devices and deployed APs. We discuss below each of these issues:

- *Association Time*: The time required for association in WiFi is about 8 -15 seconds [Bychkovsky et al., 2006] . This time can be greater than the time a device is kept in the range of an antenna. Moreover, we observed that a device stays up to 130s in the range of an AP, thus it has less than 2 minutes to transfer data before becoming out of range.

- *Authentication:* Mobile clients must authenticate using WEP/WAP, and in open WiFi hotspots there are usually proxies for authentication. Besides the association time, this step is time consuming and could force the mobile client to stay in the 3G network.

- *Handover between WiFi antennas:* Once a device gets out of the range of a WiFi antenna, it must decide between reconnecting to the 3G network or connecting to a new WiFi antenna. In both cases, there is a delay in the process and a new

authentication step could be required. Today, there are pre-handoff mechanisms, which can observe when a signal is becoming weak and try to find a new antenna in advance [Lampropoulos et al., 2005]. However, these mechanisms are not usual in off-the-shelf devices.

4.3 Conclusion

This chapter characterized the WiFi and 3G connectivity in Paris. Our main goal was to evaluate the potential of WiFi offloading in the bus routes of the city using APs deployed by the WISPs. We obtained 3G coverage in 92% of the time, while the WiFi coverage by the WISPs achieved up to 99% of the time in 82km of routes.

Since open WiFi is not a reality, we show that the WISPs play an important role in WiFi offloading. They can offload almost 30% of the traffic on the evaluated dataset. WiFi coverage by the WISP is greater than governmental APs. Thus, the WISPs should offer incentives for their customers to use their APs, supporting others customers to use it.

We observed that time and distance that a device stays in the range of an AP is low, even at low speeds, we believe that the main bottleneck for efficient WiFi offloading is the excessive time required for association and authentication. The use of an automatic authentication method such as EAP-SIM could increase the amount of traffic offload from 3G to WiFi.

However, in crowded environments, e.g. as in popular events, even the WiFi may not support thousands of users connected simultaneously. In the next chapter, we propose a solution using device-to-device mobile opportunistic communication to offload the 3G network.

Chapter 5

OppLite: An Opportunistic Mobile Data Offloading Framework

In this chapter, we propose *OppLite*, a multi-criteria decision-making framework based on utility theory [von Neumann and Morgenstern, 1953], which allows to switch between infrastructure and opportunistic modes based only on local decisions performed in the devices. Utility theory allows OppLite to map a utility function into a set of properties, allowing to quantify the user preferences when switching modes. OppLite gathers a set of network properties to decide when to use WISP infrastructure or opportunistic communication.

The rest of this chapter is organized as follows: Section 5.1 introduces OppLite and justifies the use of opportunistic networking. Section 5.2 describes the OppLite Framework and its components. Finally, concluding remarks are given in Section 5.3.

5.1 Motivation

Existing offloading solutions require special routers or new deployments, which imply significant changes in the infrastructure. We focus on traffic offloading relying on device-to-device opportunistic communication.

As mentioned previously in Section 1.4, opportunistic offloading can be used to move off data traffic from congested cells. For instance, in large agglomerations (concerts, sports, etc), where thousands of people try to connect to the infrastructure network, devices with poor connectivity may avoid highly loaded antennas forwarding their data through devices with better connection. Furthermore, relay devices may act as a cache, providing content to other devices opportunistically, saving WISPs bandwidth.

Hence, delay sensitive applications, such as VOIP, may be offloaded through WiFi and cannot be offloaded using D2D opportunistic communication since it is not possible to guarantee its delay requirements. Non-delay sensitive applications, such as participatory sensing applications and software updates, may take advantage of both WiFi and opportunistic offloading.

In this chapter, we propose *OppLite*, a framework that uses the number of neighbors, the battery lifetime and the signal strength as criteria to make the decision of sending a message or request content directly to the infrastructure or using Opportunistic communication.

One should note that here may apply issues such as: privacy - users may not want the data passing through an unknown device; and security - the existence of malicious and selfish nodes. These issues are out of scope of this work and we left their impact as future work.

5.2 OppLite Framework

This section describes our framework, all criteria considered in our decision algorithm, and the employed utility and aggregation functions. In a simple communication model, the mobile devices can be in one out of three modes: *standard*, *opportunistic* and *relay*. In *standard* mode, all data traffic is forwarded between the device and the infrastructure directly. In *opportunistic* mode, devices send or receive data through devices in *relay* mode preferably. In *relay* mode, devices communicate directly with the infrastructure and assist devices in *opportunistic* mode, relaying their data or working as a cache to them.

5.2.1 OppLite Framework

OppLite observes the radio interface in order to monitor the network surrounding the user, such as number of neighbors. Figure 5.2 illustrates the architecture of our framework. After gathering information, the *communication-mode management module* applies a utility function for each criterion observed and aggregates the results of all utility functions. From the *standard* mode, a device can switch to *relay* or *opportunistic* mode depending on the results of the aggregation of the utility functions. The decision is based on user-defined thresholds for both cases.

The *User Profile* defines the weight of each criterion and delay tolerance for the running applications. Due to the ephemeral nature of network events, such as watching

a video or downloading a software update, the profile allows OppLite to decide in real time whether an application can be offloaded opportunistically or not.

Table 5.1 summarizes all input parameters of OppLite, which are described in the next sections.

The applications on top of OppLite can also use information provided by OppLite to define whether they can be offloaded or not. For instance, an application can define that Opportunistic devices must seek for some content in the message buffer of relay devices. Furthermore, messages created by the applications are forwarded in accordance with those decisions through the *forwarding module*.

Table 5.1. OppLite Parameters.

Parameters	**Definition**
$x_i = \{c_i, s_i, w_i\} \in X$	Parameters of the utility functions for each criterion, where c_i, s_i and w_i represent the center value, the steepness and the weight of the criterion $x_i \in X$.
Γ_{relay}	Threshold that the utility function must overcome to switch a device to the relay mode.
Γ_{Opp}	Threshold that the utility function must overcome to switch a device to the opportunistic mode.
τ	Delay tolerance of a device.

OppLite can be used for two network offloading use cases: signaling offloading or data offloading. First, to achieve signaling offloading, OppLite nodes in relay mode can forward all messages to opportunistic nodes. In this case, opportunistic nodes avoid to use the infrastructure, saving signaling required to become connected. Figure 5.1 illustrates the forwarding scheme. When an application with delay tolerance $\tau \geq 0$ creates a message M at time T, if the node is opportunistic, it will attempt to send M to a relay node. If there are no available relays in its neighborhood, then the sender will wait at most $T' - T < \tau$ seconds to find a suitable neighbor, where T' is the current time.

After this, the message will be delivered using the infrastructure.

In the second case, OppLite's message buffer becomes a cache. A node in relay mode assists the infrastructure network by letting opportunistic nodes obtain contents from its cache. Therefore, saving data transmission from the infrastructure.

In both cases, in this work, we considered only one-hop between nodes in relay and opportunistic mode. To achieve multi-hop offloading, nodes must be in relay and opportunistic mode at the same time. These nodes could request content as

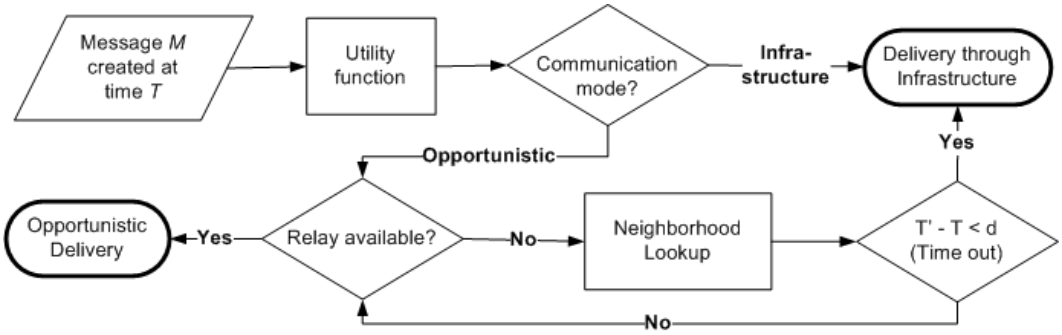


Figure 5.1. OppLite framework forwarding algorithm for signaling offloading.

opportunistic nodes and provide content as relay nodes. In Chapter 6 we propose three applications to run on top of OppLite to evaluate its signaling and data offloading feasibility.

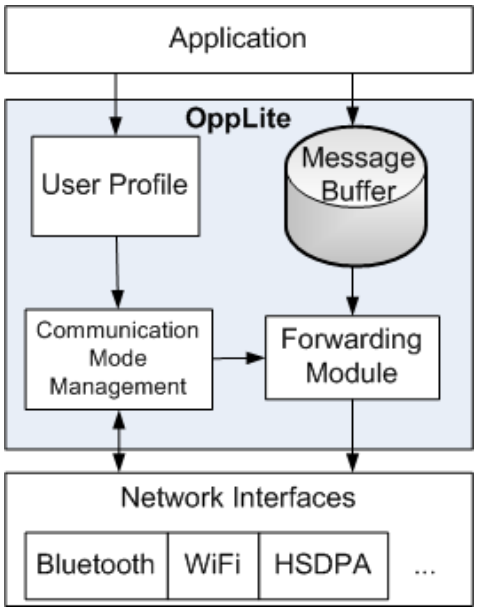


Figure 5.2. Architecture of OppLite framework.

5.2.2 Opportunistic criteria

In our work, we considered only the criteria monitored by the devices. We assume a utility function $u(x) \in [0, 1]$, where x is an upward criterion. Thus, a downward criterion x is modeled by the function $1 - u(x)$. We describe below the set of criteria considered in this work:

5.2.2.1 Number of neighbors

This criterion, gathered by the network interfaces, indicates the feasibility of using D2D to offload the network. For instance, if there are no neighbors available, it is not possible to create a *Pocket Switched Network*. On the other hand, a high number of neighbors may indicate a crowded medium, and a *pocket switched network* can be used to offload traffic from the infrastructure.

To obtain the number of neighborhood, we assume each node sends a broadcast beacon inquiring for neighbors. Nodes in the range of the beacon answer it.

5.2.2.2 Battery lifetime

This is an upward criterion for the decision for a node to become a *relay*. Moreover, the battery lifetime is a downward criterion for switching to *opportunistic communication*, since users may want to elongate the battery lifetime by sending messages opportunistically.

Balasubramanian *et al.* observed through experiments with Nokia N95 smartphones that downloading data in 3G networks consumes up to six times more energy than in WiFi [Balasubramanian et al., 2009]. Moreover, the authors showed that time between transfers significantly affects energy consumption because of the energy to start a transfer and the time the device keeps in a high state after a data transfer (tail energy). Furthermore, the authors proposed an empirical model for energy consumption. Equations 5.1 and 5.2 summarize their model for 3G and WiFi, respectively, where x is the data size in KBytes. This model is employed in our simulations to estimate the energy consumption.

$$E_{3G}(x) = 0.025(x) + 11.25J \tag{5.1}$$

$$E_{WIFI}(x) = 0.007(x) + 5.9J \tag{5.2}$$

Another power modeling was proposed in [Huang et al., 2012], where authors further modeled WiFi, 3G and LTE power levels. Huang et al. designed a network performance tool for Android devices, which allowed them to evaluate the performance of wireless networks with more than 3000 users within two months. Furthermore, they analyzed energy consumption of 20 mobile 4G (LTE) devices (Motorola Atrix or Samsumg Galaxy S) and observed that a linear model fits well for both uplink and downlink for all network interfaces. Thus, the power level (mW) for uplink, downlink and for simultaneous transfers is given by the equations 5.3, 5.4 and 5.5, respectively,

where t_u and t_d are the throughput for uplink and downlink and β is the tail energy. The tail energy represents the highest energy consumption for LTE and 3G networks, rather than energy consumed in data transfer in these networks. The best fit parameters are shown in Table 5.2.

$$P_u = \alpha_u t_u + \beta \tag{5.3}$$

$$P_d = \alpha_d t_d + \beta \tag{5.4}$$

$$P = \alpha_u t_u + \alpha_d t_d + \beta \tag{5.5}$$

Table 5.2. Best fit Parameters for power level in off-the-shelf devices [Huang et al., 2012].

	α_u (mW/Mbps)	α_d (mW/Mbps)	β (mW)
LTE	438.39	51.97	1288.04
3G	868.98	122.12	817.88
WiFi	283.17	137.01	132.86

5.2.2.3 Link quality

The received signal strength indicator (RSSI) measures the link quality, and is an upward criterion to a node become a *relay* and a downward criterion to a node become *opportunistic*. In Chapter 4 we showed that RSSI values follow a normal distribution.

5.2.3 User Profile

The User Profile module stores user-defined weights for all criteria. A weight vector v is defined, such that $v = \{w_i \in [0,1], \sum_{\forall i \in v} w_i = 1\}$, where w_i is the weight of criterion x_i [Nguyen-Vuong et al., 2008]. This vector will be employed in the aggregation function described in the next section.

5.2.4 Utility Function for Single Criterion

We use the well-known sigmoid utility function presented in Equation 5.6, where x_t is the center of the curve in the sigmoid function and $\alpha > 0$ represents the steepness

parameter (i.e. how much criterion $u(x)$ is sensitive to variations of x. The advantage of the Equation 5.6 compared to other sigmoid functions proposed in the literature, as in [Nguyen-Vuong et al., 2008; Abid et al., 2012], is that it does not need user-specific upper and lower bounds.

$$u(x) = \frac{1}{1 + e^{\alpha(x_t - x)}} \tag{5.6}$$

This utility function $u(x)$ defines a utility $u(x) \in [0, 1]$ for an *upward* criterion x. If x is a *downward* criterion, then $u'(x) = 1 - u(x)$ is used instead.

The major challenge when using sigmoid functions is tuning its parameters. Larger values of α give a more exponential behavior to the curve, while $\alpha \in]0, 1[$ presents a more linear behavior. Figure 5.3 depicts different behaviors of Equation 5.6 for different values of x_t and α.

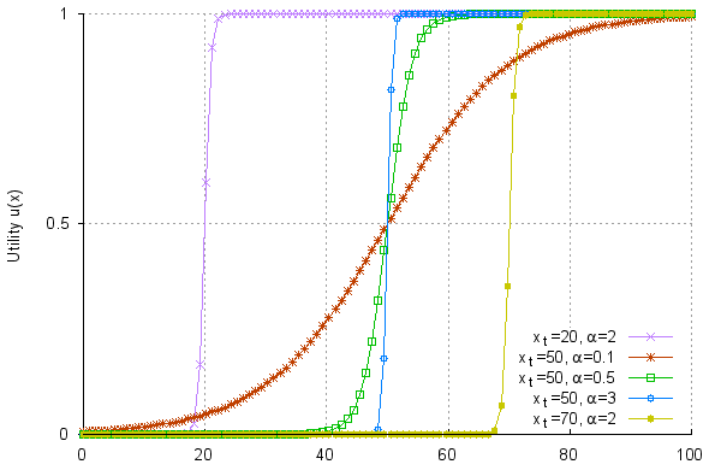


Figure 5.3. Behavior of different values in Equation 5.6.

5.2.5 Multi-Criteria Aggregation Function

OppLite aggregates all single-criterion utility value using a weighted product model to obtain the utility function U, as shown below:

$$U(x_1 \dots x_n) = \prod_{i=1}^{n} [u(x_i)]^{w_i}$$

Where w_i is the preference for the criterion x_i and $\sum_{i=1}^{n} w_i = 1$.

5.2.6 Decision Algorithm

The nodes take decisions among three alternatives: become a *relay* node, become an *opportunistic* node or resort to be a *standard* node. Given the multi-criteria utility function, the communication mode management module decides if the device should switch to opportunistic networks, thus offloading the infrastructure. The decision algorithm runs periodically in the device. In opportunistic mode, since communication disruptions can occur between a node and one relay, the node may wait up to a predefined threshold to contact a relay. Otherwise, the node switches back to the *standard* mode. This threshold avoids the ping-pong effect, where the device frequently alternates between *opportunistic* and *infrastructure* modes.

Formally, using the graph model presented earlier, a node x communicates opportunistically with a node y, if x is in the *opportunistic* mode and x is in the *relay* mode and, $\exists(e(x,y) \in E')|y \in \mathcal{S}$, where $E' = (E_t, E_{t+1}, \cdots, E_{t+\tau})$ is the set of all links created between the time t and $t+\tau$.

OppLite implements the two decision algorithms presented below.

5.2.6.1 Decision 1 – should the node become a relay node?

Relay nodes are those that forward messages from nodes operating in opportunistic forwarding. A node may become a relay based on three criteria: battery lifetime, number of neighbors and link quality. For instance, a node may elect itself as relay if it has an energy level greater than 70%.

A node becomes a relay if the aggregated utility $U_r(X) > \Gamma_{relay}$, where Γ_{relay} is a configurable threshold. Since a node does not influence the decisions of others, more nodes than necessary may become relays. Our algorithm does not attempt to produce the minimum set of relays: more relays increase the performance of nodes operating in opportunistic mode.

Figure 5.4 illustrates the value of $U_r(X)$ when the battery lifetime and the number of neighbors are considered as upward criteria, using the parameters for center and steepness summarized in Table 5.3. By choosing a low steepness and a high center value for battery lifetime, nodes may become relays even when the remaining battery levels are low.

5.2.6.2 Decision 2 – should the node use opportunistic mode?

To calculate the utility function $U_m(x)$ to switch the operating mode to opportunistic, a node will consider the number of neighbors and the inverse of the energy level. There

Table 5.3. Criteria values.

Criterion	Center	Steepness	Weight
Neighbors	3	2	0.4
Battery Level	70	0.2	0.5
Link Quality	15	2	0.1

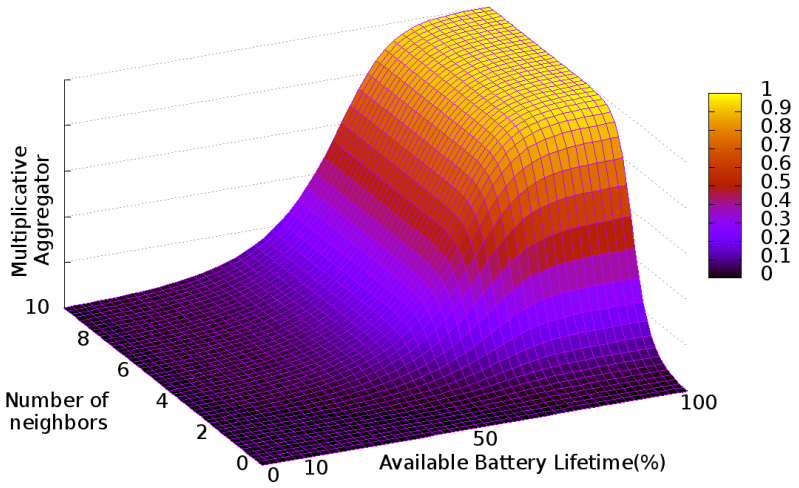


Figure 5.4. Variation of $U_r(X)$

are two reasons for this choice.

First, opportunistic forwarding typically spends less energy than infrastructured forwarding. Second, more neighbors are an indication of a more congested infrastructure. If $U_m(X) > \Gamma_{opp}$, where Γ_{opp} is a configurable threshold, the node will switch to opportunistic mode.

A node comes back to *standard* mode if there are data to be sent or some request is not answered during a delay tolerance τ, even if their utility function is lower than Γ_{opp}.

5.3 Conclusions

In this chapter, we proposed a lightweight framework to offload the traffic from cellular networks using opportunistic forwarding among mobile devices. OppLite is device-driven and, as such, avoids modifications in the infrastructure.

We modeled the offloading problem as a multi-criteria decision problem, where all inputs are collected by the user devices. Due to the delay requirements of the applications, OppLite defines a maximum delay threshold for opportunistic communi-

cation, returning to infrastructure mode when a non-delivered message achieves this delay threshold.

In the next chapter, we implemented three applications on top of OppLite to evaluate its feasibility to *i*) diminish the number of devices connected to infrastructure, saving signaling and *ii*) to provide content to opportunistic devices through the relay devices instead of using infrastructure, saving data consumption.

