MAC protocols design and a cross-layered QoS framework for next generation wireless networks

Essaïd Sabir

THÈSE

Présentée pour obtenir le grade de Docteur en Sciences de l'Université d'Avignon et des Pays de Vaucluse France & de l'Université Mohammed V-Agdal Rabat - Maroc

SPÉCIALITÉ : Informatique

École Doctorale 166 « Information Structures Systèmes »
Laboratoire d'Informatique d'Avignon (UPRES No 4128)

MAC Protocols Design and a Cross-layered QoS Framework for Next Generation Wireless Networks

par

Essaïd Sabir

Soutenue publiquement le 24 septembre 2010 devant un jury composé de :

M.	BOUYAKHF El-Houssine	Professeur, LIMIARF/FSR, Rabat	Président
M.	ALTMAN Eitan	Directeur de recherche, INRIA, Sophia Antipolis	Directeur
M.	EL-AZOUZI Rachid	Maître de Conférences, LIA, Avignon	Co-directeur
M.	CHAHED Tijani	Professeur, Telecom SudParis, Paris	Rapporteur
M.	ELKOUTBI Mohammed	Professeur, ENSIAS, Rabat	Rapporteur
M.	EL-OUAHIDI Bouabid	Professeur, Université Mohammed V-Agdal, Rabat	Rapporteur
M.	HAYEL Yezekael	Maître de Conférences, LIA, Avignon	Examinateur
M.	BELHADJ Abdenabi	Responsable Veille Technologique, Maroc Telecom	Invité

Laboratoire LIA, Avignon *Laboratoire LIMIARF*, Rabat

Preface

The research work of this doctoral thesis has been carried out in the Laboratoire d'Informatique Mathématiques Intelligence Artificielle et Reconnaissance de Formes (LIMIARF), University of Mohammed V-Agdal, Rabat, Morocco and Laboratoire Informatique d'Avignon (LIA), University of Avignon, France; During the years 2007-2010.

I am profoundly thankful to my advisors, Professor El-Houssine Bouyakhf and Associate Professor Rachid El-Azouzi. I will always be indebted for all the support, guidance, encouragement, their friendly mood that makes it enjoyable to work with them, and understanding they have given me throughout these years. It was a real pleasure to work with someones so highly motivated, smart, enthusiastic, and passionate about their work. Many thanks go to my co-advisors Professor Eitan Altman and Associate Professor Yezekael Hayel for their guidance that helped me continue my research in the right direction and the invaluable suggestions, discussions, and wonderful feedback. It is a great honor to work with you.

I am most grateful to my reviewers Professor Bouabid El Ouahidi (Faculty of sciences, Rabat), Professor Mohammed El Koutbi (ENSIAS, Rabat), and Professor Tijani Chahed (Telecom SudParis, France) for their thorough examination of the dissertation. My thanks go also Doctor Abdenabi Belhadj (R&D, Maroc Télécom). Their detailed comments significantly improved the quality of this thesis.

I thank all past and present colleagues at the LIA and LIMIARF Labs. for the joyful and pleasant working environment. In particular, I would like to thank Nabil, Lahcen, Khalil, Abdellatif, Amar, Tembiné, Sihame, Raïss, Baslam, Tania, Julio, Solan, Oussama, Mohammed, Ralph, Thierry, and Sujit. To the great staffs of LIA and LIMIARF, I am very thankful to all of them.

I would like to thank my wife Sihame Mercha for her unconditional support throughout these three years, especially during all the difficult times we went through while we had to live apart from each other in a foreign country for so long. In spite of all difficulties, I am extremely proud of all we have accomplished together in this long journey. For sure, this doctorate would have not been possible without her tremendous encour-

agement, love, patience, companionship and deep trust in me.

Most importantly, my deepest gratitude also goes to my parents Fatima Badri and Mohammed Sabir who have always filled my life with generous love, and unconditional support. This doctorate is truly theirs! My thanks go also to my lovely sisters and brothers, my uncle Saleh Sabir and all my family for their endless moral support throughout my career. To them I dedicate this thesis.