Chapter 6

OppLite Evaluation

In this chapter, we evaluate OppLite through trace-driven simulations. We implemented three applications on top of OppLite framework to assess its feasibility to *mobile signaling offloading* and *mobile data offloading*.

The first application, called *Opportunistic Relaying* (OpR), targets *mobile signaling offloading*. In OpR, all nodes in the opportunistic mode attempt to forward their data messages to relay nodes instead of sending them directly to the infrastructure. In this case, OppLite aims to reduce the number of devices connected in the cellular network.

The next two applications target *mobile data offloading*: *i) Cache-and-Forward*, in which relay nodes forward all content received to all opportunistic nodes encountered within a threshold delay (time-to-live). *ii) Relay-as-Cache*, in which opportunistic nodes seek for a determined content in the relay nodes cache until a delay tolerance instead of requesting these contents to the infrastructure.

The rest of this chapter is organized as follows: Section 6.1 describes the applications implemented to evaluate OppLite. Section 6.2 presents our evaluation methodology. Section 6.3 analyzes the traces we used in the simulations. The results and performance evaluation are discussed in Section 6.4. The concluding remarks are given in Section 6.5.

6.1 Application Scenarios

OppLite is a decision-making framework to select the communication mode of the nodes. We developed three applications on top of OppLite to measure its capabilities to offload signaling and data traffic in the network.

The first application represents use case 1, shown in Figure 1.2 in Chapter 1,

where opportunistic nodes are under low speed or congested network. In this case, opportunistic nodes leave the infrastructure network to attempt forwarding their messages through relay nodes. The next two applications represent use case 2, shown in Figure 1.3, where OppLite offloads data from the infrastructure network using relay nodes as caching mechanism. We detail each application below.

Opportunistic Relaying (OpR): This application represents a participatory sensing networks, in which devices gather data (using their embedded sensors) and forward these data to the "cloud". This kind of application sends small packets to the network constantly. Although the amount of data is small, several devices sending packets continuously may cause signaling congestion [Choi et al., 2014].

After a message creation, OpR forwards this message to OppLite. OppLite chooses how to forward this message based on the communication mode of the node. A node in opportunistic mode attempts to send the message through a relay node. In case the opportunistic node does not find a suitable relay node up to a tolerable delay (τ), OppLite sends this message directly to the infrastructure.

Cache-and-Forward (CaF): This application models HTTP request applications, in which users request content to the cloud. A node sends a request with an identification of the desired content and the cloud responds with the content requested.

CaF requests the communication mode to OppLite. In *CaF* only nodes in the *standard* or *relay* mode request content. OppLite buffers the content setting the timestamp it received. A node in the relay mode forwards the buffered content to all opportunistic nodes encountered up to a time-to-live defined by the parameter τ. Finally, OppLite drops from its buffering all content with the time-to-live expired.

A cache and Forward architecture for routers with large capacity was proposed in [Paul et al., 2008]. We consider users' devices as potential caching agent in the network.

Relay-as-Cache (RaC): *RaC* also models an HTTP request application. In *RaC* nodes in all modes are eligible to request content.

RaC sends the request messages containing the identification of desired content to OppLite. In case the nodes are in *standard* or *relay* mode, OppLite forwards the request messages to the cloud through the cellular network. In its turn, the cloud responds with the content requested. Nodes in relay mode store all content

received in their buffers. Therefore, relay nodes act as caching nodes, accepting requesting messages from other nodes.

An opportunistic node periodically broadcasts its request message searching for relays nodes with the desired content cached. This process repeats until some relay response or up to maximum delay tolerance defined by τ. After this tolerance delay expires, OppLite uses the infrastructure network to send the request message.

A *hit* occurs when a relay node forwards to the opportunistic node the content requested. On the contrary, a *miss* occurs when the relay does not have the content required. In consequence, the *hit-rate* of a request message is calculated as $\frac{hit}{hit+miss}$.

6.2 Methodology

OppLite can be implemented on top of TCP/IP stack or as clean-slate protocol. We implemented OppLite as a clean slate protocol in a simulation environment. Therefore, OppLite can access information from lower levels, such as 3G network signal status, neighborhood and remaining battery. In our simulations, we considered nodes own a 802.11 and 3G cellular communication interface. Nodes running OppLite can communicate directly among them using the 802.11 interface in ad hoc mode with no encryption in the link.

To evaluate our proposal, we measured the number of nodes selected as relays and as opportunistic for each scenario evaluated. Next, we evaluate the maximum signaling and data offloading OppLite achieves using metrics accordingly to the simulated application.

In *OpR* application, OppLite aims to reduce the number of nodes sending their messages directly to the infrastructure. Thus, we measured the ratio of *message sent through the cellular network*.

In *CaF* and *RaC* applications, OppLite aims to reduce the use of the infrastructure to request content already downloaded by a set of users. In common for both applications, we quantify the number of *messages Delivered Opportunistically* as the offloading ratio below:

$$OffloadingRatio = \frac{\text{Delivered Opportunistically}}{\text{Delivered Opportunistically} + \text{Delivered by Infra}}$$

Exclusively in *CaF* application, we further analyzed the *average number of messages forwarded* by each relay node.

In *RaC* application, we defined a small, medium and large set of content available to download. We analyzed how the size of this set impacts the offloading ratio. We further analyzed the hit-ratio for each set of content available to download. The maximum hit ratio a cache mechanism may reach is defined by the cacheability of the network [Ager et al., 2010]. Cacheability is given by Equation 6.1, where n is the number of unique content requested at least once and k_i denotes the number of times content i was requested.

$$Cacheability = \frac{\sum_{i=1}^{N}(k_i - 1)}{\sum_{i=1}^{N}(k_i)} \tag{6.1}$$

To assess OppLite, we compare its results against two strategies: *Random* and *Oracle*. In the first one, relay and opportunistic nodes are chosen randomly. In *Random*, a device becomes relay if $X > \Gamma_{relay}$, and opportunistic if $Y > \Gamma_{opp}$, where $X, Y \in [0, 1]$ are random variables following a uniform distribution, Γ_{relay} and Γ_{opp} are configurable thresholds as in OppLite. The remainder of the decision algorithm is identical to OppLite. Furthermore, *Random* algorithm selects nodes beside their willingness to become a relay. Thus, *Random* works as a benchmark to evaluate relay node selection.

In the second strategy, *Oracle* selects a subset of nodes to act as relay using a global view of the network. *Oracle* creates a graph $G = (V, E)$ accordingly to the trace used, where vertices represent nodes and edges represent existing links between nodes. Then, it calculates the minimum dominating set of the graph each time a message is created. The nodes in the dominating set act as a *relay*. The transmission delay is not considered in this strategy.

In order to refine our evaluations, we used two real traces of network contacts. We characterized the graph properties of each trace to understand the effect of these properties in our evaluation. Furthermore, we implemented trace-driven models for traffic, content popularity and network quality.

6.2.1 Simulations

We used the ONE simulator [Keränen et al., 2009] to evaluate our proposal. In the ONE simulator a message is either totally delivered or not, without fragmentation. We extended the simulator by implementing the energy consumption model presented

in [Balasubramanian et al., 2009], where the transmission cost depends on the size of the data, as well by adding support for infrastructure networks. For simplicity, we assume that all users have devices equipped with batteries of 1500mAh and output voltage of 3.7V, which are typical values for popular smartphones.

Table 6.1. Criteria Parameters

Criterion	Center	Steepness	Weight
Number of Neighbors	0.3	2.0	0.4
Battery Level	30	0.1	0.5
Link Quality	15	2.0	0.1

6.2.2 Parameters

We assume that each device has two interfaces: a WiFi interface for ad hoc communication with a bit-rate of 2Mbps, which is compatible rate of IEEE 802.11, and a cellular network interface with 100kB/s (according to real bandwidth measurements [Whitbeck et al., 2012]).Furthermore, nodes have 500MB of disk space (available for buffering). Each message has between 100Kb and 1Mb of payload, representing text and images. The initial battery level of each node is uniformly distributed in the range $[40, 90]\%$, and link quality measurements change every 10s using the normal distribution (due to our findings in [Mota et al., 2013]).

The set of criteria in OppLite was configured as described in Table 6.1. Furthermore, we evaluate OppLite and Random performance by combining the parameters threshold relay (Γ_{relay}), threshold opportunistic (Γ_{opp}) and delay tolerance (τ) with the following values:

$$\Gamma_{relay} = [0, 0.2, 0.5, 0.7]$$
$$\Gamma_{opp} = [0.01, 0.2, 0.5, 0.7]$$
$$\tau = [1, 100, 600, 1200]\text{s}$$

The threshold opportunistic begins with 0.01 to avoid OppLite switches nodes without neighbors to opportunistic mode. We run each simulation scenario 15 times, and show the results with 95% confidence intervals.

6.2.3 Traffic Model

To increase the reality of our simulations we investigate the inter-message generation time distribution in a real dataset. We analyzed a dataset of 22 million *checkins* during

six months in several location sharing services such as Foursquare provided by [Cheng et al., 2011].

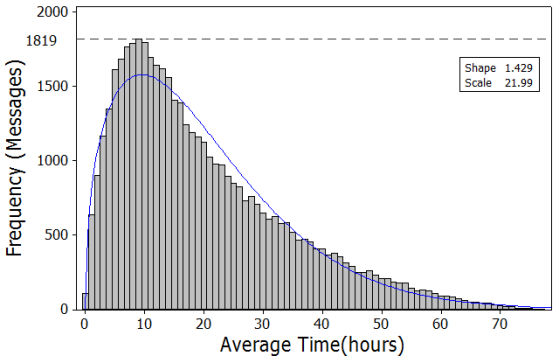


Figure 6.1. Distribution of inter-message generation time.

We observed for users that created more than 100 messages, the message generation interval follows an exponential inter-generation time distribution, as depicted in the histogram of Figure 6.1. We fitted a Weibull curve into the data, with the shape parameter $k = 21.99$ and scale $\lambda = 1.429$, obtaining a goodness of fit $R^2 = 0.994$. Thus, users typically create a new message within 10h of the last message. In our simulations, we employ the fitted Weibull distribution to model the inter-message generation time of each node in the network.

6.2.4 Content Request Pattern

Content popularity is an important metric for caching mechanisms. Several web caching studies have shown that HTTP Requests or video requests follow a Zipf probability distribution [Breslau et al., 1999; Sinha and Pan, 2006; Ben Abdesslem and Lindgren, 2014; Valerio et al., 2015].

Valerio et al. analyzed the probability distribution of 524,787 video requests made by 398,329 users from a European cellular operator [Valerio et al., 2015]. The authors showed that popularity of contents follows a Zipf distribution with parameter $\rho = 1.47$.

Hence, in the *Relay-as-Cache* application the content request pattern follows a Zipf distribution with exponent $\rho = 1.47$

6.2.5 Traces

We evaluate our framework using the following contact traces available in the Community Resource for Archiving Wireless Data (Crawdad) [Crawdad, 2013].

INFOCOM06: In 2005 and 2006 the Haggle Project gathered a set of traces from volunteers in the INFOCOM conference. The authors distributed iMote devices

to 41 and 98 participants, respectively [Haggle, 2013]. We used the 2006 INFOCOM edition since it has a higher number of nodes. INFOCOM06 is a three-day dataset with 20 static nodes and 78 users [Chaintreau et al., 2007]. We used the first day of the conference, from 12:00-18:00. The application stops generating messages at 17:00, ensuring enough time for message delivery.

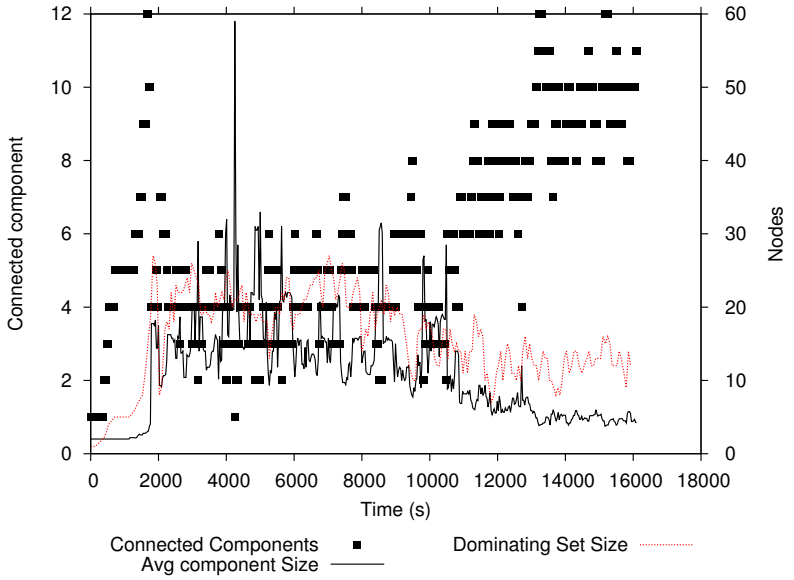
ROLLERNET: The Pari-Roller tour brings together thousands of people to rollerblade through Paris. The Rollernet project distributed 62 iMotes to volunteers in August 2006 among 2500 rollerbladers, logging Bluetooth encounters every 15 seconds [Tournoux et al., 2009]. Since this trace logs every Bluetooth device in range, it has 1112 devices in total. The contact trace has approximately three hours. We use the first 30 minutes as warm up time and stopped message creation after 9500s. The simulation runs until 12500s to ensure message delivery.

6.3 Trace Analysis

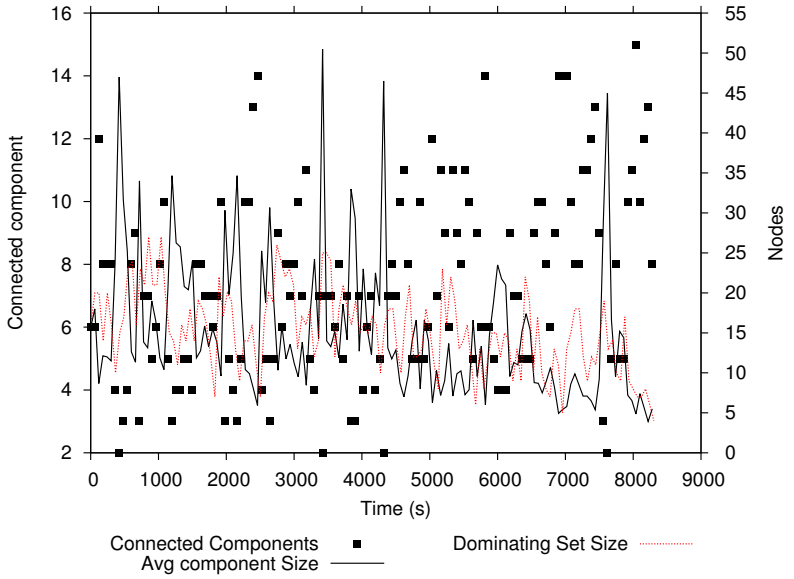
Figure 6.2 shows the number of connected components, the average number of nodes in these components and the number of nodes in the dominating set of the graph over time for both traces.

We observed 5.67 connected components with 11.18 nodes on average in Infocom06 trace (Figure 6.2a). This behavior remains until 10000s approximately, later on the nodes tend to become isolated, which indicates the ending of a conference day. The dominating set has 15.56 nodes and node degree of 3.32 nodes on average. The Rollernet trace has larger connected components (Figure 6.2b), with 6.20 connected components with 20.35 nodes on average. Although the Rollernet trace contains more than 1000 nodes, few nodes remain active in the network. The dominating set contains 15.55 nodes on average.

In this scenario, the behavior of the participants follows the rules defined by the staff of Pari-Roller (e.g. speed, pauses and paths). This causes oscillations in the connectivity when local density is low (participants are rolling) and high when some deceleration happens, in an effect known as the *accordion phenomenon* [Tournoux et al., 2009].



(a) Infocom06 graph properties



(b) Rollernet graph properties

Figure 6.2. Properties of the traces used in our evaluation. The black squares show the number of connected components (left y-axis), while the lines show the average number of nodes in the connected components and in the dominating set (right y-axis).

6.4 Simulation Results

6.4.1 Amount of Relays and Opportunistic Nodes

This section analyzes the effect of the relay threshold (Γ_{relay}) and opportunistic threshold (Γ_{opp}) in the number of nodes in relay or opportunistic mode, respectively. We fixed each threshold $\Gamma_x = 0$ while evaluating the other threshold Γ.

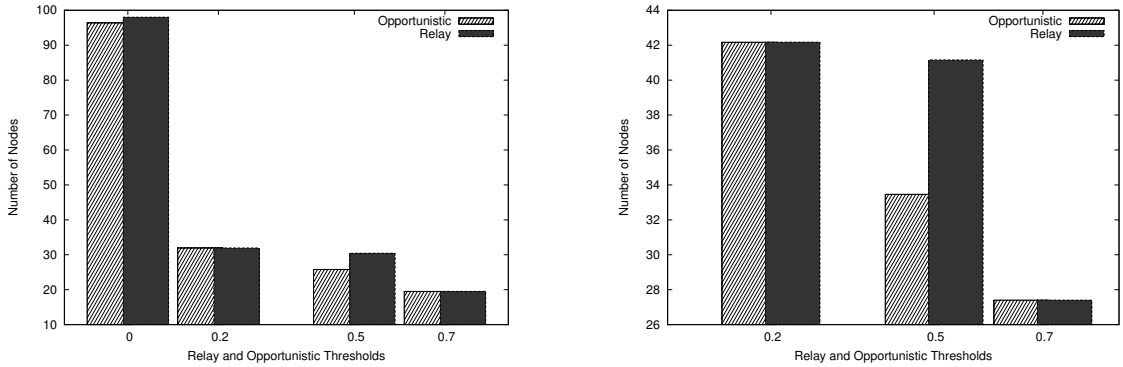
Figure 6.3 shows the influence of Γ_{relay} and Γ_{opp} in the number of nodes in the mode relay and opportunistic for INFOCOM and ROLLERNET dataset.

When $\Gamma_{relay} = 0$, all nodes switch to relay mode, assisting to offload the infrastructure. In turn, when $\Gamma_{opp} = 0$, all nodes switch to opportunistic mode. Higher values of the thresholds tighten the number of nodes selected as relays or opportunistic, since to bypass the thresholds the utility function of OppLite must evaluate the number of neighbors, remaining energy and link quality.

We observed that number of neighbors is the criterion which defines the utility function of OppLite, since the number of neighbors changes quickly ranging from zero up to fifty neighbors in some cases. We noted that in the INFOCOM scenario (Fig. 6.3a), $\frac{2}{3}$ of the nodes do not have neighbors for most of the time, and as such are ineligible to become relay/opportunistic nodes. Similarly, ROLLERNET dataset contains 1112 nodes, however with the increase of the thresholds only 40 nodes become opportunistic or relay, as shown in Figure 6.3b. We omitted Γ_{relay} and Γ_{opp} zero values in 6.3b for the sake of its clarity since in $\Gamma_{relay} = 0$ OppLite selects 1112 nodes as relay.

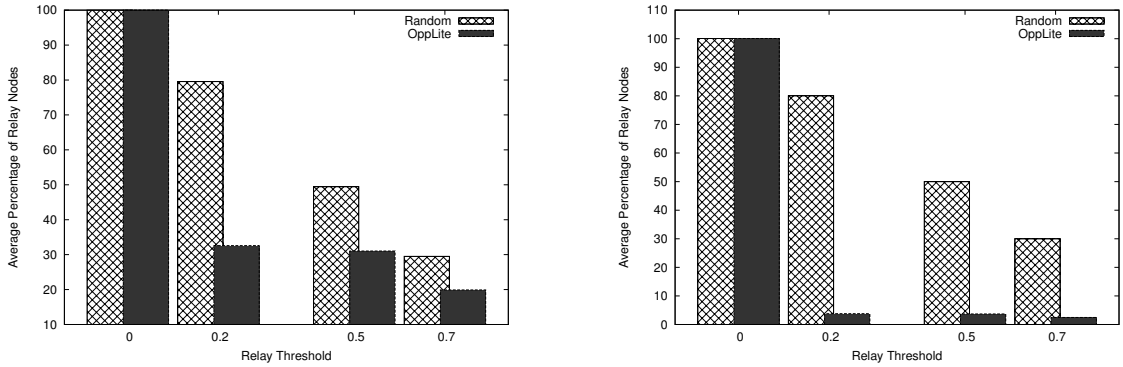
Figure 6.4 shows the comparisons of the amount of relay and opportunistic nodes managed by OppLite against Random solution. In INFOCOM dataset, shown in Figure 6.4a, we observed that in Random, any node can become a relay based only on the probability, e.g. if the threshold is 0.2, then 80% of the nodes switch to relay mode on average, thus Random has a higher number of relays. On the other hand, increasing the threshold in OppLite, only nodes with a considerable number of neighbors, high signal and with remaining battery lifetime become relays.