Finally, I want to thank the financial support I have received from the Moroccan-French cooperation driven by the Pôle de Compétences en Sciences et Technologies de l'Information et de la Communication (STIC). I would also thank the financial support I have received from the Maroc Télécom R&D project under grant N^o 10510005458.06PI and the European WINEM project.

Rabat, September 24, 2010

Essaïd Sabir.

Abstract

The present dissertation deals with the under-utilization problem of medium access control in wireless collision channels and other closely related problems known in wireless networks. It deals with the design of random access protocols for wireless systems and provides a mathematical framework for performance evaluation of multi hop based heterogeneous wireless networks. A new modeling framework is also introduced for the analytical study of MAC protocols operating in multihop wireless ad hoc networks, i.e., wireless networks characterized by the lack of any pre-existent infrastructure and where participating devices must cooperatively provide the basic functionalities that are common to any computer network. To show the applicability of our modeling framework, we model wireless ad hoc networks that operate according to the IEEE 802.11 standard. To accomplish this, we present a comprehensive analytical modeling of the IEEE 802.11 and the derivation of many performance metrics of interest, such as delay, throughput, and energy consumption.

The rest of this dissertation is divided into three parts. The first part comprises Chapters 1, 2 and 3. We first propose and evaluate four new power control enabled algorithms of slotted aloha with priority and capture effect. Both team problem (common objective function is maximized) and game problem (each user maximizes its own objective) were discussed. Extensive simulations were important to understand the behavior of such a system and the real impact of involved parameters (transmit power, transmit rate, arrival rate). Next, we present a new hierarchical slotted aloha version. Indeed, we consider two classes of users (leaders and followers) and compute the Stackelberg equilibrium of the constructed game. Introducing hierarchy seems to provide many promising improvement without (virtual controller) or with a low amount of external/common information (the case of several leaders). Later we analyzed a collision channel system where each user has some throughput demand to fulfill in order to maintain its service. Considering non saturated users, we showed existence of infinitely many Nash equilibria. Another interesting feature is that the stability region coincides with the Nash equilibria region. In this context and regarding the energy investment, we showed existence of an efficient Nash equilibrium for all active users. Later, we proposed two distributed algorithms to converge to the best Nash equilibrium point. The first algorithm is derived from the best response strategy, whereas the second learning algorithm is fully distributed and uses only the user's own information.

The second part comprising Chapters 4 and 5, is devoted to issues related to perfor-

mance evaluation of heterogeneous wireless networks composed of a cellular system extended by a multi-hop ad hoc network. We based our study on an analytical model that takes into account topology, routing, random access in MAC layer, forwarding probability and a finite retries per packet per path. We distinguish three key features of the network model that make our contribution in this thesis novel when considered all together. First, packet scheduling in the network layer. Using a Weighted Fair Queueing, we mainly addressed the cooperation effect and the stability region of forwarding probabilities. Second, the asymmetry of the multi-hop heterogeneous network in terms of topology, traffic and nodes intrinsic parameters. Lastly, we built a cross-layer architecture that makes important information available to concerned layers and allows to benefit from this latter. The case of a homogeneous ad hoc network study is also presented to derive the distribution of delay, it could be straightforwardly extended to heterogeneous networks.

In a spirit similar to that of part 2, the third part presents a more realistic cross-layered model. We indeed developed a new analytical modeling of the IEEE 802.11e DCF/EDCF in the context of multi-hop ad hoc networks. We formulated two coupled systems for the NETWORK and the MAC layers. The attempt rate and collision probabilities are now functions of the traffic intensity, of topology and of the routing decision. For more generality, we also considered a finite retries per packet per path. On one hand, this latter is responsible of asymmetric queue service rate and therefore of service time distribution. On the other hand, it has a direct impact on the performance of a loaded network. Finally, we showed how to take benefit from the interaction between NETWORK, MAC and PHY layers. Extensive simulations and numerical results are carried out to assist and confirm our work. A Fountain code-based MAC layer is also proposed to improve the throughput and establish better fairness properties. Our scheme aims also to reduce the expected number of retransmissions per packet and takes benefit from incremental redundancy of previous received copies of a given packet. Through several simulations, we showed that using Fountain codes in the context of multi-hop ad hoc networks improves the throughput of paths that suffer from bad channel conditions, i.e., high collision probability. It may nevertheless decrease it on other paths. However this latter scheme achieves better fairness index (Jain's index) in the whole network, so a throughput/fairness tradeoff can be efficiently defined.