In ROLLERNET scenario, increasing Γ_{relay} from zero to 0.2 causes a reduction of 1112 nodes acting as relay to 42 nodes, which represent 67% of the nodes that officially belong to the experiment, as can be seen in Figure 6.4b. Finally, the significant number of Bluetooth contacts with devices not belonging to the experiment increases the proportion of nodes acting as relays in Random algorithm, but not in OppLite, since OppLite is tied to the network properties of each node.



(a) Thresholds variation INFOCOM (b) Thresholds variation ROLLERNET

Figure 6.3. Average number of nodes in relay and opportunistic mode based on threshold variation in OppLite.



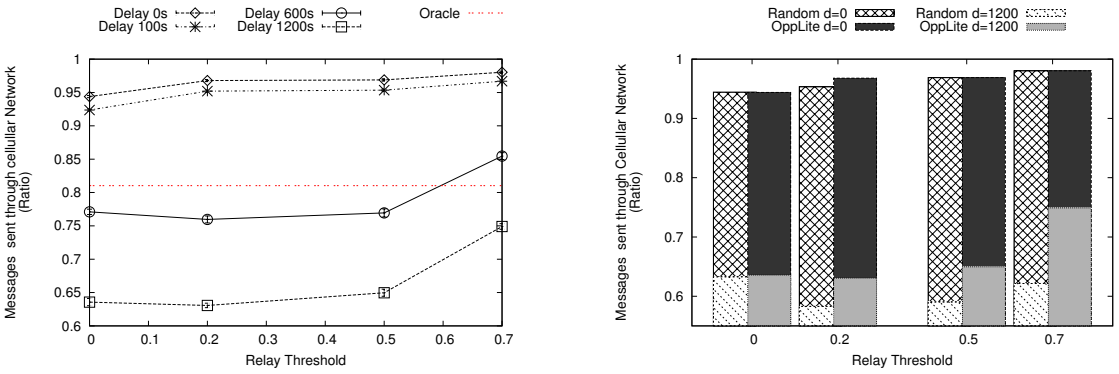
(a) INFOCOM (b) ROLLERNET

Figure 6.4. Comparison of average relay nodes between Random and OppLite.

6.4.2 Opportunistic Relaying (OpR) Evaluation

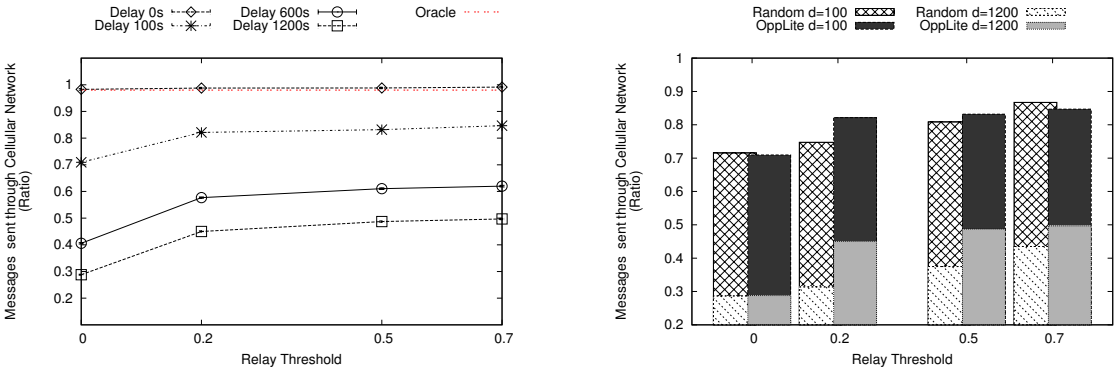
In this application, opportunistic nodes avoid data connection on the WISPs in order to save signaling in the infrastructure. In this case, opportunistic nodes shall forward their messages through relay nodes. In case the opportunistic node fails to find a relay node up to the delay tolerance τ, OppLite forwards the message to infrastructure. We analyzed the number of messages sent directly from the source to infrastructure. In the Rollernet scenario, only the 62 devices which officially belong to Rollernet experiments create messages during our simulation.

In this scenario, offloading occurs every time an opportunistic node gets a message forwarded by some relay node.



(a) Delay tolerance effect (b) Random vs OppLite

Figure 6.5. OpR: Offloaded messages based on the Relay Threshold in INFOCOM dataset



(a) Delay tolerance effect (b) Random vs OppLite

Figure 6.6. OpR: Offloaded messages based on the Relay Threshold variation in ROLLERNET dataset

6.4.2.1 Cellular Networking Utilization

We show the feasibility of our framework for different delay tolerances in Figures 6.5 and 6.6. If the application tolerates delay, it can be offloaded from the infrastructure in both scenarios analyzed. Opportunistic nodes with high delay tolerance have a greater chance to offload their messages.

Here, we emphasize that the *Oracle* solution considers as relay the nodes in the dominating set t belonging to the the graph at instant of the creation of a message. Hence, there is zero delay tolerance for message delivery in the *Oracle* solution. On the other hand, the Random solution gives us an insight of the gains of choosing relays based on the network characterization instead of choosing them randomly.

In INFOCOM scenario, a total of 5700 messages were sent over the network.

Figure 6.5a represents our evaluation for increasing values of Γ_{relay} and fixed $\Gamma_{opp} = 0.01$ with different delay tolerance values. The maximum offload is obtained when an application supports higher delays because opportunistic nodes must wait until making contact with a relay node. In case this contact does not happen, OppLite sends the message directly to the infrastructure. With zero delay tolerance OppLite offloads 6% of messages when $\Gamma_{relay} = 0$ and 3% when $\Gamma_{relay} = 0.7$. All nodes belonging to the dominating set as relays offload up to 20% of the messages with zero delay tolerance in the INFOCOM dataset, but in this case, there are more nodes acting as relays in Oracle than in OppLite approach. Since Oracle considered only the dominating set with zero delay tolerance, OppLite beats Oracle solution when the delay tolerance is greater than 600s with only 28 nodes in relay mode ($\Gamma_{relay} = 0.5$).

Figure 6.5b compares OppLite against Random algorithm for zero and 1200s of delay tolerance τ (for better readability we omitted the other delay values). When $\tau = 1200$ OppLite offloads the same amount of messages than Random after $\Gamma_{relay} \geq 0.5$. For $\tau = 0$ Random performed better than OppLite. We highlight Random selects on average the double of nodes as relay.

In ROLLERNET scenario, 2390 messages were created. Figure 6.6a shows the influence of delay threshold in message offloading. We observed a constant behavior for zero delay tolerance; this is explained by the fact that nodes have shorter contact times in such scenario. However, when the application can tolerate 600s or 1200s delay, for $\Gamma_{relay} \geq 0.5$ OppLite offloads up to 50% and 65%, respectively. Here, the *Oracle* solution has poor performance to offload messages because of the high number of isolated nodes.

Figure 6.6b shows the results for Random and OppLite with delays of 100s and 1200s in ROLLERNET scenario. We omitted from the figure results with zero delay, since OppLite and Random achieve similar results, offloading only 2% of the messages on average. In this scenario, if all nodes act as relays, OppLite can offload 74% of the traffic when the application tolerates up to 20 minutes of delay. OppLite ties the performance with Random when $\Gamma_{relay} \geq 0.5$, However OppLite selects 42 nodes as relay and Random selects 580 nodes.

We highlight the dominating set represents the best candidates to offload messages. However, Rollernet dataset has a high number of nodes isolated in the most of the time. Furthermore, there is a tradeoff between the number of nodes selected as relays and the amount of messages offloaded: a high number of relays means high data offload, but this requires the users' willingness to switch their device to relay mode.

We note that for both scenarios the acceptable delivery delay of an application might limit the amount of messages offloaded. For instance, participatory sensing

applications support 20 minutes of delay before uploading data. On the other hand, users may require lower delay for sending an email.

OpR application further represents nodes without connection with the infrastructure being favored by relay nodes. The results presented above showed OppLite may help to extend network coverage for users with delay tolerance acceptance.

6.4.3 Cache-and-Forward (CaF) Evaluation

In CaF, data offloading occurs every time a relay has opportunity to replicate a message to nodes in opportunistic mode. Offloading ratio quantifies the amount of messages forwarded opportunistically. Again, we show the feasibility of our framework for different delay tolerances. In this scenario, OppLite aims to maximize message forwarding by relay nodes with minimum delay.

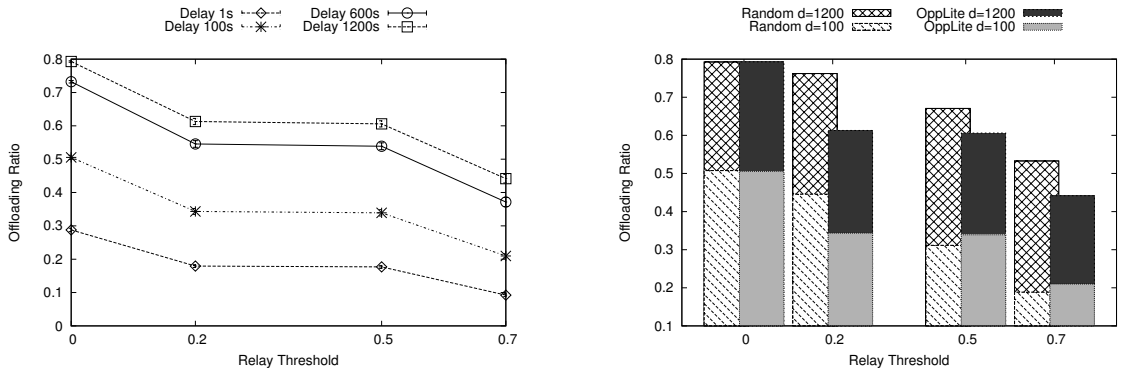
6.4.3.1 Data Offloading

In INFOCOM scenario, nodes requested 865 messages directly to infrastructure. Figure 6.7 shows OppLite offloading ratio with different values of delay tolerance (τ) considering $\Gamma_{opp} = 0.01$. In CaF, τ represents the time-to-live of each message received, thus, how long a relay node forward a given message to opportunistic nodes. Therefore, high values of τ cause higher offloading rate. Indeed, as shown in Fig. 6.7a when all nodes are in relay mode ($\Gamma_{relay} = 0$), message replications represent from 28% up to 79% of the network data traffic with delay tolerance of 1 and 1200 seconds, respectively. The offloading ratio drops to 9% and 44% when $\Gamma_{relay} = 0.7$ for $\tau = 1s$ and $\tau = 1200s$, respectively.

We observed that increasing τ from 1 to 1200 seconds improved offloading ratio in 182% for all Γ_{relay} values on average. Meanwhile, this improvement is only 12% on average when increasing τ from 600 to 1200 seconds.

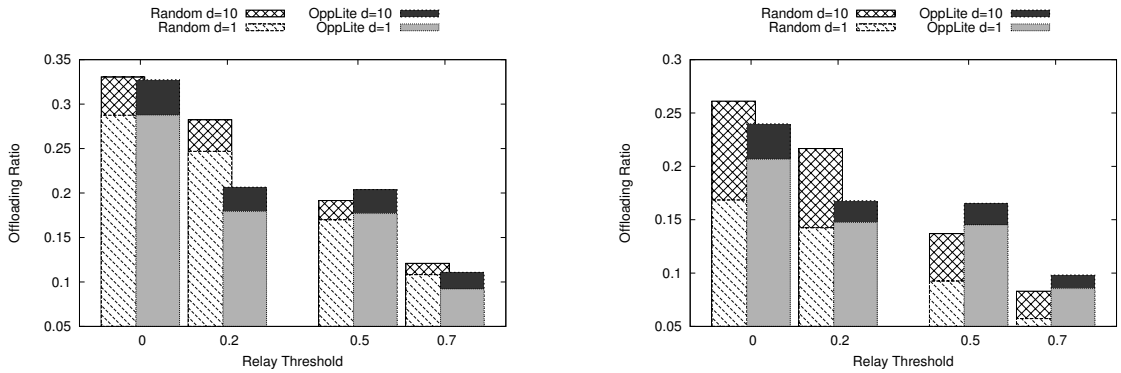
Figure 6.7b shows a comparison between Random and OppLite with 100 and 1200 seconds of delay tolerance (similarly as shown for OpR scenario). OppLite beats Random's offloading ratio when $\Gamma_{relay} \geq 0.5$ with small values of delay tolerance.

In order to support this observation, Figure 6.8 shows the offloading ratio comparison between OppLite and Random with 1 and 10s of delay tolerance and $\Gamma_{opp} \in \{0.01, 0.7\}$. In Fig. 6.8a $\Gamma_{opp} = 0.01$ and 96 nodes are in opportunistic mode. In this case, OppLite overcomes Random algorithm when $\tau = 1$ and $\Gamma_{relay} = 0.5$. On the other hand, $\Gamma_{opp} = 0.7$ represents a conservative opportunistic threshold value, where only 30 and 16 nodes are in opportunistic modes in Random and OppLite solution, respectively. As shown in Fig. 6.8b, OppLite can offload 20% of traffic for $\Gamma_{relay} = 0$



(a) Delay tolerance effect (b) Random vs OppLite

Figure 6.7. Cache-and-Forward data offloading: Relay Threshold impact in INFOCOM dataset



(a) $\Gamma_{opp} = 0.01$. (b) $\Gamma_{opp} = 0.7$.

Figure 6.8. Cache-and-Forward data offloading: Random vs OppLite in INFOCOM dataset.

and 8% for $\Gamma_{relay} = 0.7$ against 16% and 5% in Random for the same thresholds. We emphasize Random solution tends to select more nodes as relays, 48 versus 18 when $\Gamma_{relay} = 0.5$, and 29 versus 14 when $\Gamma_{relay} = 0.7$.

Each relay node replicated four distinct messages in INFOCOM dataset on average. The average number of messages forwarded by nodes in relay mode can be seen in Figure 6.9. High τ values allow nodes replicate messages for longer than lower values. The number of messages forwarded for all Γ_{relay} is almost constant when the delay tolerance is low. The amount of replications of a message in lower delays is defined by the size of connected components and nodes in opportunistic mode belonging to this component.

Figure 6.10 shows the offloading ratio for several τ values and a comparison among

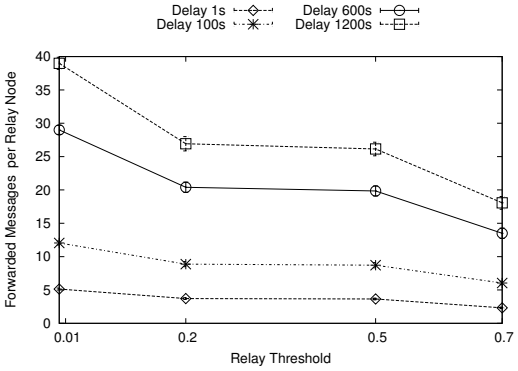
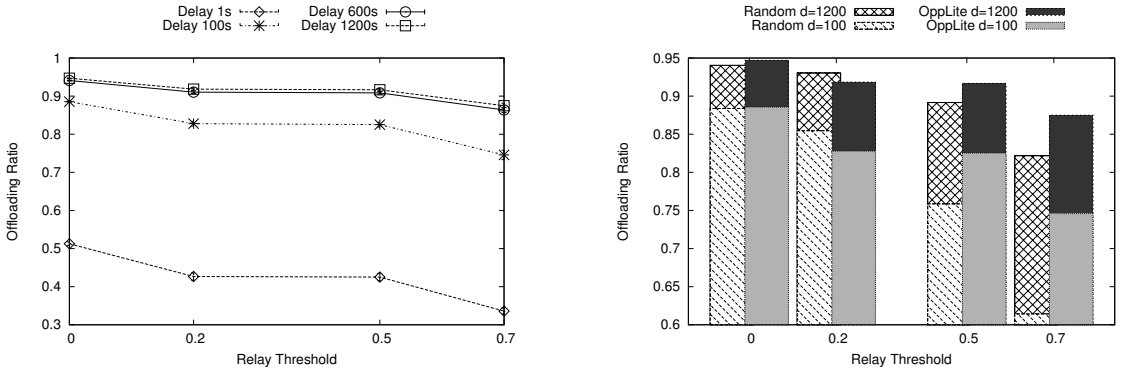


Figure 6.9. Cache-and-Forward INFOCOM: Average messages forwarded by relay nodes.



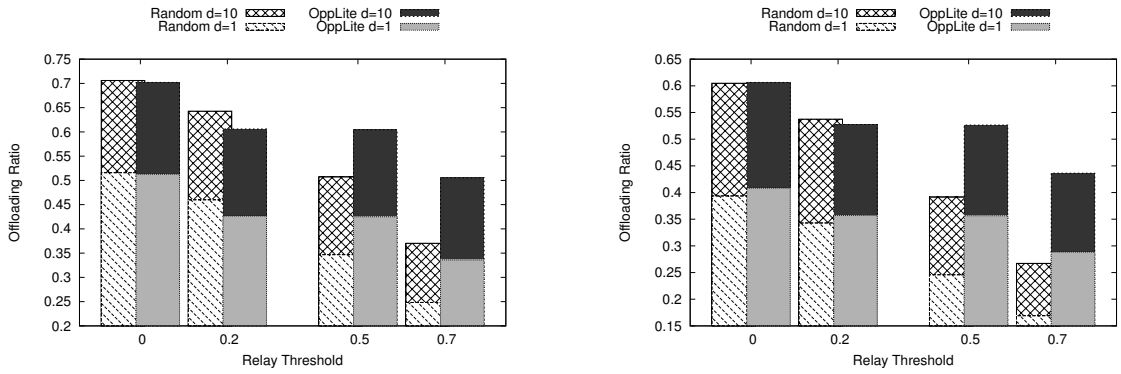
(a) Delay tolerance effect (b) Random vs OppLite

Figure 6.10. Cache-and-Forward data offloading: Relay Threshold impact in ROLLERNET dataset

OppLite and Random. As shown in Fig. 6.10a $\Gamma_{relay} = 0$ causes OppLite to offload (51, 88, 94, 95)% for $\tau = \{1, 100, 600, 1200\}$, respectively. We highlight the difference in the offloading ratios between $\tau = 600$ and $\tau = 1200$ are less or equal 1% $\forall\Gamma_{relay}$. Therefore, OppLite reaches its near-maximum data offloading within 10 minutes of delay tolerance.

We observed OppLite outruns Random for $\Gamma_{relay} \geq 0.5$, Fig. 6.10b. In $\Gamma_{relay} = 0.5$ and $\Gamma_{relay} = 0.7$ Random selects 556 and 333 nodes as relays, while OppLite selects 54 and 34 nodes, respectively. These results show choosing nodes based on the set of criteria used by OppLite has better performance than choosing a larger set of nodes.

Comparisons with different opportunistic thresholds between Random and OppLite with lower delay tolerance values are shown in Figure 6.11. In both figures, 6.11a and 6.11b, OppLite outperforms Random when $\Gamma_{opp} \geq 0.5$. In Fig. 6.11b $\Gamma_{opp} = 0.7$,



(a) $\Gamma_{opp} = 0.01$. (b) $\Gamma_{opp} = 0.7$.

Figure 6.11. Cache-and-Forward data offloading: Random vs OppLite in ROLLERNET dataset.