Résumé et organisation de la thèse

Mots clés : WiMAX, UMTS, IEEE 802.11, Réseau hétérogène, Couche MAC, Aloha, Théorie des jeux, Théorie des chaines de Markov, Théorie de l'ordonnancement, Architecture inter-couches, Evaluation de performances, Débit, Délai, Stabilité, Simulation.

Désormais, les organismes de normalisation, les équipementiers, les fournisseurs de contenu ainsi que les opérateurs ont commencé à avoir la coutume de passer à une nouvelle génération de systèmes de télécommunication au bout de chaque décennie. A ce rythme, les futures générations de réseaux seront donc inévitablement hétérogènes ou encore ubiquitaires, qui se caractérisent par des entités mobiles (terminaux, routeurs, PDA, stations de base, téléphones cellulaires, etc.) communicantes et de taille parfois très variable. De tels systèmes posent des problèmes de mobilité, de sécurité et de sûreté (confidentialité des données, fiabilité des applications), de continuité de services (serviabilité, tolérance aux pannes, reconfigurabilité, etc.), de passage à l'échelle et de qualité de service. Faire fonctionner ces réseaux hétérogènes complexes (hétérogénéité des infrastructures, protocoles, applications, etc.) nécessite le développement de recherches fondamentales et appliquées en conception des architectures et des protocoles, ainsi qu'en dimensionnement, optimisation et planification des réseaux. Encore faut-il trouver des réponses à la problématique de la coopération ou la concurrence entre technologies. La conception et l'utilisation de tels systèmes informatiques constituent donc un défi scientifique majeur pour les prochaines années auquel nous avons décidé d'y apporter notre contribution.

Ce manuscrit est centré sur la conception, l'amélioration et l'évaluation des protocoles des couches RESEAU, MAC et PHY. En particulier, nous nous focalisons sur la conception de nouveaux protocoles distribués pour une utilisation optimale/améliorée des ressources radio disponibles. Par ailleurs, nous caractérisons les performances des réseaux ad hoc à accès aléatoire au canal en utilisant des paramètres de plusieurs couches avec aptitude de transfert d'information (data forwarding). La majeure partie de nos analyses se base sur le concept d'interaction entre les couches OSI (cross-layer). En effet, cette nouvelle et attractive approche est devenue en peu de temps omniprésente dans le domaine de recherche et développement et dans le domaine industriel. Les métriques de performances qui nous intéressent sont la stabilité des files d'attentes de transfert, le débit, le délai et la consommation d'énergie. Principale-

ment, la compréhension de l'interaction entre les couches MAC/PHY et routage du standard IEEE 802.11e DCF/EDCF, d'une part, et l'interaction entre nœuds en terme d'interférences, d'autre part, constituent le cœur central de notre travail. Cette thèse est divisée en sept chapitres contextuellement structurés en trois parties.

Amélioration des techniques d'accès aléatoire sans fil : Contrôle de puissance, Hiérarchie entre utilisateurs et Apprentissage

Dans la première partie, nous proposons et analysons de nouvelles versions distribuées du fameux protocole slotted aloha. En particulier ces nouvelles versions incorporent la diversité de puissance, le contrôle de puissance, la priorité entre paquets, la hiérarchie entre utilisateurs, la qualité de service et ne se restreint pas aux cas particulier des nœuds saturés massivement utilisée dans la littérature.

Slotted aloha coopératif et non coopératif avec priorité, diversité de puissance et effet de capture : Communiquant via slotted aloha standard, les utilisateurs transmettent toujours à la même puissance. Ceci implique la perte de tous les paquets entrés en collision. Notre protocole amélioré suppose que N niveaux de puissance sont disponibles pour les transmissions. Ainsi, chaque utilisateur choisit aléatoirement un niveau de puissance avant de transmettre son prochain paquet. Les niveaux de puissance disponibles sont supposés communs à tous les utilisateurs et sont donnés par le vecteur $\mathcal{P} = [p_1, p_2, \cdots, p_N]$. Par conséquent, même impliqué dans une collision, le paquet transmis avec la plus grande puissance pourrait être correctement décodé si son rapport signal sur bruit (RBS) est supérieur à un certain seuil. Nous distinguons deux types de paquets : 1) Les nouveaux paquets, c-à-d les paquets nouvellement entrés dans le système en vue d'être transmis, et 2) les paquets backloggés, c-à-d les paquets qui sont entrés en collision avec d'autres transmissions simultanées et qui attendent d'être retransmis après un temps aléatoire. Cette différenciation permet d'établir une certaine priorité en terme du niveau de puissance utilisé pour la transmission du paquet en question. Ensuite, nous étudions deux configurations : 1) Problème d'équipe où tous les utilisateurs maximisent la même fonction objective (débit, délai ou une combinaison convexe de ces deux métriques), et 2) problème du jeu où chaque utilisateur maximise sa propre fonction d'utilité. Ici, nous supposons que les usagers sont rationnels et donc agissent de telle sorte pour maximiser leur propre profit. Quatre algorithmes sont alors proposés :

- Choix aléatoire de la puissance de transmission parmi les N niveaux disponibles. Aucune priorité entre paquets n'est considérée dans cet algorithme.
- Retransmission avec plus de puissance. Les nouveaux paquets sont toujours transmis avec la plus faible puissance p_1. Tandis que les paquets backloggés sont prioritaires et sont retransmis avec une puissance choisie aléatoirement parmi les $N - 1$ plus hauts niveaux.

- Retransmission avec moins de puissance. Les nouveaux paquets sont toujours transmis avec la plus grande puissance p_N. Tandis que les paquets backloggés sont retransmis avec une puissance choisie aléatoirement parmi les $N-1$ plus faibles niveaux.
- Retransmission avec la plus faible puissance. Les paquets backloggés sont toujours transmis avec la plus faible puissance p_1. Tandis que les paquets backloggés sont retransmis avec une puissance choisie aléatoirement parmi les $N-1$ plus grands niveaux.

Nous modélisons le système par un jeu stochastique où le nombre de paquets backloggés est pris comme état du système. Puis, nous calculons la chaine de Markov associée à chaque algorithme. La distribution stationnaire est ensuite calculée et utilisée pour déterminer le débit moyen ainsi que le délai moyen. En terme de stabilité et de probabilité de succès, nos algorithmes sont toujours meilleurs que slotted aloha classique. En outre et contrairement à slotted aloha, ils peuvent ne pas souffrir du problème de bistabilité. Cependant, à partir d'une certaine valeur de la probabilité de retransmission, même nos algorithmes deviendraient bistables. Une multitude d'exemples numériques est fournie en fin de ce chapitre et témoigne de l'intérêt de nos algorithmes. Opérant avec nos algorithmes, les mobiles sont généralement moins agressifs en comparaison avec slotted aloha où les mobiles transmettent avec probabilité 1 à moyen et fort trafic. Ceci a pour effet de réduire les collisions et par la suite garantir un débit non nul et un délai borné.