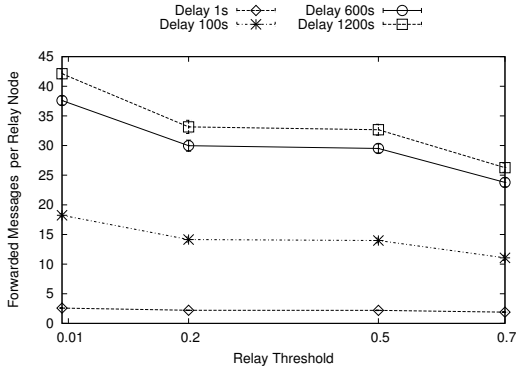


Figure 6.12. Cache-and-Forward ROLLERNET: Average messages forwarded by each relay node.

OppLite presented 28% offloading ratio against 16% in Random, even with only 32 nodes in opportunistic mode in OppLite against 330 opportunistic nodes in Random.

Due to ROLLERNET trace characteristics OppLite performed better in lower delay tolerance scenarios. In the ROLLERNET scenario, Random switches nodes to opportunistic mode which rarely will find a suitable node in relay mode.

OppLite forwards up to 43 messages by each relay in ROLLERNET scenario, Fig. 6.12. On average, each node forwards 2 distinct messages. However, ROLLERNET provides more contact opportunities and as a consequence, relay nodes can forward their messages more often than in the INFOCOM scenario.

6.4.4 Relay-as-Cache (RaC) Evaluation

RaC represents an HTTP request application, where all nodes request content to the cloud. However, the amount of content available to request, referred as pool size (S), was limited to $S = [100, 1000, 10000]$ distinct contents. Each content has an key identification and nodes choose a key to request following a Zipf distribution, thus, some keys are requested more frequently.

In *RaC*, delay tolerance (τ) means how long an opportunistic node waits until finding the desired content in some relay node's cache. After this delay tolerance, opportunistic nodes request the content using the infrastructure. Furthermore, opportunistic nodes request only one content until getting a response, opportunistically or directly from the infrastructure.

The total of content request is affected by this restriction as shown in Table 6.2. This table shows the average of total requests based on delay tolerance τ for INFOCOM and ROLLERNET dataset for all pool sizes when $\Gamma_{relay} = 0$ and $\Gamma_{opp} = 0.01$. We emphasize the difference in total requests between different content pool sizes and other threshold values is negligible.

Table 6.2. Average content requests for each delay tolerance.

Delay (τ)	INFOCOM	ROLLERNET
1	16890.7	60528.42
10	11797.30	53757.30
100	3822.82	10791.39
600	1037.09	1358.03
1200	660.06	367.94

6.4.4.1 Caching Results

We used *Hit Rate* as metric to evaluate caching OppLite performance. Hit rate measures the probability an opportunistic node finds a content among nodes in relay mode in its vicinity.

Figure 6.13 shows the hit ratio for all pool sizes (S) in INFOCOM and ROLLERNET dataset. We observed that Hit Ratio performed better in graphs with more stable connectivity. In INFOCOM trace, most of the nodes have degree greater than 1 during entire network lifetime. Meanwhile, Rollernet represents a sparse network and several nodes are isolated in the most of the time.

In the best case, as shown in Figures 6.13a and 6.13b, when the pool size is small ($S = 100$) and all nodes are in relay mode ($\Gamma_{relay} = 0$), the probabilities of opportunistic nodes find a content within relay's cache are 48% and 22% in INFOCOM and ROLLERNET trace, respectively. Even with $\Gamma_{relay} = 0.7$, OppLite reaches up to 38% of hit ratio in INFOCOM trace (only 14 relays) and 18% in ROLLERNET (with 33 relays).

In Figure 6.13a, we observed OppLite with 1 second of delay tolerance outperformed the hit ratio obtained with 100 seconds of delay tolerance. In fact, in the scenarios where $\tau = 1$ and $S = 100$, a small group of contents are requested more times (due to Zipf distribution). Since the pool size is small, in the beginning of the simulated time opportunistic nodes with low delay tolerance request messages directly to infrastructure quickly, which increase the probability of nodes been carrying the requested content in the future.

As shown in Table 6.3, $\tau = 1$ causes 99% of cacheability. In other words, there are 99% of chance a node finds a content in other nodes cache in a perfect caching mechanism.

For pool size of 1000 and 10000, the hit ratio increase accordingly to the delay tolerance, as shown in Figures 6.13c and 6.13e. To increase the pool from 1000 to 10000 contents available reduced 5% of the hit ratio approximately when $\Gamma_{relay} \geq 0$ for all delay tolerance values evaluated in the INFOCOM trace.

In ROLLERNET trace, the higher hit ratio was obtained with 600s of delay tolerance (22%), shown in Figure 6.13b. To understand this result, we remind the simulation parameters discussed in Section 6.2: Rollernet trace has 12500 seconds of duration, nodes begin to request content after a warm up period of 1800 seconds and stop requesting at 9500 sec. Since nodes need to wait until receiving a response before requesting novel content, nodes sent only 368 messages request on average with $\tau = 1200$, as shown in Table 6.2. Therefore, the probability of a content be stored in the nodes cache decays with $\tau = 1200$, due the fact the inter content requesting time is too long when compared with the network lifetime.

Figures 6.13d and 6.13f showed a similar hit ratio for $S = 1000$ and $S = 10000$ in ROLLERNET scenario, respectively. OppLite obtained between 2 and 18% hit ratio, which are considered low when compared with the cacheability of ROLLERNET in Table 6.3. This fact is mainly caused by the high number of isolated nodes (rollernet trace contains 1112 nodes), which can request content but are unable to become opportunistic because they have zero neighbors.

Overall, OppLite reached higher hit ratio when opportunistic nodes supported 10 minutes or more of delay to find a content. In restricted scenarios, with few nodes as

in INFOCOM, even with $\Gamma_{relay} = 0.7$, OppLite reached a hit ratio of 6%.

However, OppLite was designed as a communication mode selection algorithm and it has space to be improved as a caching mechanism. In the next section, we discuss these results in terms of data traffic savings.

Table 6.3. Cacheability of the network in all analyzed scenarios.

Delay (τ) / Size (S)	INFOCOM			ROLLERNET		
	100	1000	10000	100	1000	10000
1	0.99	0.95	0.82	0.99	0.98	0.93
10	0.99	0.93	0.80	0.99	0.98	0.93
100	0.97	0.84	0.77	0.98	0.94	0.90
600	0.91	0.79	0.75	0.94	0.89	0.87
1200	0.88	0.78	0.75	0.87	0.81	0.80

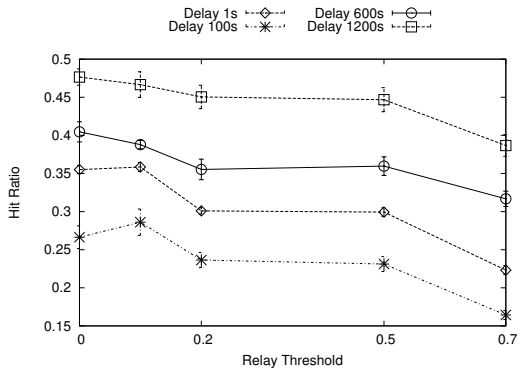
6.4.4.2 Data Offloading

We first discuss INFOCOM scenario results, shown in Figure 6.14. When there are only 100 contents available to download (Fig. 6.14a), OppLite offloads up to 42% of the traffic whether opportunistic nodes support 1200s of delay tolerance and $\Gamma_{relay} = 0$ or 34% whether $\Gamma_{relay} = 0.7$. In this scenario, we observed that OppLite offloads from 22% to 35% of the traffic when $\tau = 1s$. Figure 6.14b shows a comparison between Random and OppLite with delay tolerance of 1 and 1200 seconds. OppLite outperforms Random solution in 1 second of tolerance when $\Gamma_{relay} \geq 0.1$. In $\Gamma_{relay} \geq 0.5$ and $\tau = 1$s, OppLite offloads 52% more than Random with half of the nodes selected as relays, 14 and 28 relays, respectively.

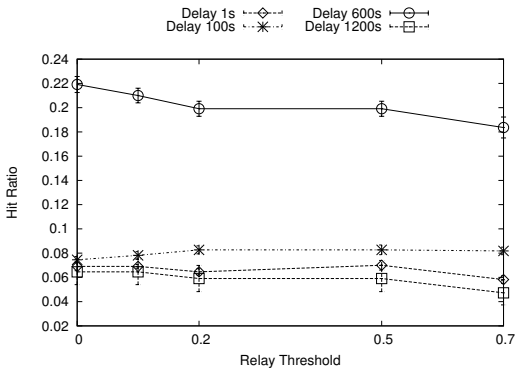
When the pool of contents increases the offloading ratio decays, Figures 6.14c-f. Random solution keeps almost constant its offloading ratio when $\tau = 1200$s due to the high number of nodes selected as relays. OppLite achieves higher offloading rate when opportunistic nodes have low delay tolerance.

Although Random showed better offloading ratio in high delay tolerance scenarios, the difference for OppLite is less than 5% for all scenarios. Taking Figure 6.14c as an example, with $\Gamma_{relay} = 0.7$ Random selected 28 nodes as relays and offloaded 31% of the traffic, while OppLite selected just 14 nodes as relays and offloaded 28% of the traffic.

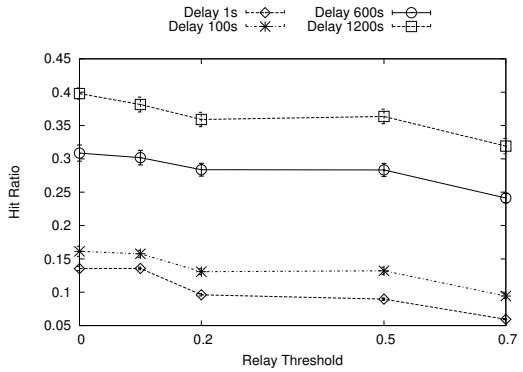
Figure 6.15 shows the offloading ratio for all pool of contents evaluated in Rollernet trace. Similar to the hit ratio, the higher offloading ratio was obtained within a



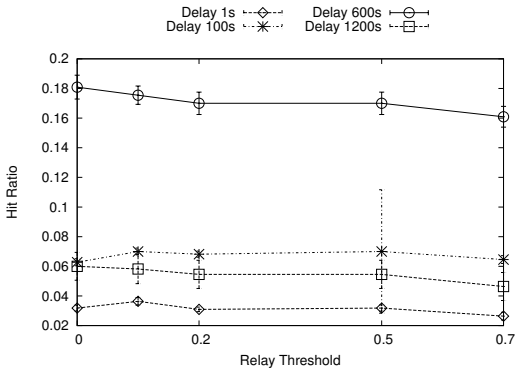
(a) Infocom. $S = 100$.



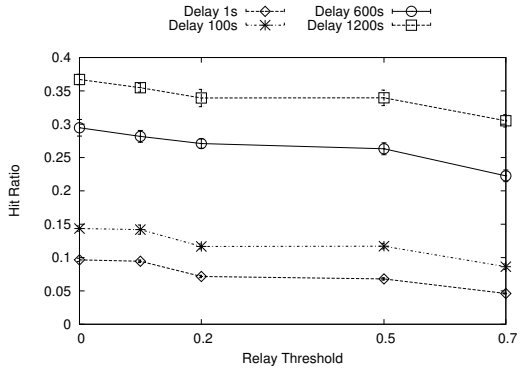
(b) Rollernet. $S = 100$.



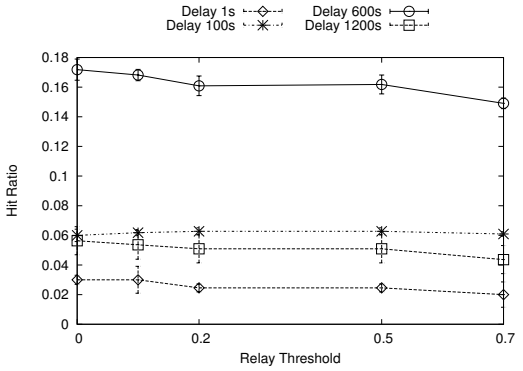
(c) Infocom. $S = 1000$.



(d) Rollernet. $S = 1000$.



(e) Infocom. $S = 10000$.

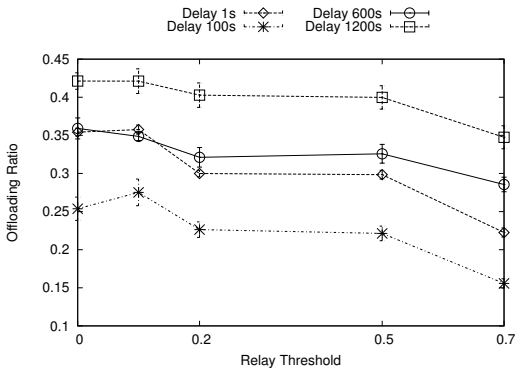


(f) Rollernet. $S = 10000$.

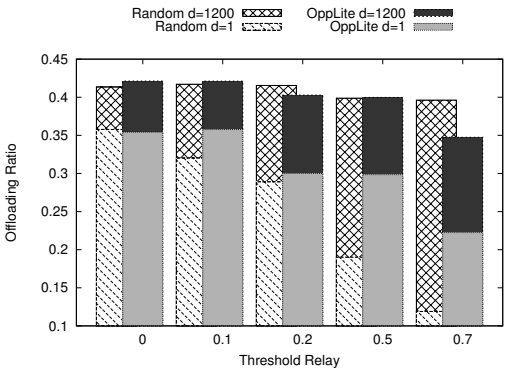
Figure 6.13. Relay-as-Cache: Hit Ratio performance. Left column shows results for INFOCOM trace while right column shows hit ratio for Rollernet. This figure aggregates all sizes of the content pool.

delay tolerance of 600 seconds in all scenarios. With $\tau = 600$, OppLite offloads up to 17% of network data traffic in the best case (Fig. 6.15a) and 12% in the worse scenario (Fig. 6.15e) .

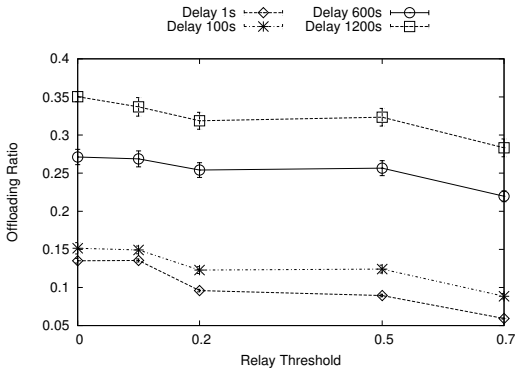
As can be seen in Figures 6.15b, 6.15d and 6.15f, OppLite offloaded 10% less



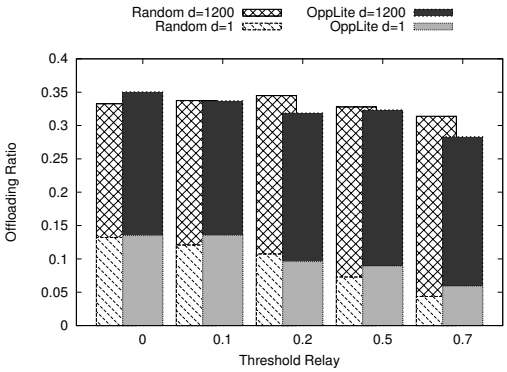
(a) Delay tolerance effect. $S = 100$.



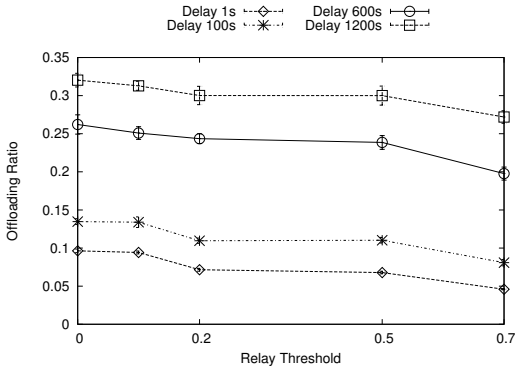
(b) Random vs OppLite. $S = 100$.



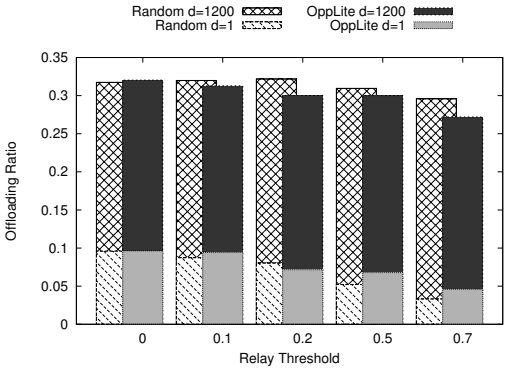
(c) Delay tolerance effect. $S = 1000$.



(d) Random vs OppLite. $S = 1000$.



(e) Delay tolerance effect. $S = 10000$.

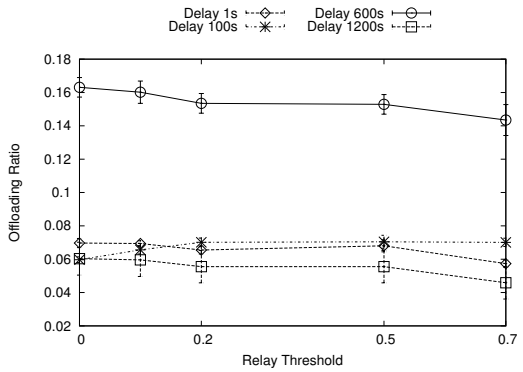


(f) Random vs OppLite. $S = 10000$.

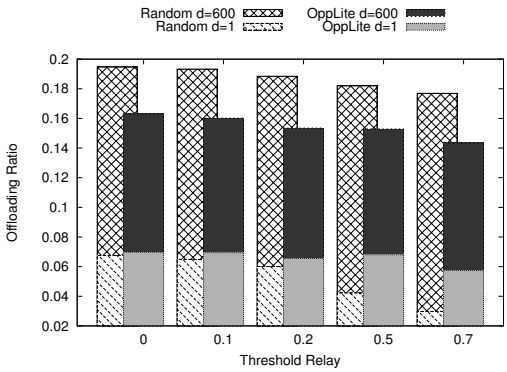
Figure 6.14. Relay-as-Cache INFOCOM dataset: Data offloading.

than the Random solution for all values of S In Rollernet with $\tau = 600$ and equal or greater than Random with $\tau = 1$. As mentioned before, Random solution selects much more nodes than OppLite in Rollernet trace, with $\Gamma_{relay} = 0.7$ there are 330 relay nodes in Random against 34 relays in OppLite.

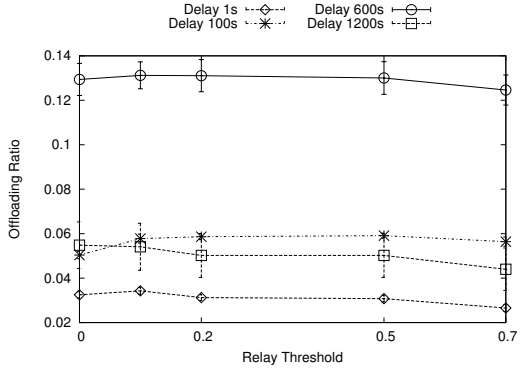
Although ROLLERNET trace contains 1112 nodes, only 62 nodes are active during the network lifetime. Furthermore, connections and disconnections are more



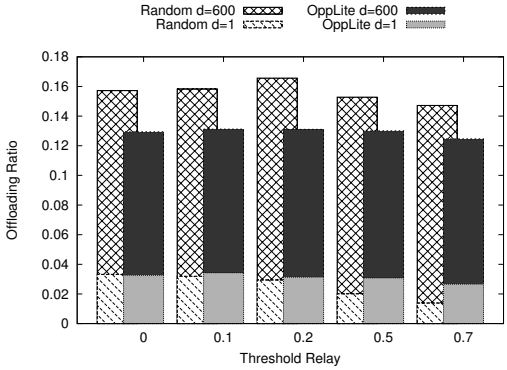
(a) Delay tolerance effect. $S = 100$.