L'équilibre de Stackelberg et la hiérarchie entre utilisateurs dans slotted aloha : Nous reprenons le même modèle développé dans le chapitre 1 et y introduisons la notion de hiérarchie entre utilisateurs. Slotted aloha est connu par une chute du débit à moyen et fort trafic, ceci est dû au fait que les utilisateurs deviennent très agressifs à l'équilibre et transmettent avec probabilité proche de 1. Notre idée de base étant de réduire le taux de collision en réduisant le nombre d'utilisateurs accédants simultanément au canal. Avec une formulation en un jeu de Stackelberg, les utilisateurs sont divisés en deux groupes : 1) les maîtres (leaders) ayant la hiérarchie haute, et 2) les esclaves (followers) qui sont des subordonnés dont la stratégie est directement dépendante de celle des maîtres. En réalité le jeu est récursive dans le sens où les leaders jouent un équilibre de Nash sachant (prédisant) le profile de stratégies qui serai décidé par le groupe des followers. Le profile de stratégies décidé par les followers est en fait un équilibre de Nash connaissant la décision des leaders. En fin de compte, il est clair que ce sont les leaders qui décident réellement du sort du jeu. Le point d'équilibre est appelé équilibre de Stackelberg. Pour une raison de degré de liberté (nombre de paramètre agissant sur le système) et de faisabilité de calcul, nous nous restreindrons au équilibre de Stackelberg symétrique.

Puis, nous discutons l'implémentation de ce genre de protocole dans des réseaux réels. En effet nous avons défini le concept de « contrôleur virtuel » qui permettra de régulariser, d'une manière distribuée, le taux de transmission des autres usagers. Son inter-

vention serait donc faire des transmissions qui ont pour but de bruiter le canal, ainsi les usagers vont détecter une mauvaise qualité du canal et donc se rendre compte qu'ils devront réduire leurs taux de transmission. Ensuite, nous étendons notre analyse aux cas de plusieurs leaders. Nous constatons que les deux solutions que nous proposons offrent une meilleure utilisation du canal et promettent des performances plus élevées comparées au standard slotted aloha (meilleur débit et délai inférieur). Nous aussi montrons que la meilleure partition des rôles leader/followers étant de définir un seul leader et plusieurs followers.

Apprentissage de l'équilibre de Nash dans les canaux à collision non saturés : Le troisième chapitre considère les techniques d'accès aléatoire de type slotted aloha ou CSMA, et traite la région des équilibres de Nash en considérant le cas général d'usagers non saturés. Chaque usager dispose d'une file d'attente d'une capacité infinie et qui recueille les paquets provenant des couches hautes. Une caractérisation de la région de stabilité est aussi présentée. Ici, chaque utilisateur i demande une qualité de service (QdS) à satisfaire en terme d'un débit fixe r_i. Ensuite, nous considérons un système distribué où chaque utilisateur décide de sa probabilité de transmission en vue de satisfaire sa QdS. Le concept de solution le mieux adapté pour ce genre de situation conflictuelle est incontestablement l'équilibre de Nash avec contraintes. En effet, contrairement au cas d'utilisateurs saturés largement étudié (c-à-d ayant toujours des paquets à transmettre) dans la littérature et où l'existence de deux équilibres a été démontré [103], nous prouvons l'existence d'un continuum d'équilibres de Nash dans le cas non saturé. Ici, les utilisateurs peuvent être dans deux états différents (ayant un paquet à transmettre ou étant libre). Un résultat très important étant la coïncidence de la région de stabilité avec celle des équilibres de Nash. En outre nous prouvons l'existence d'un équilibre énergiquement efficace (minimise la consommation d'énergie totale) pour tous les utilisateurs. Ce dernier résultat nous a motivé à décrire deux algorithmes d'apprentissage pour converger d'une manière distribuée vers le meilleur point de fonctionnement. Etant basé sur la meilleure réponse, le premier algorithme est semi-distribué. En effet, chaque utilisateur a besoin d'estimer la probabilité que le canal ne soit pas occupé par ses concurrents. Par contre, le deuxième est complètement distribué et chaque utilisateur a besoin seulement de recevoir le feed-back accusant réception de la transmission, si elle avait lieu, au cours de l'itération précédente. Evidemment, la demande et débit, la probabilité de transmission et taux de saturation sont des données intrinsèques connues par l'usager en question. Quand la durée d'observation devient conséquente, même l'algorithme basé sur la meilleure réponse devient complètement distribué. Une étude analytique et de nombreuses simulations valident nos résultats et démontrent la convergence et re-convergence, après perturbation, de nos algorithmes vers l'équilibre efficace. Dans le cas où aucun équilibre efficace n'existe, nos algorithmes mènent à des situations conflictuelles. Les mobiles deviennent agressifs et transmettent avec probabilité 1, et aucun d'eux ne satisfait sa propre demande en débit. Cette situation ne pourrait avoir de solution même en utilisant un mécanisme de type multiplexage temporel (TDMA) ou en présence d'une entité centralisée qui ordonnance les transmission. Ceci étant simplement dû à l'insuffisance des ressources.