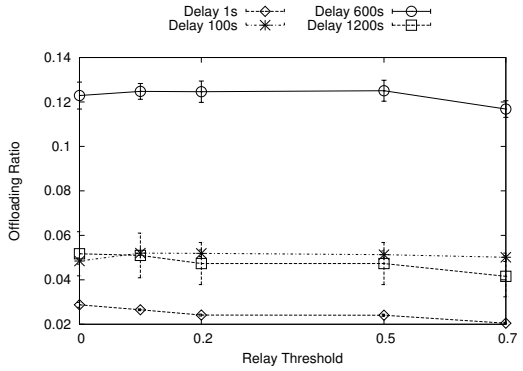
(b) Random vs OppLite. $S = 100$.



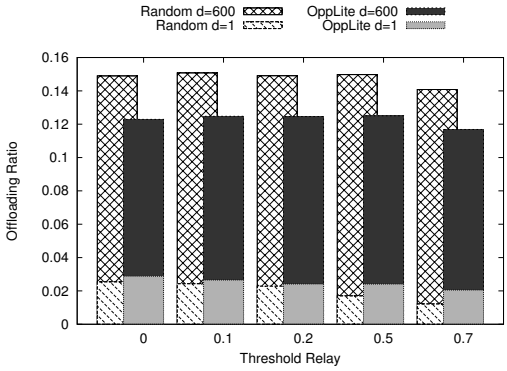
(c) Delay tolerance effect. $S = 1000$.



(d) Random vs OppLite. $S = 1000$.



(e) Delay tolerance effect.$S = 10000$.



(f) Random vs OppLite. $S = 10000$.

Figure 6.15. Relay-as-Cache ROLLERNET dataset: Data offloading.

frequent due the *accordion phenomenon* observed in this trace [Tournoux et al., 2009]. However, in *RaC* all nodes can request contents, which decreased the offloading ratio when compared with INFOCOM trace.

Overall, the offloading ratio followed the hit ratio results. Indeed, the duration of pair-wise node connection is enough for a relay to transmit a content requested by an opportunistic node. Furthermore, a content is entirely forwarded, without fragmen-

tation.

OppLite allows users to set their thresholds to switch to relay and opportunistic mode and their delay tolerance through the Γ_{relay}, Γ_{opp} and τ parameters. However, WISPs should encourage users to become opportunistic nodes and tolerate some delay and further, encourage users to become relay and offer their resources (such as cache and energy) to others.

This evaluation showed the potential of the using personal devices as cache to offload data traffic from cellular to opportunistic network within a low delay.

6.5 Conclusions

In this chapter, we presented and discussed an extensive evaluation of OppLite. We proposed three applications to evaluate OppLite: *i*) *Opportunistic Relaying* (OpR), opportunistic nodes forward their messages to infrastructure through relay nodes; *ii*) *Cache-and-Forward* (CaF), relay nodes forward all content they receive to opportunistic nodes; and *iii*) *Relay as Cache* (RaF), opportunistic nodes seek for content in relay nodes cache.

To refine our simulations we used a Weibull distribution to model users inter requesting time, a Zipf distribution to choose which content to request (only for *CaF* application) and two real traces of contact: Infocom, which represents a conference scenario; and Rollernet, which represents people rollerblading around the city of Paris. We characterized these traces and show they have different connectivity characteristics. INFOCOM presents a stable graph with few connected components with many nodes, while ROLLERNET has dynamic graphs with moments of more connected components with fewer nodes.

OppLite reduced the number of nodes using the infrastructure in *OpR* application. We observed relay nodes can forward up to 45% of the traffic in a conference scenario within a delay of less than 20 minutes. In a scenario representing a moving crowd in the streets of Paris, relay nodes can forward up to 70% of the traffic.

In *CaF* application, OppLite offloads between 30 and 80% of the data traffic in INFOCOM and between 52 and 94% in ROLLERNET within a delay of 20 minutes.

In *RaC* application, OppLite offloads from 32 to 44% of the data traffic when opportunistic nodes tolerate 20 minutes of delay in INFOCOM dataset. In case opportunistic nodes wait for only 1 second until receiving a response from a relay, OppLite offloads from 5 to 35% of the data traffic. Meanwhile, in ROLLERNET trace, the offloading ratio ranged from 7 to 17% with 10 minutes of delay.

In all evaluated applications, OppLite expects user cooperation to switch its device to relay or opportunistic mode. Although OppLite allows users to configure their willingness to cooperate, it is uncertain whether users would be cooperative (setting low values of Γ) or not (setting high values of Γ). Incentive mechanisms may deal with these issues, which are discussed in the next chapter.

Chapter 7

Towards Incentive Mechanisms for Opportunistic Mobile Data Offloading

Opportunistic offloading relies on cooperative opportunistic communication between devices. In this chapter, we discuss how to encourage users to participate actively during network lifetime. However, since cooperative relaying or opportunistic cellular offloading scenario spends devices' resources, a natural question arises: *How to engage user willingness to relay data to others*?

The user willingness to cooperate depends on the benefits received. We proposed MINEIRO, an incentive mechanism that engages users to forward third-party messages. To achieve user cooperation for opportunistic mobile data offloading, we propose mechanisms based on data reward paid by WISPs and further, an integration between OppLite and MINEIRO to provide user-centric incentive mechanisms.

The rest of this chapter is organized as follows: Section 7.1 proposes a reciprocity based incentive mechanism, called MINEIRO, to engage users to forward messages to other users in general opportunistic networking. Section 7.2 discusses how to encourage user cooperation based on benefits and costs for users act as relay or opportunistic. Furthermore, it proposes two approaches as incentive mechanisms: i) data reward-based mechanism, where WISPs can manage the award value for cooperative users and ii) OppLite integrated with MINEIRO, which ranks cooperative and non-cooperative users. Finally, Section 7.3 concludes this chapter.

7.1 Incentive Mechanism for Opportunistic Forwarding

In a generic manner, opportunistic networking researchers have already proposed dozens of forwarding algorithms, which differ in the way they decide when and to

whom forwarding the messages [Mota et al., 2014]. However, the majority of these algorithms assumes all nodes cooperate with the network by forwarding messages correctly. This assumption fits in networks where all nodes cooperate to achieve a certain task, for instance, in an environmental monitoring sensor network or in a rescue mission after a disaster. However, in applications where individuals perform unrelated tasks, the human behavior should be considered.

We propose an incentive mechanism, called MINEIRO[1], a ***M****essage-based* ***IN****centive mechanism for* ***E****nd-user* ***I****mprovement of* ***R****outing* ***O****pportunities* in opportunistic networking. MINEIRO builds a reputation rank based on the source of messages received by the forwarding nodes. Our motivation is to provide a framework aiming to increase user willingness to forward messages to others.

Thus, MINEIRO provides a **technical** benefit for users, since forwarding messages for others provides performance improvement for itself. To show that be cooperative is the best behavior for users, we model MINEIRO as a Bayesian game and show the conditions where MINEIRO achieves the Bayesian equilibrium. Furthermore, we analyze the impact of selfish behavior in the well-known opportunistic Epidemic forwarding algorithm in two mobility scenarios: RandomWay point mobility model and a trace-based mobility model.

7.1.1 MINEIRO - Reciprocity based Incentive Mechanism

MINEIRO classifies the forwarding nodes according to the source of the messages they transmitted, thus punishing those nodes that forward only their own messages. As a consequence, if nodes wish to increase their chances of having their messages delivered to the destination, then they should forward messages from other nodes.

A node increases its reputation by relaying third-party messages, while decreases it by forwarding its own messages. For instance, in Figure 7.1, after node B relays the message x to C, C increases the reputation of B.

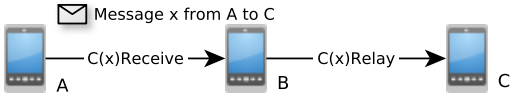


Figure 7.1. An example of transmission in an ad hoc network. Node B considers the cost to receive and to relay the message x.

[1]The acronym MINEIRO is also a tribute for the people born in the state of Minas Gerais, Brazil, known by their hospitality and trust in strangers.

Algorithm 1 describes our incentive mechanism. MINEIRO follows a proof-of-trust rule: Initially, when a pair of nodes, u and v, meet for the first time, both of them mark the timestamp of the event and assume a non-zero reputation of each other (Lines 6-8). In further meetings between u and v, the node u will analyze the messages that v wants to forward. Node u increases the reputation of v for each message where the source is different from v. Otherwise, node u decreases the reputation of node v, that is, v is forwarding its own messages (Lines 10-18).

A node recognizes all nodes in the ranking with reputation equal to zero as selfish, and messages coming from them are rejected, with the exception of the node itself being the destination. The node forwards correctly all messages coming from nodes with reputation greater than zero. Periodically, a node evaluates its reputation table to decrease the reputation of nodes that were not contacted for a long time.

Algorithm 1 Node Reputation Algorithm

```
1:  input: Node u receives the messages M forwarded by v
2:  T_v = now(); {Updating the timestamp}
3:  if R_v ∉ R then
4:      R_v = r_init;
5:  end if
6:  for all M_j ∈ M do
7:      if destination(M_j) ≠ u ∧ source(M_j) ≠ v then
8:          R_v = min(R_v + Δ_increase, 1);
9:      else
10:         if destination(M_j) ≠ u ∧ source(M_j) = v then
11:             R_v = max(R_v − Δ_decrease, 0);
12:         end if
13:     end if
14: end for
```

The reputation value indicates the behavior of a node v:

- **Altruist -** $\mathcal{R}_v = 1$: The node accepts and forwards all messages from all nodes.
- **Rational -** $\mathcal{R}_v =]0, 1[$: The node rejects the messages from selfish nodes.
- **Selfish -** $\mathcal{R}_v = 0$: The node rejects all messages from all nodes.

Since a node always receives an initial reputation, it has the opportunity to forward its own messages in a first meeting. A high initial reputation value leads the algorithm to take a long time to indicate other nodes as selfish, while low values may indicate nodes as selfish quickly (false positives). If a node's reputation value is zero, it will have its messages rejected, consequently, decreasing its message delivery probability. Meanwhile, rational nodes tend to balance their reputation value, thus increasing the network performance.

Assuming that users are rational, if an uncooperative behavior causes degradation on the network quality for a user, then this user tends to collaborate with the network to improve its network quality. Further, our game-theoretical model in the next section shows that forwarding third-party messages is the best strategy for a node to increase its reputation value, and consequently, the probability of having its messages delivered.

MINEIRO differs from other proposals because a node does not require having previous knowledge about the others. In open networks, where nodes can get in and get out anytime, it is a challenge for a node to gather the information about other nodes, such as their public keys. Moreover, the algorithm avoids central entities to guarantee the reputation of other nodes, since the time to reach a third node in an opportunistic network can be very high. This makes the reputation calculation much faster, since it relies only on local information. However, the reputation takes longer to converge, since the node cannot take hints from other nodes.

This convergence time impacts on the decisions made by each node. In a first contact, a node always trust in other node and take more contacts and message exchanges to both nodes realize the reputation of each other.

7.1.2 MINEIRO as a Bayesian Game

Game theory aims to address situations in which the outcome of a participant of the game is affected by his or her decision, and the decisions made by all other participants they are interacting with [Easley and Kleinberg, 2010]. The situation where players have private information and they are uncertain about the preferences and intentions of others is referred as a *Bayesian game*. The *Bayesian equilibrium* is the strategy profile that maximizes the payoffs of all players given a type and a strategy chosen by one player and and his/her beliefs about the type and strategy chosen by the other players.

We model MINEIRO as a Bayesian game to analyze which parameter values lead the algorithm to a Bayesian equilibrium. Thus, consider the Bayesian game $\mathcal{G}_{MINEIRO} = \langle \mathcal{N}, \mathcal{T}, \mathcal{S}, \mathcal{Q}, \mathcal{U} \rangle$, where:

$\mathcal{N} = \{n_1, n_2\}$ is the set of players.

$\mathcal{T} = \{Se : Selfish, Co : Cooperative\}$ is the set of types.

$\mathcal{S} = \{Own, Third\}$, Forwarding its own messages or third-party message strategies, respectively.

$\mathcal{Q}, where \mathcal{Q}_i \in [0, 1]$ is the set of distribution probabilities according the type of the player. Let $\mathcal{Q}(\mathcal{T}_1 = Se) = \mathcal{R}$, thus $\mathcal{Q}(\mathcal{T}_1 = Co) = 1 - \mathcal{R}$.

$\mathcal{U}$, the set of payoffs.

The game considers the meeting of players n_1 and n_2, each player having in its buffer a set of own and third-party messages to be forwarded and none of the players are the destinations. Furthermore, each player has a type (behavior), selfish or cooperative, and can choose the strategy of forwarding its own messages only or also the messages of third-parties.

Table 7.1 represents the payoff matrices of the game. The matrix $\mathcal{T}_1 = Selfish$ assumes that player 1 behaves as selfish with probability $\mathcal{R}_1$, while in the matrix $\mathcal{T}_1 = Cooperative$, player 1 behaves as cooperative with probability $1 - \mathcal{R}_1$. The lines and columns represent the strategies chosen by players 1 and 2, respectively. Each tuple (u_1, u_2) represents the payoff received by player n_k in relation to his/her strategy and the strategy chosen by the other player. Let I be the initial payoff that each node receives when they meet each other, the r_{init} MINEIRO's parameter. Furthermore, Δ_{inc} and Δ_{dec} are the MINEIRO's parameters to increase and decrease the reputation, respectively.

From these payoff matrices, we analyzed which strategies maximize the payoff for both players and we derived Theorem 7.1.1.

Type $\mathcal{T}_1 = Selfish \therefore \mathcal{Q}(\mathcal{T}_1) = \mathcal{R}_1$

Strategy $\mathcal{S}$	Own	Third
Own	$I - \Delta_{dec}, I - \Delta_{dec}$	$I - \Delta_{dec}, I + \Delta_{inc}$
Third	$0, I - \Delta_{dec}$	$0, I + \Delta_{inc}$

Type $\mathcal{T}_1 = Cooperative \therefore \mathcal{Q}(\mathcal{T}_1) = 1 - \mathcal{R}_1$

Strategy $\mathcal{S}$	Own	Third
Own	$I - \Delta_{dec}, I - \Delta_{dec}$	$I - \Delta_{dec}, I + \Delta_{inc}$
Third	$I + \Delta_{inc}, I - \Delta_{dec}$	$I + \Delta_{inc}, I + \Delta_{inc}$

Table 7.1. Payoff's Matrices.

Theorem 7.1.1. *User cooperation is motivated when* $\Delta_{inc} \geq \Delta_{dec}$ *and* $\mathcal{R}_k < \frac{\Delta_{inc}+\Delta_{dec}}{I+\Delta_{inc}}$, *in this case, the strategy profile (Cooperative:Forwarding third-party messages, Forwarding third-party messages) is a Bayesian equilibrium for the game* $\mathcal{G}_{MINEIRO}$.

To demonstrate Theorem 1, we need to evaluate the perspective of each player. From the player n_2 perspective, the strategy *Forwarding third-party messages* is strictly dominant if it satisfies the equation $I + \Delta_{inc} \geq I - \Delta_{dec}$, $\forall(\Delta_{inc} \geq \Delta_{dec})$.

Player n_1 has different payoffs for each type and strategy chosen. In this case, $\sigma_1(Own)$ and $\sigma_1(Th)$ are the mixed-strategies profiles that represent the probabilities of the player n_1 to choose the strategy *Forwarding its own messages* or *Forwarding third-party messages*, respectively. Thus:

$$\begin{aligned}\sigma_1(Own) &= \sigma_1(Own|Se) + \sigma_1(Own|Co) \\ &= \mathcal{R}_1(I - \Delta_{dec}) + (1 - \mathcal{R}_1)(I - \Delta_{dec}) \\ &= (I - \Delta_{dec})\end{aligned}$$

$$\sigma_1(Th) = \sigma_1(Th|Se) + \sigma_1(Th|Co) = (1 - \mathcal{R}_1)(I + \Delta_{inc})$$

Since we want to know the conditions that the strategy *Forwarding third-party messages* is better for player 1, then we need to find the parameter values of the algorithm that satisfy the equation $\sigma_1(Te) > \sigma_1(Pr)$. This equation is satisfied when

$$0 < \mathcal{R}_1 < \frac{\Delta_{inc}+\Delta_{dec}}{I+\Delta_{inc}}$$

In these conditions, n_1 and n_2 earn their maximum payoffs by being cooperative and forwarding third-party messages. Thus, the strategy profile (*Cooperative:forward third-party messages, forward third-party messages*) leads to the Bayesian equilibrium.

7.1.3 MINEIRO Evaluation

The objective of this evaluation is to understand how Mineiro improves relaying. In other words, how Mineiro engages users to forward messages from other users.

7.1.3.1 Simulation Setup

Mineiro's performance was evaluated using the Opportunistic Networking Evaluator (ONE) [Keränen et al., 2009]. The network has 50 nodes in a 1000m x 1000m area. The simulated time was 28800s. Nodes are equipped with 802.11 interfaces with a radio range of 50m and bandwidth of 10Mbps. A random source-destination pair was drawn each 10-50s up to 20000s. This limit guarantees that the last generated message has time to be delivered. For the sake of simplicity, all nodes have infinite buffers, $\Delta_{increase}$ and $\Delta_{decrease}$ are equal 0.1, and $r_{init} = 0.5$, which were derived from Theorem 7.1.1. These parameters affect the amount of third-party messages a node sends before other nodes consider it selfish.

Nodes use the Epidemic forwarding protocol, which forwards all messages in the buffer to all other nodes encountered[2] [Vahdat and Becker, 2000]. We compared MINEIRO against two other policies:

[2]In the majority of store-and-forwarding protocols, before forwarding the messages, two nodes exchange their summary of messages being buffered, then each node compares the summary with the

- **No Incentive**: Epidemic algorithm with no incentive. In this case, nodes with selfish behavior deny to forwarding messages from third-parties.

- **Barter-Based**: When a pair of nodes gets in contact, first they exchange the list with the summary of messages that each one carries. Next, each node removes from its list the messages already stored in its buffer. Each node determines a value for each message in the list based on their age and the nodes exchange the result list with the summary of messages each one wishes to download. Finally, each node forwards the messages requested by the other node while they are in each other's radio range [Buttyán et al., 2010].

Selfish nodes with no incentive deny to forward any third-party message. Meanwhile, nodes with an incentive mechanism are rational. We consider two mobility models:

i) Random-Waypoint model (RWP), which is commonly used in mobile simulations. However, it abstracts the real human mobility pattern. We set the nodes moving with walking speed ranging from 0.5m/s up to 1.5m/s.

ii) Small World in Motion (SWIM) [Mei and Stefa, 2009], which exploits the regularity of human mobility and generates synthetic traces similar to real mobility traces. We set the parameters of SWIM as the default parameters shown in [Mei and Stefa, 2009].

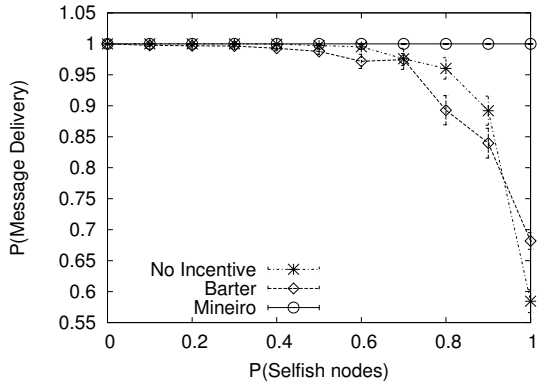
7.1.3.2 Impact of selfishness in Opportunistic Networks

We ranged the percentage of selfish nodes in the network from 0 to 100%. A selfish node, $\mathcal{R}_v = 0$, may take two actions after receiving a message: *i)* The node keeps the messages in its buffer without forwarding them, referred as *non-forwarding behavior*; or *ii)* The node drops the messages received, referred as *dropping message behavior*.

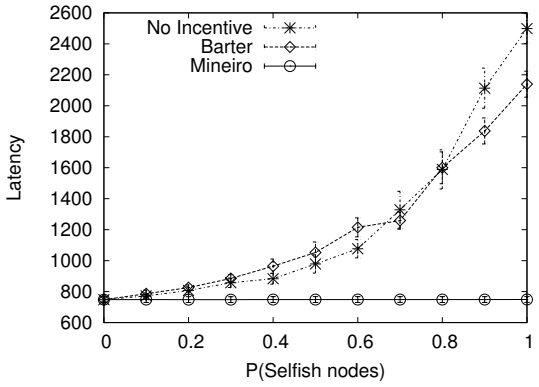
The difference between these actions is that a node that drops the message can receive the same message several times.

Figures 7.2 and 7.3 show the results for the RWP mobility model scenario. In this scenario, the delivery ratio keeps constant for up to 60% of selfish nodes, Figure 7.2a. This shows that in the RWP scenario, the network supports up to 60% of nodes with selfish behavior without degrading its delivery ratio. After this value, Epidemic with *No incentive* and the *barter-based* policy decrease rapidly. Moreover, Figure 7.2b shows that the delay increases linearly for these policies, while MINEIRO presents

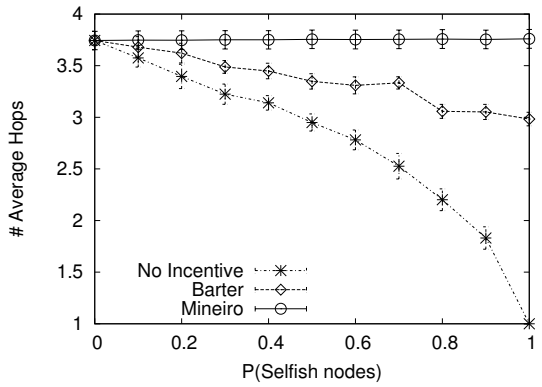
messages already in its buffer, and forwards only the set difference between the messages in its buffer and the messages in the summary, avoiding unnecessary retransmissions.



(a) Messages delivered



(b) Latency



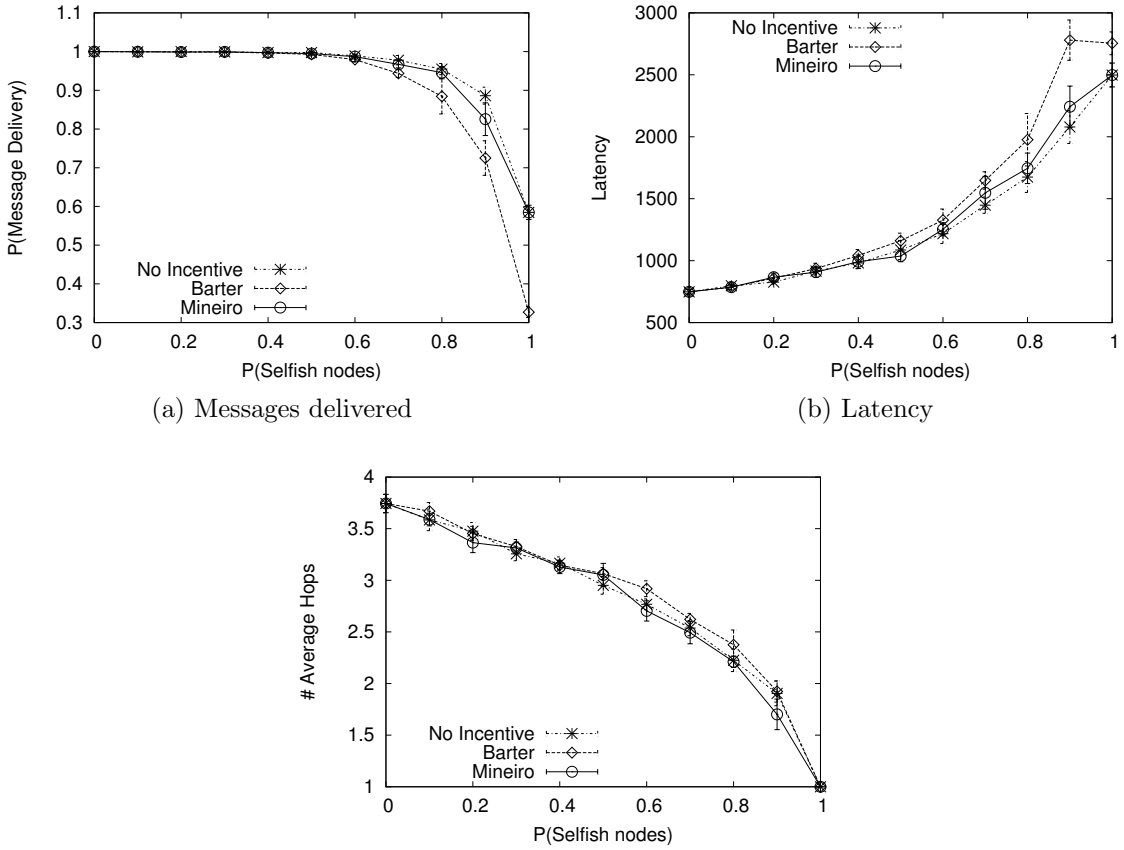
(c) Average hops per message

Figure 7.2. Non-forwarding selfish behavior in RandomWay mobility model.

constant values for these metrics. Figure 7.2c explains the reasons for this: while the selfishness percentage increases, the number of hops decreases for *No incentive* and *Barter-based policy*. In the worst case, when all nodes are selfish, the average number of hops is one, which means that all messages are directly delivered to the destination. MINEIRO keeps the number of hops constant, thus, it encourages rational users to relay messages.

The results of the SWIM model are presented in Figures 7.4 and 7.5. The delivery ratio is lower than in RWP due to the creation of social communities. Messages generated to nodes external to the community have a less probability to be delivered. Moreover, the delivery ratio decreases linearly for the *No incentive* and *Barter-based* approaches, while MINEIRO presents a linear result. Again, MINEIRO keeps the number of hops constant, stimulating nodes to forward third-party messages.

As shown in Figures 7.3 and 7.5, the selfish behavior of dropping messages presented the worst results for all policies. When a node drops a message, it becomes eligible to receive the same message from the same node in future encounters or from



(a) Messages delivered

(b) Latency

(c) Average hops per message

Figure 7.3. Dropping message selfish behavior in RandomWay mobility model.

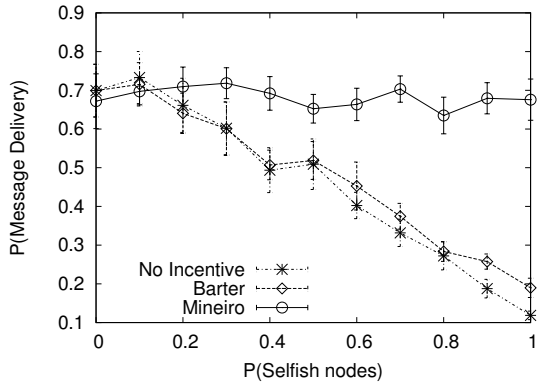
others. Thus, dropping messages increases the number of retransmissions of the same message.

7.2 Engaging Cooperation in Opportunistic Offloading

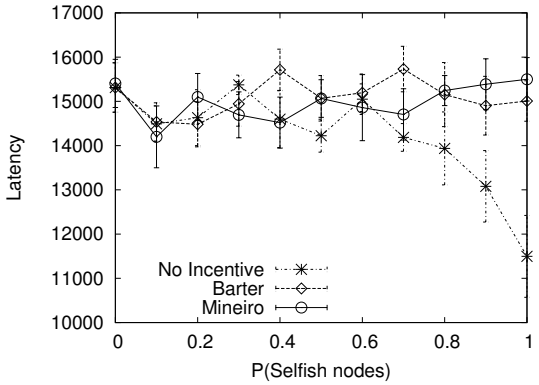
Two actors drive the effectiveness of cooperation in mobile offloading: the WISP and its customers, where three behaviors are expected for these customers, become relay device, become opportunistic device or be non-cooperative.

In this section, we present two incentive mechanisms based on each one of these actors. WISPs can encourage customers through rewards, thus, we propose a data reward model without cost explosion problem. Meanwhile, users may cooperate among themselves to improve their own performance. We proposed an integration between *OppLite* and *MINEIRO* in a centralized and decentralized fashion, with minimal or no intervention on the WISP infrastructure.

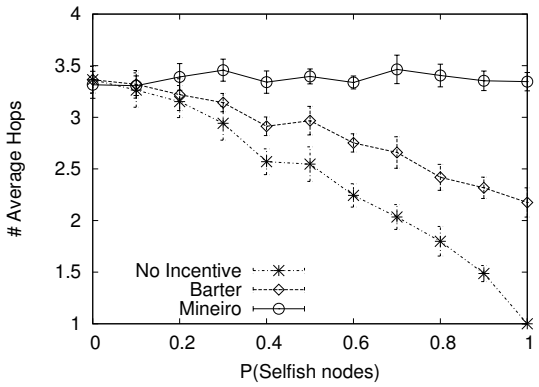
Figure 7.6 illustrates these actors and their distinct behaviors. The natural be-



(a) Messages delivered



(b) Latency



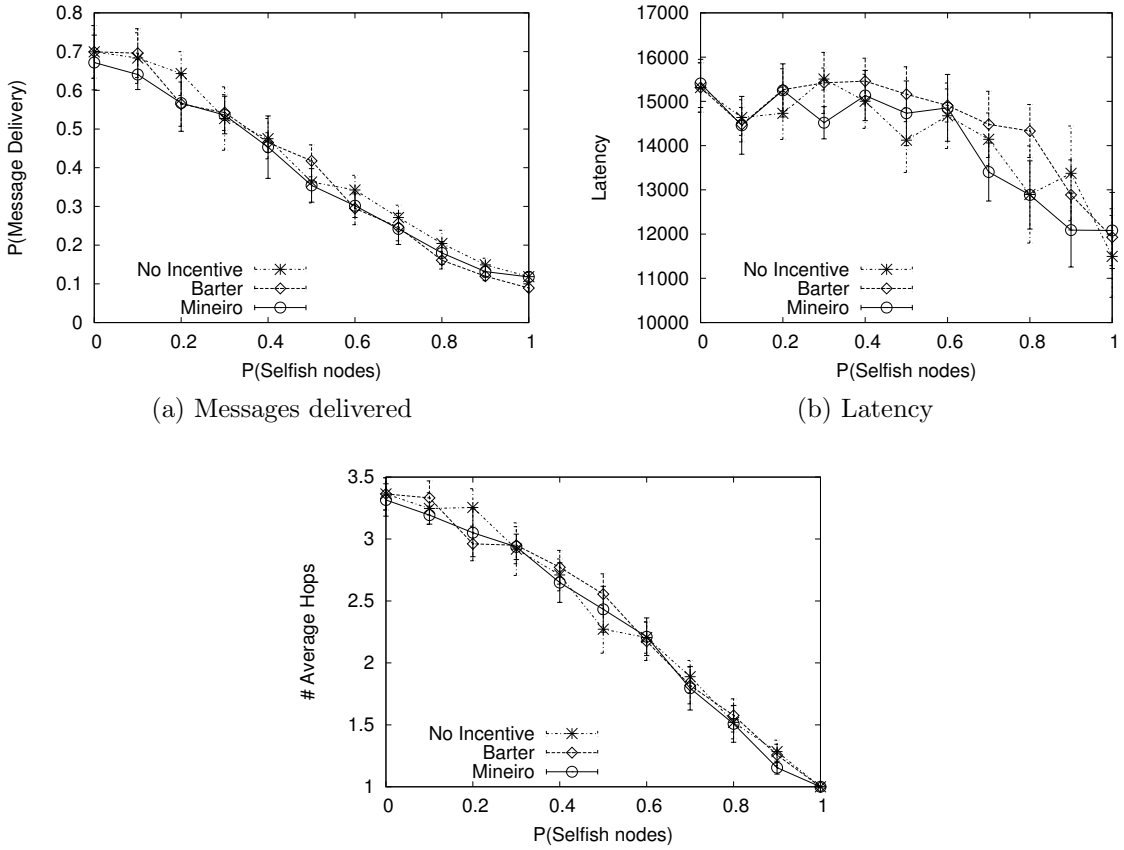
(c) Average hops per message

Figure 7.4. Non-forwarding selfish behavior in swim mobility model.

havior of every node in the network is to be *non-cooperative*. Again, cooperation arises when benefits is greater than costs. Therefore, the benefits to become *opportunistic* or *relay* must be clear for users. Table 7.2 describes intrinsic benefits and costs for each node behavior.

Based on the benefit-cost rule, cooperation in the opportunistic mobile offloading context arises when: *i*) users are *altruists*; *ii*) there is a natural willingness to cooperate; *iii*) *forced*, e.g. WISPs may force users to act as relays; *iv*) *technical*, users perceive performance improvements when cooperating; and *v*) *social*, cooperation brings social rewards.

In the best case, all users are altruists and always cooperate when the conditions to become a relay or opportunistic node hold. However, this assumption is unrealistic and users will cooperate only when forced or based on a clear benefit. Thus, in the next sections, we discuss how WISPs can encourage cooperation among 3G network participants, and how *Opplite* and *MINEIRO* working together may provide user-centric incentive mechanism.



(a) Messages delivered

(b) Latency

(c) Average hops per message

Figure 7.5. Dropping message selfish behavior in swim mobility model.

7.2.1 Data Reward to Engage Users' Cooperation

From the WISPs standpoint, offloading means decrease data consumption in their network. Indeed, there are redundant costs to delivering identical content to devices in the same vicinity. Opportunistic offloading helps WISPs improve their mobile data service provision without new infrastructure deployment. Therefore, cost reducing is a clear benefit for WISPs adopt opportunistic data offloading.

Furthermore, WISPs may improve their services, such as to provide connection for users with lack of coverage or at higher rates for users without devices that support the macrocell tier (e.g. devices without 4G). WISPs can further cooperate among themselves. In this case, cooperation means share network infrastructure or user devices as relays among them. Thus, customers without coverage or under link constraints in some WISP could take advantage of exploiting relays from other WISPs. Assuming WISPs have similar infrastructure and amount of customers, cooperation may be

[3]In cases where opportunistic devices can access only low-speed cellular network.

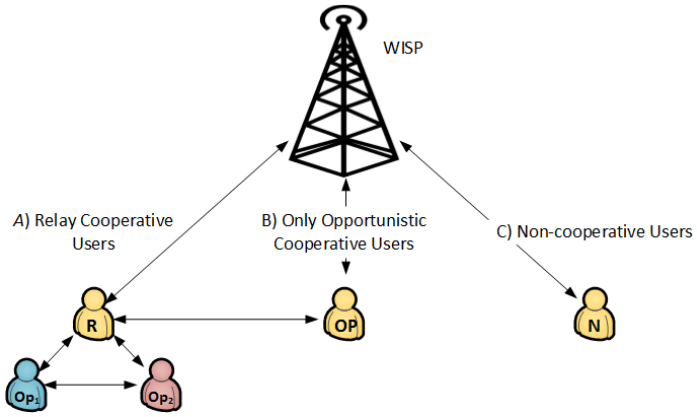


Figure 7.6. Actors and their behaviors. Opportunistic offloading relies on the assumption that a set of users will cooperate with others.

Node Behavior	Benefits	Costs
Non-Cooperative	Bandwidth Low Delay	3G/4G energy consumption Data consumption
Opportunistic	Save data consumption Save energy Improve bandwidth[3]	WiFi energy consumption Incurs Delay
Relay	Bandwidth Low Delay	WiFi energy consumption 3G/4G energy consumption Data consumption

Table 7.2. Benefit-cost for each node behavior in opportunistic offloading.

obtained by *reciprocity*. We left the analysis of cooperation among several WISPs as future works, and focused only on modeling the behavior between a WISP and its customers.

Service providers can force their customers to participate of the offloading process or can encourage their cooperation through rewards. *Forced-based* represents the easiest manner to obtain users' participation, especially whether WISP control all steps to select relays and to choose content desired to be offloaded. As mentioned in Chapter 3, the majority of previous work concerning algorithms to select nodes to act as relays rely on this assumption. However, in forced-based, users have their devices used in the offloading process without consent, which abstracts their willingness to cooperate or not.

WISPs play an important role to encourage user cooperation with *Reward-based* incentive mechanisms. These rewards may be monetary, offering discounts, free services or points in loyalty programs. However, the challenge is achieving a fair reward to

balance WISPs interests with users' expectations.

Plan pricing of mobile Internet are based on different advertised service speeds, monthly data allowances, charges contract periods and even by promotions and bundling strategies (as shown in Table 1.1). When users exceed their monthly data allowances, the common strategies are to blocking user data access or bandwidth throttling, which reduces users' bandwidth to low speeds. Thus, WISPs must design novel business models before they can push forward opportunistic offloading techniques.

Although novel billing or business models are out of scope of this thesis, we present below a naive data reward model where WISP provides data awards for its customers within a cost limit for the WISP, which overcomes the costs for users.

Consider an operator with a finite set of service plans $S = \{(P_1, D_1), ..., (P_j, D_j)\}$, where P_i is the fixed monthly price and D_i represents the maximum data allowance (MB). Let $N = 1, 2, ..., n$ be the set of devices, each device is mapped to a service plan $s \in S$, $n \rightarrow s(P, D)$, and it consumes $d \leq n^D$ MB per month. For simplicity, let the cost to provide one MB of data be $C(WISP) = n^P/n^D$. Thus, the monthly revenue of a WISP is given by:

$$R_{WISP} = \sum_{i=1}^{N} n_i^P - d_i \frac{n_i^P}{n_i^D} \tag{7.1}$$

In this simple model, if all devices consume all their data allowances the revenue for the operator is zero. Meanwhile, the maximum revenue is obtained when all users consume zero data of their service plan.

Data offloading occurs when a *relayer* - *R* device moves data directly to an *opportunistic* - *Op* device. Every time an opportunistic get a content k from a relay, it saves the size of the content ($|k|$ MB) from its monthly data allowance. The reward for the opportunistic node is given by how much data traffic was saved, that is it

$$\forall k \in K, R_{Opportunistic} = \sum_{i=1}^{K} |k_i| \tag{7.2}$$

Another benefit for opportunistic devices occur when they already exceeded their data allowance, in this case, to be an opportunistic device allows further utilization of the network. It is important to note, an opportunistic node needs to find a relay with the content it seeks instead of direct communication with the WISP, which incurs in some delay t. Therefore, there is a tradeoff between data transfer saving and delay expectation. Now, we tailor the utility function presented in Eq. 5.6 to define a satisfaction of an opportunistic node to wait for content within a delay $t_k \leq \tau$, limited

to τ, to find a suitable *relayer* node as:

$$u(t) = 1 - \frac{1}{1 + e^{\alpha(\frac{\tau}{2} - t)}} \tag{7.3}$$

The value of α defines the steepness of the function and the center, τ, represents the maximum delay tolerated. Hence, as faster an opportunistic node receives offloaded data greater the utility value.

Let consider as a reward for the relay, a proportional amount of the traffic it forwarded[4]. The data reward to relay can be expressed as a function of the size of the content forwarded $|k|$:

$$R_{Relay}(k) = \sum_{i=1}^{n} \beta^i |k| \tag{7.4}$$

In Equation 7.4, $\beta \in [0, 1]$ represents a discount factor, which it limits the reward to a maximum value and allows WISPs control how much to spend as incentive reward.

WISP saves $n|k|$ bytes of traffic, where n is the number of times a relay forwarded $|k|$. Thus, the reward offered by the WISP should be less or equal the traffic the *relayer* saved. In this case, the maximum value of β in which WISPs avoid cost explosion is given by:

$$n|k| \geq \sum_{i=1}^{n \to \infty} \beta^i |k|$$

$$\therefore$$

$$\beta \leq \frac{n}{1+n}$$

Assuming a relay forwards k at least once, $\beta = 0.5$, which makes $|k|$ be the upper bound reward for the *relayer*. Figure 7.7 shows relay rewards when offloading data $|k| = 100$ with two values of β.

OppLite has two parameters that allow users to define their willingness to switch to relay or opportunistic mode, Γ_{relay} and Γ_{opp}. These parameters are thresholds to define user awareness for cooperation in the decision-making algorithm. The decision algorithm in OppLite analyzes all criteria discussed in Chapter 5, and based on the threshold defined choose the communication mode. We can extend *OppLite* including the utility function 7.3 and β as criteria for the decision algorithm.

The delay tolerance criterion increases the aggregate function (weighted product

[4]WISP shall certify or authenticate each message forwarded by relay nodes due to security issues.