Evaluation de performances des réseaux multi-sauts hétérogènes: Extension locale de la couverture d'un réseau cellulaire WiMAX

Nous consacrons la deuxième partie du manuscrit à l'évaluation de performances d'une famille de réseaux hétérogènes composés d'un réseau WiMAX interconnecté à un réseau multi-sauts ad hoc en vue d'étendre/améliorer la couverture. Nous avons basé notre étude sur un modèle analytique qui prend en compte la topologie, le routage, l'accès aléatoire dans la couche MAC et une probabilité de transfert (forwarding probability). Nous distinguons divers aspects clés qui font que notre contribution dans cette deuxième partie de thèse est aussi nouvelle s'ils sont considérés ensemble. Le premier point fort étant l'ordonnancement des transmissions de paquets au niveau de la couche réseau. En utilisant un ordonnancement WFQ (Weighted Fair Queueing), nous avons principalement étudié l'impact de la coopération et la région de stabilité dans le réseau. Ensuite, la considération d'une topologie générale, d'un trafic et des paramètres intrinsèques propres à chaque nœud, nous a permis d'élaborer un cadre théorique puissant et général contrairement à l'hypothèse de la symétrie communément utilisée dans la littérature.

Modélisation et évaluation de performance de bout-en-bout dans les réseaux hétérogènes WiMAX/ad hoc : Une des contributions majeures qu'apporte le chapitre 4 est le développement d'un modèle mathématique mettant en évidence deux systèmes interconnectés entre eux. D'après notre étude bibliographique, la littérature relative à l'étude et l'analyse des systèmes hétérogènes est souvent limitée aux simulations et aux expérimentations. D'où la motivation de développer un cadre théorique pour l'analyse de tels systèmes. Au niveau de la couche réseau, chaque nœud dispose de deux files d'attente; La première stocke ses propres paquets (provenant des couches hautes) et la deuxième est réservée au stockage temporaire des paquets provenant des nœuds voisins et qui doivent être transférés via un mécanisme de multi-sauts jusqu'à la destination finale. Nous étudions d'abord la stabilité des files de transfert des nœuds intermédiaires. Puis, nous définissons la région de stabilité du système comme étant l'ensemble de valeur de la probabilité de transfert garantissant la stabilité de toutes les files de transfert. Ensuite nous calculons le débit de bout-en-bout et terminons ce chapitre par le calcul du délai de bout-en-bout moyen. Ce modèle peut aussi être très utile pour étudier le multihoming (connexion à plusieurs réseaux simultanément). Ainsi les utilisateurs se trouvant dans la région couverte par les deux systèmes, pourraient décider de transmettre une fraction α de leurs trafics vers le réseau WiMAX alors qu'ils transmettent $1-\alpha$ via le réseau multi-sauts ad hoc. Cette solution peut en effet réduire la charge et alléger le réseau, réduire le problème d'interférence et aussi offrir une plus large gamme de services aux usagers. Des simulations et exemples numériques montrent comment ceci permettrait de mieux allouer les ressources en fonction de la qualité perçue du canal. Ils montrent aussi l'existence d'une fraction optimale α^* de trafic à transmettre sur chacun des sous-réseaux.