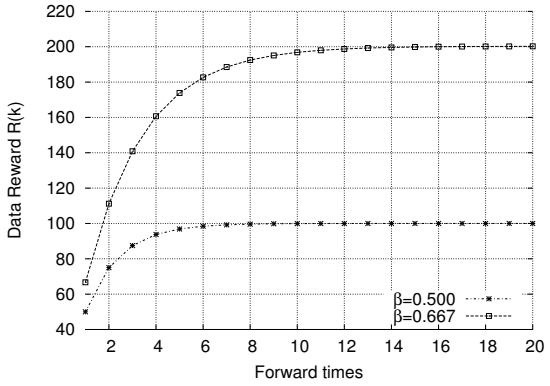


Figure 7.7. Reward received by the relay when forwarding data of size $|k| = 100$.

model) aiming to reach the threshold value (Γ_{opp}) faster, which defines users' willingness to become opportunistic nodes. The same occurs with the reward criterion for the relay, which users can set the expected reward (β) and its weight in the aggregate function.

WISPs can define β as a fixed value or run auctions where participants of the network submit bids with expected reward. In these cases, game theoretical tools may help to find a solution where user and WISP maximize their profit with no deviation.

Incentive mechanisms proposals are usually evaluated through theoretical validation or small controlled experiments. However, these experiments may not be able to predict with high accuracy the participation of users over time on the network. Thus, evaluate the real efficiency of these mechanisms may be costly and a time consuming task. We left modeling of reward models based on auction and the evaluation of this data reward mechanism as future work.

7.2.2 User Centric Incentive Mechanisms

In the previous section, WISPs take responsibility to engage user cooperation through rewards. This solution requires deployment of new infrastructure and billing mechanisms to control and manage rewards for users, which incurs extra costs to WISPs. Now, we aim to encourage cooperation with no or minimal WISP intervention.

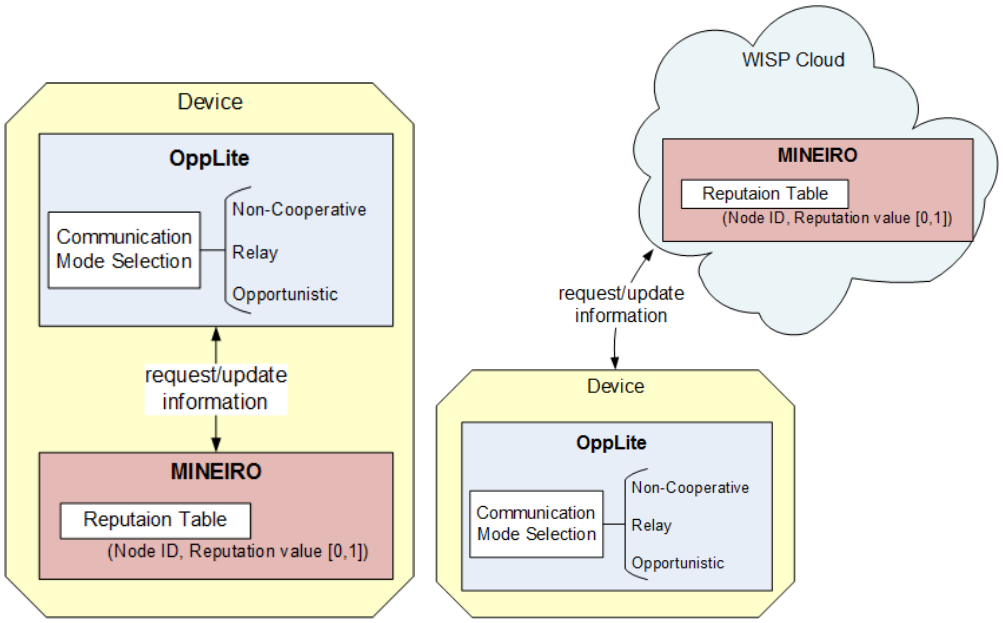
Advantages to become opportunistic devices include energy and monthly data allowance savings, which represent benefits with the potential to overcome delay costs.

Since a cooperative relay device has more costs than a non-cooperative device, some benefit that overcome costs must be considered. Assuming users are rational, *reciprocity* is the main driving force to achieve user willingness to turn his/her device as a relay with no or minimal WISP intervention.

Benefit in the *reciprocity* is given as equilibrium between how much a user offered

their resources to others and how much this user was benefited by others. In other words, all costs for a device as relay must be less or equal than the benefits obtained by a device while acting as opportunistic. As long as the majority of the nodes becomes relays when having enough resources and becomes opportunistic in case of scare resources, the reciprocity is reached. However, the main problem is how to determine whether a node is cooperative or it is just profiting by network resources.

In this sense, MINEIRO fills the gap of measuring nodes cooperation by keeping a reputation table with the previous behavior of the nodes. We extend OppLite architecture, shown in Figure 5.2, in order to deal with MINEIRO as a separate module, which contributes to the decision algorithm module. Therefore, MINEIRO assists OppLite decision module to determine the behavior of the node. Figure 7.8 shows two integrations between OppLite and MINEIRO: In a distributed fashion and in a centralized fashion.



(a) OppLite and MINEIRO Self contained in user device.

(b) MINEIRO runs on the cloud.

Figure 7.8. Distributed and centralized approaches to integrate OppLite and MINEIRO

In both types of integration, OppLite must request information about nodes before deciding its communication mode. MINEIRO was designed to engage message forwarding in generic opportunistic network. To integrate MINEIRO with OppLite we applied small changes in MINEIRO algorithm. Instead of increasing reputation of nodes forwarding third-party messages, MINEIRO increases reputation of a relay for

every data this relay offloaded. Contrarily, MINEIRO decreases opportunistic nodes' reputation whenever OppLite forwards data for it.

This integration makes the decision about the instantaneous behavior of a relay to become dynamic. Even whether all criteria satisfy the threshold to become relay, before cooperating with opportunistic nodes, the decision module requests to MINEIRO past behavior information about the opportunistic device. A relay node forwards data to an opportunistic node only when its reputation in MINEIRO is greater than zero.

In **distributed solution**, Fig. 7.8a, MINEIRO runs parallel to OppLite on the user device in order to provide information about past behavior of the devices encountered. At a first meeting, both nodes give an initial reputation for each other, relay cooperates and forwards data to opportunistic nodes. This initial cooperation is defined by MINEIRO. After this threshold, if an opportunistic node was always non-cooperative, the relay node denies to cooperate with it. Each MINEIRO node knows only the reputation of other nodes that interacted with it in the past. This approach requires no intervention or changes on the WISP side.

A drawback of this totally distributed solution is its scalability. In cellular networks, novel nodes can appear in the network any time, which it makes unfeasible to keep track about all nodes encountered. Furthermore, there might be situations where a pair of nodes meets only once in all network lifetime, for instance, people who do not know each other crossing each other in the cities. A centralized approach overcomes this drawback putting MINEIRO to run in the cloud, e.g. WISPs take responsibility to keep MINEIRO on their servers. This solution incurs minimal intervention on the WISP side, since MINEIRO can execute on the WISPs' servers.

In **centralized solution**, shown in Fig 7.8b, MINEIRO has a global view of the network. All nodes receive an initial reputation. Thus, nodes can be opportunistic some day and be relay days after, and its reputation will be updated on the server. An OppLite node in relay mode queries the reputation of an opportunistic node at the MINEIRO reputation table before forwarding data for it. A relay informs MINEIRO, through an update message, that it forwarded data for an opportunistic node. MINEIRO increases the reputation of relay nodes and decreases the reputation of opportunistic nodes accordingly to the algorithm 1.

One should note that this centralized approach incurs in data overhead to relay and security issues to avoid cheating. First, since OppLite needs to inform the server about message forwarding, these messages generate novel overhead and data consumption. This overhead can be diminished by aggregating information about data offloading during a period instead of sending messages for each data offloaded. For instance, a relay node sends only one message per day informing opportunistic node

identification and amount of data offloaded.

A malicious relay could update MINEIRO with false information about the amount of data offloaded. To avoid cheating, a relay must request a digitally signed receipt to the opportunistic node for every data forwarded. Relay forwards this receipt to MINEIRO, which checks its validity.

7.3 Conclusions

We proposed an incentive mechanism for generic opportunistic networking, called MINEIRO, which classifies the nodes based on the messages they forward. We modeled MINEIRO as a Bayesian game and showed the conditions in which the strategy profile of cooperation, that is, to forward third-party messages, leads to a Bayesian Equilibrium. Thus, MINEIRO provides a benefit for users, since forwarding messages for others improve its own performance.

Towards incentive mechanisms for opportunistic offloading, we showed how WISP can provide rewards for cooperative relaying within an upper cost bound. We show OppLite can be easily extended to support novel criteria that encourage users to act as relay or opportunistic nodes. Our naive data reward incentive mechanism makes clear the benefits for both relay and opportunistic nodes, while avoiding explosion cost on the WISP side.

Furthermore, we showed an integration between MINEIRO and OppLite in a decentralized and centralized fashion. In this case, cooperation brings technical benefit for users. Supposing the users are rational, users collaborate when they have enough resources to gain positive reputation and use its reputation as an opportunistic node to save its resources or improve bandwidth when necessary.

Beyond the technical challenges inherent for cooperative relaying, to convincing users switch their devices to relay mode also brings social and psychological challenges. An incentive mechanism is efficient if it recruits more participants and keeps these participants with cooperative behavior for a longer time. However, when proposing incentive mechanisms several challenges may be faced. For instance, costs and validation of the mechanism.

For effectiveness of monetary incentive mechanisms, it must be considered the costs for the WISP (or other central entity), as well, the earnings and costs for participants. However, finding and deciding a value that minimizes the cost to the WISP and at the same time, motivate the user requires further investigations.

Chapter 8

Conclusions and Future Work

This chapter summarizes this thesis and discusses directions for future research. Section 8.1 revisits the topics covered and the obtained results in this thesis. Section 8.2 presents the future work and research perspective.

8.1 Conclusions

Mobile Web applications are each day more bandwidth eager and as result, forecasts indicate an exponential mobile data traffic growth. Wireless Internet Services Providers (WISPs) have been facing a bottleneck to provide all resources required by their customers. To alleviate the cellular backhaul, WISPs can use mobile data offloading techniques, which means move off the data traffic from the cellular infrastructure to other types of network such as WiFi access points or opportunistic communication.

This dissertation discussed techniques to offload the struggled cellular network. Our three major contributions are: A study of the feasibility of WiFi infrastructure assisted offloading, a framework for opportunistic infrastructure-less offloading and for opportunistic networking demands users' cooperation, we developed incentive mechanisms to encourage cooperative behavior. Each contribution complements each other.

We conducted a case study to analyze the feasibility of cellular offloading through WiFi networks deployed in the cities. Nowadays, several WISPs have deployed thousands of WiFi access point worldwide in order to provide connectivity to their customers. We discussed how WiFi access points deployed in the cities can offload WISPs data traffic. Furthermore, we showed how much data traffic WISPs can offload through their already deployed WiFi infrastructure.

To achieve this goal, we implemented an application to Android devices and gathered information about access points and cellular network quality through several

bus routes in Paris, France. We showed that the city is almost entirely covered by private WiFi APs, public WiFi APs provided by government bodies and WISPs' WiFi APs. Our results showed that if all access points gathered were open access, WiFi APs offload up to 30% of the network traffic. However, most of WiFi access points require authentication and each WISP has its own APs deployment, making difficult to obtain the best offloading results in practice. Therefore, WISPs can explore better these already deployed WiFi APs, for instance, sharing their APs among themselves.

Opportunistic device-to-device communication arose as an approach to offload data traffic using devices carried by WISP customers. Here, the problem is how to select a subset of nodes that guarantees message delivery under a given delay constraint.

In this direction, we proposed *OppLite*, a user-awareness multi-criteria decision framework, which switches device communication mode based on the user preference. Devices can be in *standard*, *opportunistic* or *relay* mode. OppLite aims to provide opportunistic offloading using a set of devices as relay to forward data or to act as distributed cache and a set of devices as opportunistic to use relay devices instead of the cellular network. To achieve its goal, OppLite gathers information about device vicinity, battery lifetime and signal quality to infer through utility theory the device's communication mode. This set of criteria is extensible and novel criterion can be added just informing its utility function parameters, such as center, steepness and weight of the criterion.

To the best of our knowledge, OppLite is the first decentralized opportunistic offloading framework based on local information only.

We implemented three applications in a simulation environment to evaluate OppLite: *i*) *Opportunistic Relaying* (OpR), nodes in opportunistic mode forward their messages to infrastructure through relay nodes; *ii*) *Cache-and-Forward* (CaF), nodes in relay mode keep all content received in their buffer and deliver these contents for all nodes in opportunistic mode encountered; and *iii*) *Relay as Cache* (RaF), nodes in opportunistic mode request for a given content to nodes in relay mode and a relay node responds with the content requested if it owns the content in its cache. All applications consider a time out tolerance delay.

Overall, our evaluation shows that OppLite reaches higher offloading ratio in the following conditions: crowd scenarios, high numbers of nodes in relay mode and when nodes in opportunistic mode support high delay (10 minutes or more). Indeed, in these conditions, OppLite offloaded up to 70% of the messages in *OpR* application and 94% of the data traffic in *CaF* application. In *RaC* application, OppLite offloaded up to 44% of the data traffic.

Random solution, in which nodes switch to relay or opportunistic mode randomly,

achieved high offloading ratios in all applications. OppLite beats Random in all applications with low delay tolerance, and the contrary, Random beats OppLite in the most of cases the applications supported higher delay tolerance. However, Random switches much more nodes to relay or opportunistic mode than OppLite, which increases the number of nodes participating in the offloading process.

Therefore, OppLite can extend coverage through data relaying as shown in *OpR* and to distribute content efficiently, offloading data from the network infrastructure as shown in *CaF* and *RaC*.

Since opportunistic offloading requires users' willingness, we proposed incentive mechanisms and show how these mechanisms fit on the offloading problem. Our first incentive mechanism, called MINEIRO (***M****essage-based* ***IN****centive mechanism for* ***E****nd-user* ***I****mprovement of* ***R****outing* ***O****pportunities*), builds a reputation rank based on the source of messages received by the forwarding nodes. This algorithm was mainly deployed to understand general message forwarding incentive mechanisms. MINEIRO kept the delivery rate and the delay constant even in scenarios where more than 60% of the nodes present selfish behavior.

Towards incentive mechanisms to opportunistic mobile data offloading, we proposed mechanisms based on data reward offered by WISPs and further, an integration between OppLite and MINEIRO to provide user-centric incentive mechanisms. Data reward incentive mechanism provides clear benefits for both relay and opportunistic nodes, while avoiding explosion cost on the WISP side. Furthermore, OppLite integrated with MINEIRO creates reputation ranks, which allows users cooperate only with nodes with positive reputation. We proposed an integration between OppLite and MINEIRO in a centralized fashion, in which MINEIRO runs on the cloud and has a global view of nodes behavior; and in a decentralized fashion, in which both applications run on users' devices.

Assuming users are rational, they cooperate to gain positive reputation or rewards when they have enough resources, and use their reputation to switch to opportunistic mode in order to save their resources when necessary. However, cooperation in mobile network contexts is still an open issue and requires further investigations.

8.2 Future Work and Research Perspective

This section presents the future work based on our contributions. In this thesis, we presented how WiFi deployed in the city can help WISPs offload the data traffic, proposed a framework to offload the data traffic opportunistically through device-to-

device communication and provided incentive mechanisms to encourage users to adopt opportunistic communication. However, each study realized during the making of this thesis opened room for other research opportunities, such as novel analyses, OppLite extension and implementation in real environment. We describe below our research perspective:

8.2.1 WiFi Infrastructured Offloading

Chapter 4 analyzed WiFi deployment and cellular network only in one city. In order to understand better the connectivity in the cities, the same analysis can be done in other cities with different properties such as population size, Gross Domestic Product (GDP) or other traffic conditions.

We aim to publish the application developed to gather information about WiFi hotspots as open source. Hence, we can recruit more volunteers to collect data from several cities.

Volunteers spread over several cities would allow us to analyze and compare the WiFi connectivity in different cities. This research could shed light on which properties of the city make it more feasible to use WiFi connection instead of cellular network.

This research requires improvement of the application to gather information about access points, publication on the market for smart devices such as *Play Store* or *App Store*, volunteers recruiting and result analyses.

8.2.2 Opportunistic Mobile Offloading

Opportunistic mobile data offloading is a recent research topic which emerged in the last five years and requires further investigation to understand its pros and cons. Hence, we identified short and long term research based on this subject.

Chapter 5 and 6 proposed and evaluated OppLite, a framework to opportunistic offloading. As short term research, we plan to investigate the following extensions in OppLite:

- **Dynamic Self-Configuration**: Today, OppLite utilizes fixed thresholds to switch to relay or opportunistic mode. These thresholds could be dynamic considering the context of each node. We intend to make OppLite self-configurable based on the user context and relationship among users.
- **Novel Criteria**: Include novel criteria, such as monthly data allowance and monthly price, can improve users' willingness to cooperate since users that exceed their data allowance could switch their devices to opportunistic mode.

- **Applications on top of OppLite**: We presented three applications to show the feasibility of Opportunistic offloading. However, some applications can be extended to improve the offloading ratio of the network. In *CaF*, nodes in relay mode could forward messages in n-hop fashion. Furthermore, *RaC* application could use distributed hash tables to index content in relay nodes' cache, as used in distributed peer-to-peer networking.

Privacy and security are complex issues even in ad hoc network research. As long term research, we consider investigating these issues and the challenges they bring to the feasibility of opportunistic mobile data offloading. Furthermore, we aim to deploy a test-bed to evaluate our proposals.

- **Privacy**: Since the data is passing through unknown devices, the question that arises is: how to guarantee the privacy of the content requested or sent? In [Merwe et al., 2007], the authors classified privacy mechanisms as: *authority based*, where trusted authorities distribute keys; and *full-self-organized*, where mobile clients generate and manage the keys among themselves. In case of WISPs being the authorities for key distribution to their clients, OppLite would cypher the messages using the key, but this incurs in change of the infrastructure.

- **Security**: Opportunistic networks are somehow resilient to security attacks such as flooding and false route tables since they build multiple paths between source and destination. However, these attacks may drain the resources of mobile devices acting as *relays*. For instance, a flood attack to a *relay* could quickly drain its energy. In this case, by design OppLite would solve this problem changing the mode of a *relay* to *standard* based on the threshold. Moreover, malicious nodes can deliver false data. In this case, reputation mechanisms can be used to rank nodes delivering untruthful content or that act as spammers.

- **Implementation of OppLite in a Test-Bed**: Besides the efforts done to increase the reality of our simulations, we aim to implement OppLite in a real environment to assess its performance. We believe today's off-the-shelf devices have enough storage capacity to act as distributed cache and they are ready to support device-to-device communication through Bluetooth or WiFi in ad hoc mode.

8.2.3 Incentive Mechanisms Evaluation

Mechanism to engage users' cooperation can define the success of opportunistic offloading. Furthermore, the success of the techniques proposed to offload cellular data traffic depends mainly of WISP and user adoption. Hence, as long term research we plan a deeper investigation on users behavior characterization.

Chapter 7 proposed MINEIRO, a data reward and an integration between MINEIRO and OppLite. We aim to investigate and extend MINEIRO to become a cheat proof protocol against malicious users who want to increase their reputation by forwarding fake third-party messages. Furthermore, due to limitations such as time and resources, our evaluation is based on simulations.

Real experiments may shed light on user behavior. We intend to implement and deploy OppLite integrated with MINEIRO to analyze and to characterize incentive elements, such as reputation, rewards, among others. However, these types of experiments involve recruiting volunteers, which alone represents a challenge.

Bibliography

3GPPP (2013). Study on architecture enhancements to support proximity services (prose) (release 12). Report 23.703 V0.4.1, 3rd generation partnership project.

Abid, M., Yahiya, T., and Pujolle, G. (2012). A utility-based handover decision scheme for heterogeneous wireless networks. In *IEEE Consumer Communications and Networking Conference*, pages 650–654.