Distribution du délai de bout-en-bout dans les réseaux ad hoc multi-sauts : Le

chapitre 5 est une extension de l'article [49] et qui peut facilement être généralisé au cas des réseaux hétérogène traité dans le chapitre précédent. Il est dédié au calcul de la distribution de délai dans un réseau multi-sauts ad hoc. Contrairement au chapitre 4 où nous sommes contentés du calcul du délai moyen qui pourrait s'avérer parfois imprécis et insuffisant pour donner une vision fiable des performances du réseaux en terme de latence, délai ou de gigue. Nous détaillons ici un calcul plus rigoureux pour trouver une forme explicite du délai de bout-en-bout. En effet, à l'aide de l'approche des fonctions génératrices de probabilité (Probability Generating Funcion), nous obtenons une formule simple pour le temps de service et pour le temps d'attente dans les files d'attente de transfert. Ceci permet de calculer le délai de bout-en-bout ainsi que sa distribution. Comme application, notre résultat est directement exploité pour définir un contrôle d'admission de paquets en se basant sur le délai de bout-en-bout. En d'autres mots un paquet qui arrive après le dépassement d'un délai fixé est automatiquement détruit. Ce mécanisme permet de supporter le trafic temps réel et les services de streaming (Vidéo ou télévision à la demande, etc.). Ensuite et dans le but d'augmenter le débit utile de bout-en-bout, nous avons proposé un algorithme simple rendant dynamique le nombre de retransmission par paquet. L'idée est simple, l'algorithme favorise les paquets qui sont proches de leurs destination finale en les attribuant un nombre de retransmission de plus en plus grand. Plusieurs résultats numériques et simulations prouvent la précision de notre approche et fournissent des politiques et heuristiques pour un choix optimal des différents paramètres.

Analyse de performances du standard IEEE 802.11e dans les réseaux ad hoc multi-sauts asymétriques avec ordonnancement WFQ

Le développement du modèle mathématique de la partie précédente, nous a permis de concevoir un modèle analytique assez complet pour représenter/évaluer les réseaux ad hoc multi-sauts fonctionnant avec la norme IEEE 802.11e DCF/EDCF. En effet, nous réexaminons dans la troisième partie les réseaux ad hoc en considérant une architecture protocolaire communicante de type inter-couche (cross-layer). Pour optimiser la transmission de bout-en-bout, il a été démontré que l'optimisation d'une couche séparée sans considération des autres couches n'améliore pas les performances, la robustesse ni l'efficacité du réseau.

Modélisation inter-couche des réseaux IEEE 802.11e en mode ad hoc multi-sauts avec ordonnancement WFQ : Le chapitre 6 considère l'interaction entre les couches Application, Réseau, MAC et Physique. Ainsi, par exemple, les paramètres des couches Application, Réseaux et MAC peuvent être reconfigurés selon l'information sur l'état du canal provenant de la couche Physique. Notre modèle prend en compte aussi le problème des nœuds cachés et celui des nœuds exposés. Le taux d'accès au canal et la probabilité de collision sont maintenant exprimés en fonction de l'intensité du trafic, de la topologie, du routage et des interférences. Par ailleurs le nombre limite de retransmissions dans la couche MAC dédié pour chaque connexion est fini. D'une part, ce dernier est responsable de l'asymétrie du taux de service des files d'attentes et de

la distribution générale du temps de service. D'autre part, il a un impact direct sur les performances d'un réseau chargé, et qui s'avère être une bonne solution pour établir l'équité dans le réseau. Nous arrivons à deux systèmes d'équations non linéaires mettant en dépendance des paramètres de la couche Réseau et MAC. Ensuite, nous avons décris un algorithme simple qui permet de résoudre conjointement les deux systèmes. Des résultats numériques et des simulations sont ensuite présentés pour assister et confirmer nos résultats. L'impact des paramètres des couches Réseau, MAC, et Physique a été discuté et décrypté pour montrer l'utilité du modèle inter-couche.