Adamopoulou, E., Demestichas, K., Koutsorodi, A., and Theologou, M. (2006). Intelligent access network selection in heterogeneous networks-simulation results. In *Wireless Conference -Enabling Technologies for Wireless Multimedia Communications*, pages 1--6.

Adar, E. and Huberman, B. A. (2000). Free riding on gnutella. *First Monday*, 5(10).

Ager, B., Schneider, F., Kim, J., and Feldmann, A. (2010). Revisiting cacheability in times of user generated content. In *INFOCOM IEEE Conference on Computer Communications Workshops, 2010*, pages 1--6. IEEE.

Akyildiz, I., Su, W., Sankarasubramaniam, Y., and Cayirci, E. (2002). Wireless sensor networks: a survey. *Computer networks*, 38(4):393--422.

Akyildiz, I. F., Akan, Ö. B., Chen, C., Fang, J., and Su, W. (2003). Interplanetary internet: state-of-the-art and research challenges. *Computer Networks*, 43(2):75 – 112. ISSN 1389-1286.

Akyildiz, I. F., Pompili, D., and Melodia, T. (2005). Underwater acoustic sensor networks: research challenges. *Ad Hoc Networks*, 3(3):257 – 279. ISSN 1570-8705.

Andrews, J., Claussen, H., Dohler, M., Rangan, S., and Reed, M. (2012). Femtocells: Past, present, and future. *IEEE Journal on Selected Areas in Communications*, 30(3):497 –508. ISSN 0733-8716.

Asadi, A. and Mancuso, V. (2013). Wifi direct and lte d2d in action. In *IFIP/IEEE Wireless Days*.

Balasubramanian, A., Mahajan, R., and Venkataramani, A. (2010). Augmenting mobile 3g using wifi. In *8th international conference on Mobile systems, applications, and services (MobiSys)*, pages 209–222.

Balasubramanian, N., Balasubramanian, A., and Venkataramani, A. (2009). Energy consumption in mobile phones: a measurement study and implications for network applications. In *9th Internet measurement conference (IMC)*, pages 280--293.

Barbera, M. V., Stefa, J., Viana, A. C., De Amorim, M. D., and Boc, M. (2011). Vip delegation: Enabling vips to offload data in wireless social mobile networks. In *Distributed Computing in Sensor Systems and Workshops (DCOSS), 2011 International Conference on*, pages 1--8. IEEE.

Ben Abdesslem, F. and Lindgren, A. (2014). Cacheability of youtube videos in cellular networks. In *Proceedings of the 4th workshop on All things cellular: operations, applications, & challenges*, pages 53--58. ACM.

Bigwood, G. and Henderson, T. (2011). Ironman: Using social networks to add incentives and reputation to opportunistic networks. In *Privacy, security, risk and trust Workshop in IEEE Socialcom*, pages 65--72. IEEE.

Bowles, S. and Gintis, H. (2003). *Origins of human cooperation. In: Genetic and cultural evolution of cooperation*. MIT Press Cambridge, MA.

Breslau, L., Cao, P., Fan, L., Phillips, G., and Shenker, S. (1999). Web caching and zipf-like distributions: Evidence and implications. In *INFOCOM'99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 1, pages 126--134. IEEE.

Burleigh, S., Hooke, A., Torgerson, L., Fall, K., Cerf, V., Durst, B., Scott, K., and Weiss, H. (2003). Delay-tolerant networking: an approach to interplanetary internet. *Communications Magazine, IEEE*, 41(6):128 – 136. ISSN 0163-6804.

Buttyán, L., Dóra, L., Félegyházi, M., and Vajda, I. (2010). Barter trade improves message delivery in opportunistic networks. *Ad Hoc Networks*, 8(1):1--14.

Bychkovsky, V., Hull, B., Miu, A., Balakrishnan, H., and Madden, S. (2006). A measurement study of vehicular internet access using in situ wi-fi networks. In

Proceedings of the 12th annual international conference on Mobile computing and networking, pages 50--61. ACM.

Chaintreau, A., Hui, P., Scott, J., Gass, R., Crowcroft, J., and Diot, C. (2007). Impact of human mobility on opportunistic forwarding algorithms. *IEEE Transactions on Mobile Computing*, 6(6):606--620.

Chandrasekhar, V., Andrews, J., and Gatherer, A. (2008). Femtocell networks: a survey. *Communications Magazine, IEEE*, 46(9):59--67.

Chen, X., Proulx, B., Gong, X., and Zhang, J. (2014). Exploiting social ties for cooperative d2d communications: A mobile social networking case. *Networking, IEEE/ACM Transactions on*, PP(99):1–1. ISSN 1063-6692.

Chen, Y., Towsley, D., Nahum, E. M., Gibbens, R. J., and Lim, Y. (2012). Characterizing 4g and 3g networks: Supporting mobility with multi-path tcp. Technical report, UMass Amherst Technical Report: UM-CS-2012-022.

Cheng, Z., Caverlee, J., Lee, K., and Sui, D. Z. (2011). Exploring millions of footprints in location sharing services. In *Fifth International AAAI Conference on Weblogs and Social Media (ICWSM)*.

Choi, Y., Yoon, C.-h., Kim, Y.-s., Heo, S. W., Silvester, J., et al. (2014). The impact of application signaling traffic on public land mobile networks. *Communications Magazine, IEEE*, 52(1):166--172.

Cisco (2013). Global mobile data traffic forecast update, 2012-2017. `http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html`.

Cisco (2015). Global mobile data traffic forecast update, 2014-2019. `http://www.cisco.com/assets/sol/sp/vni/forecast_highlights_mobile/index.html`.

Clausen, T. and Jacquet, P. (2003). Optimized link state routing protocol (olsr). RFC 3626, IETF.

Crawdad (2013). *Crawdad*. A Community Resource for Archiving Wireless Data At Dartmouth. http://crawdad.cs.dartmouth.edu/.

Crowcroft, J., Gibbens, R., Kelly, F., and Östring, S. (2004). Modelling incentives for collaboration in mobile ad hoc networks. *Performance Evaluation*, 57(4):427--439.

Deterding, S., Dixon, D., Khaled, R., and Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In *International Academic MindTrek Conference: Envisioning Future Media Environments*, pages 9--15. ACM.

Doppler, K., Rinne, M., Wijting, C., Ribeiro, C., and Hugl, K. (2009). Device-to-device communication as an underlay to lte-advanced networks. *Communications Magazine, IEEE*, 47(12):42--49.

Doppler, K., Yu, C.-H., Ribeiro, C. B., and Janis, P. (2010). Mode Selection for Device-To-Device Communication Underlaying an LTE-Advanced Network. *IEEE Wireless Communication and Networking Conference*, pages 1--6.

DTNRG (2013). *Delay-Tolerant Network Research Group*. DTN. http://www.dtnrg.org/.

Easley, D. and Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press.

Fall, K. (2003). A delay-tolerant network architecture for challenged internets. *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 27--34.

Fitzek, F. H., Heide, J., Pedersen, M. V., and Katz, M. (2013). Implementation of network coding for social mobile clouds [applications corner]. *Signal Processing Magazine, IEEE*, 30(1):159--164.

Gao, H., Liu, C., Wang, W., Zhao, J., Song, Z., Su, X., Crowcroft, J., and Leung, K. (2015). A survey of incentive mechanisms for participatory sensing. *Communications Surveys Tutorials, IEEE*, PP(99):1–1. ISSN 1553-877X.

Gass, R. and Diot, C. (2010). An experimental performance comparison of 3g and wi-fi. In *Passive and Active Measurement*, pages 71--80. Springer.

Gazis, V., Houssos, N., Alonistioti, N., and Merakos, L. (2003). On the complexity of always best connected in 4g mobile networks. In *58th IEEE Vehicular Technology Conference (VTC-Fall)*, volume 4, pages 2312–2316 Vol.4. ISSN 1090-3038.

Ghosh, A., Zhang, J., Andrews, J. G., and Muhamed, R. (2010). *Fundamentals of LTE*. Pearson Education.

Gil-Castineira, F., Gonzalez-Castano, F., and Franck, L. (2008). Extending vehicular can fieldbuses with delay-tolerant networks. *Industrial Electronics, IEEE Transactions on*, 55(9):3307--3314.

Haggle (2013). *Haggle Project*. Haggle. http://www.haggleproject.org.

Haldar, K. L., Li, H., and Agrawal, D. P. (2013). A cluster-aware soft frequency reuse scheme for inter-cell interference mitigation in lte based femtocell networks. In *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2013 IEEE 14th International Symposium and Workshops on a*, pages 1--6. IEEE.

Hale, R. V. (2004). Wi-fi liability: Potential legal risks in accessing and operating wireless internet. *Santa Clara Computer & High Tech. LJ*, 21:543.

Han, B., Hui, P., Kumar, V., Marathe, M., Shao, J., and Srinivasan, A. (2012). Mobile data offloading through opportunistic communications and social participation. *IEEE Transactions on Mobile Computing*, 11(5):821–834.

Han, J. and Moraga, C. (1995). The influence of the sigmoid function parameters on the speed of backpropagation learning. In *From Natural to Artificial Neural Computation*, pages 195--201. Springer.

Huang, J., Qian, F., Gerber, A., Mao, Z. M., Sen, S., and Spatscheck, O. (2012). A close examination of performance and power characteristics of 4g lte networks. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*, MobiSys '12, pages 225--238, New York, NY, USA. ACM.

Hui, P., Chaintreau, A., Scott, J., Gass, R., Crowcroft, J., and Diot, C. (2005). Pocket switched networks and human mobility in conference environments. In *ACM SIGCOMM workshop on Delay-tolerant networking (WDTN)*, pages 244–251.

Jacquet-Lagreze, E. and Siskos, J. (1982). Assessing a set of additive utility functions for multicriteria decision-making, the uta method. *European journal of operational research*, 10(2):151--164.

Juang, P., Oki, H., Wang, Y., Martonosi, M., Peh, L., and Rubenstein, D. (2002). Energy-efficient computing for wildlife tracking: design tradeoffs and early experiences with zebranet. *ACM SIGPLAN Notices*, 37(10):96--107.

Kamenica, E. (2012). Behavioral economics and psychology of incentives. *Annu. Rev. Econ.*, 4(1):427--452.

Katabi, D., Handley, M., and Rohrs, C. (2002). Congestion control for high bandwidth-delay product networks. In *Proceedings of the conference on applications, technologies, architectures, and protocols for computer communications (SIGCOMM)*, pages 89–102.

Keränen, A., Ott, J., and Kärkkäinen, T. (2009). The one simulator for dtn protocol evaluation. In *Proceedings of the 2nd International Conference on Simulation Tools and Techniques*, Simutools '09, pages 55:1--55:10, ICST, Brussels, Belgium, Belgium. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Khabbaz, M., Assi, C., and Fawaz, W. (2011). Disruption-tolerant networking: A comprehensive survey on recent developments and persisting challenges. *IEEE Communications Surveys & Tutorials*, 14(2):1–34. ISSN 1553-877X.

Lampropoulos, G., Passas, N. I., Merakos, L. F., and Kaloxylos, A. (2005). Handover management architectures in integrated wlan/cellular networks. *IEEE Communications Surveys and Tutorials*, 7(1-4):30--44.

Lee, J.-S. and Hoh, B. (2010). Dynamic pricing incentive for participatory sensing. *Pervasive and Mobile Computing*, 6(6):693--708.

Lee, K., Lee, J., Yi, Y., Rhee, I., and Chong, S. (2010). Mobile data offloading: how much can WiFi deliver? In *6th International Conference on emerging Networking EXperiments and Technologies (CoNEXT)*, pages 26:1–26:12.

Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of economic dynamics*, 1(3):593--622.

Lin, X., Andrews, J. G., Ghosh, A., and Ratasuk, R. (2013). An overview on 3gpp device-to-device proximity services. *arXiv:1310.0116*.

Manfredi, V., Crovella, M., and Kurose, J. (2011). Understanding stateful vs stateless communication strategies for ad hoc networks. In *Proceedings of the 17th annual international conference on Mobile computing and networking*, MobiCom '11, pages 313--324, New York, NY, USA. ACM.

Marti, S., Giuli, T. J., Lai, K., and Baker, M. (2000). Mitigating routing misbehavior in mobile ad hoc networks. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, MobiCom '00, pages 255--265, New York, NY, USA. ACM.

Martinez, F., Toh, C., Cano, J., Calafate, C., and Manzoni, P. (2011). A survey and comparative study of simulators for vehicular ad hoc networks (vanets). *Wireless Communications and Mobile Computing*, 11(7):813--828.

Mei, A. and Stefa, J. (2009). Swim: A simple model to generate small mobile worlds. In *IEEE INFOCOM*, pages 2106--2113. IEEE.

Merwe, J. V. D., Dawoud, D., and McDonald, S. (2007). A survey on peer-to-peer key management for mobile ad hoc networks. *ACM computing surveys*, 39(1):1.

Miao, J., Hasan, O., Mokhtar, S. B., Brunie, L., and Yim, K. (2013). An investigation on the unwillingness of nodes to participate in mobile delay tolerant network routing. *International Journal of Information Management*, 33(2):252--262. ISSN 02684012.

Min, H., Lee, J., Park, S., and Hong, D. (2011). Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks. *Wireless Communications, IEEE Transactions on*, 10(12):3995--4000.

Mota, V. F., Cunha, F. D., Macedo, D. F., Nogueira, J. M., and Loureiro, A. A. (2014). Protocols, mobility models and tools in opportunistic networks: A survey. *Computer Communications*, 48:5--19.

Mota, V. F. S., Macedo, D. F., Ghamri-Doudane, Y., and Nogueira, J. M. S. (2013). On the feasibility of wifi offloading in urban areas: The paris case study. In *IFIP Wireless Days conference*.

Mukherjee, S. (2011). Analysis of ue outage probability and macrocellular traffic offloading for wcdma macro network with femto overlay under closed and open access. In *Communications (ICC), 2011 IEEE International Conference on*, pages 1--6. IEEE.

Nash, J. (1951). Non-cooperative games. *Annals of mathematics*, pages 286--295.

Nguyen-Vuong, Q., Ghamri-Doudane, Y., and Agoulmine, N. (2008). On utility models for access network selection in wireless heterogeneous networks. In *IEEE Network Operations and Management Symposium (NOMS)*, pages 144–151.

Ormond, O., Murphy, J., and Muntean, G.-M. (2006). Utility-based intelligent network selection in beyond 3g systems. In *IEEE International Conference on Communications (ICC)*, volume 4, pages 1831–1836.

Osseiran, A., Braun, V., Hidekazu, T., Marsch, P., Schotten, H., Tullberg, H., Uusitalo, M., Schellman, M., et al. (2013). The foundation of the mobile and wireless communications system for 2020 and beyond: Challenges, enablers and technology solutions. In *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*, pages 1--5. IEEE.

Panagakis, A., Vaios, A., and Stavrakakis, I. (2007). On the effects of cooperation in dtns. In *Communication Systems Software and Middleware, 2007. COMSWARE 2007. 2nd International Conference on*, pages 1--6. IEEE.

Partan, J., Kurose, J., and Levine, B. N. (2006). A survey of practical issues in underwater networks. In *Underwater Networks*, pages 17–24. ACM.

Paul, S., Yates, R., Raychaudhuri, D., and Kurose, J. (2008). The cache-and-forward network architecture for efficient mobile content delivery services in the future internet. In *Innovations in NGN: Future Network and Services, 2008. K-INGN 2008. First ITU-T Kaleidoscope Academic Conference*, pages 367--374. IEEE.

Perkins, C. and Bhagwat, P. (1994). Highly dynamic destination-sequenced distance-vector routing (dsdv) for mobile computers. *Proceedings of the conference on Communications architectures, protocols and applications*, pages 234--244.

Perkins, C. E. and Royer, E. M. (1999). Ad-hoc on-demand distance vector routing. *proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications*, 2:90--100.

Qiao, J., Shen, X. S., Mark, J. W., Shen, Q., He, Y., and Lei, L. (2015). Enabling device-to-device communications in millimeter-wave 5g cellular networks. *Communications Magazine, IEEE*, 53(1):209--215.

Qualcomm (2012). Study on lte device to device proximity service. Report 122009, 3rd generation partnership project.

Ramadas, M., Burleigh, S., and Farrell, S. (2008). Licklider transmission protocol - specification. RFC 5326, IETF.

Reddy, S., Estrin, D., Hansen, M., and Srivastava, M. (2010). Examining micro-payments for participatory sensing data collections. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, Ubicomp '10, pages 33--36, New York, NY, USA. ACM.

Sesia, S., Toufik, I., and Baker, M. (2009). *LTE: the UMTS long term evolution*. Wiley Online Library.

Sinha, S. and Pan, R. K. (2006). How a hit is born: The emergence of popularity from the dynamics of collective choice. *Econophysics and Sociophysics: Trends and Perspectives*, pages 417--447.

Sinnott, R. W. (1984). Virtues of the haversine. *Sky and Telescope*, 68(2):159+.

Small, T. and Haas, Z. (2003). The shared wireless infostation model: a new ad hoc networking paradigm (or where there is a whale, there is a way). In *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pages 233--244. ACM.

Tournoux, P., Leguay, J., Benbadis, F., Conan, V., De Amorim, M., and Whitbeck, J. (2009). The accordion phenomenon: Analysis, characterization, and impact on dtn routing. In *IEEE INFOCOM*, pages 1116--1124.

Triantaphyllou, E. and Mann, S. H. (1989). An examination of the effectiveness of multi-dimensional decision-making methods: a decision-making paradox. *Decision Support Systems*, 5(3):303--312.

Tsolkas, D., Liotou, E., Passas, N., and Merakos, L. (2013). Enabling d2d communications in lte networks. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*, pages 2846–2850.

Turlenet (2013). *UMass TurtleNet*. UMASS, http://prisms.cs.umass.edu/dome/turtlenet.

Vahdat, A. and Becker, D. (2000). *Tec. Report, Duke University*.

Valerio, L., Abdesslemy, F. B., Lindgreny, A., Bruno, R., Passarella, A., and Luoto, M. (2015). Offloading cellular traffic with opportunistic networks: a feasibility study. In *Ad Hoc Networking Workshop (MED-HOC-NET), 2015 14th Annual Mediterranean*, pages 1--8. IEEE.

Vaughan-Nichols, S. J. (2004). Achieving wireless broadband with wimax. *Computer*, 37(6):10--13.

von Neumann, J. and Morgenstern, O. (1953). *The Theory of Games and Economic Behavior*. Princeton University Press.

Wang, R., Taleb, T., Jamalipour, A., and Sun, B. (2009). Protocols for reliable data transport in space internet. *IEEE Communications Surveys and Tutorials*, 11(2):21–32.

Wei, L., Cao, Z., and Zhu, H. (2011). MobiGame: A User-Centric Reputation Based Incentive Protocol for Delay/Disruption Tolerant Networks. *Communications Society*.

Werbach, K. and Hunter, D. (2012). *For the win: How game thinking can revolutionize your business*. Wharton Digital Press.

Whitbeck, J., Lopez, Y., Leguay, J., Conan, V., and Amorim, M. (2012). Push-and-track: Saving infrastructure bandwidth through opportunistic forwarding. *Pervasive and Mobile Computing*, 8(5):682 – 697. ISSN 1574-1192.

Yang, D., Xue, G., Fang, X., and Tang, J. (2012). Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing. In *International conference on Mobile computing and networking (Mobicom)*, pages 173--184.

Yu, C.-H., Doppler, K., Ribeiro, C. B., and Tirkkonen, O. (2011). Resource sharing optimization for device-to-device communication underlaying cellular networks. *Wireless Communications, IEEE Transactions on*, 10(8):2752--2763.

Zhang, S., Wu, J., Qian, Z., and Lu, S. (2015). Mobicache: Cellular traffic offloading leveraging cooperative caching in mobile social networks. *Computer Networks*, 83(0):184 – 198. ISSN 1389-1286.

Zhang, Z. (2006). Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges. *IEEE Communications Surveys & Tutorials*, 8(1):24--37.

Zhuo, X., Gao, W., Cao, G., and Hua, S. (2014). An incentive framework for cellular traffic offloading. *Mobile Computing, IEEE Transactions on*, 13(3):541--555.