Utilisation des codes Fountain pour améliorer le débit et l'équité dans les réseaux ad hoc : Enfin, pour établir l'équité entre utilisateurs/connexions, nous détaillons dans le chapitre 7 une solution simple et qui ne sollicite aucune modification du standard IEEE 802.11e DCF/EDCF. L'idée de base est, au niveau de la couche MAC, de coder les paquets à l'aide d'un codage Fontaine (Fountain code). A notre connaissance, nous sommes les premiers à avoir proposé l'utilisation du codage Fontaine dans la couche MAC. En effet, notre étude bibliographique nous a montré que le codage Fontaine est souvent employé pour les diffusions (3GPP MBMS, DVB,etc.). Le codage de l'information étant au niveau de la couche Application ou Transport. Ici, avec un mécanisme de décodage incrémentale de type H-ARQ (Hybrid Automatic Repeat reQuest), les différentes copies d'un même paquet sont combinées pour assurer une plus grande probabilité de succès de décodage. Une des propriétés fabuleuses du codage Fontaine étant la nécessité d'un faible taux de redondance. Cette solution nous a permis donc de réduire le nombre moyen de retransmissions par paquets et donc établirait plus d'équité en terme de débit, en particulier pour les connexions qui souffrent d'une probabilité de collision élevée.

Contents

List of Figures

List of Tables

Introduction

General introduction

Approximately every five or ten years we are blessed with the implementation of a new technology that can have a major bearing upon how we live and work. Figure 1 presents a non-exhaustive list of the most implemented cellular telecommunications systems. The development of cellular networks can be classified in different classes (grouped together by technology), each class is known as a generation:

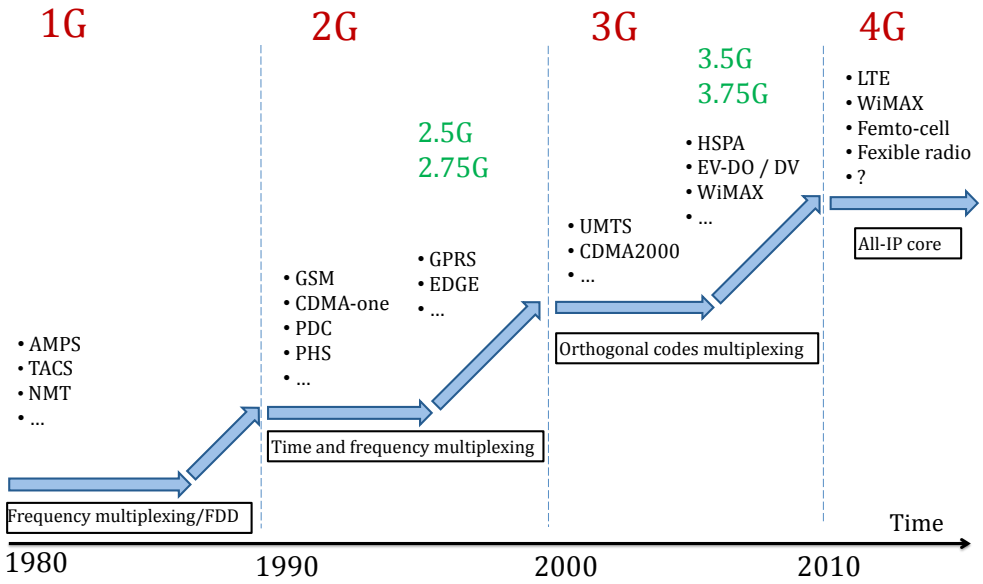


Figure 1: *Evolution of Telecommunication systems and rise of network heterogeneity.*

1G : Voice communication through analog FM transmission. Frequency Division Multiple Access/Frequency Division Duplexing (FDMA/FDD) was used to accommodate multiple users and channels (1980's). Advanced Mobile Phone System (AMPS), Total Access Communication System (TACS) and Nordic Mobile Telephony (NMT) were the most used 1G networks.

2G : Were purely digital networks using digital modulation and Time Division Multiple Access (TDMA) or Code Division Multiple Access (CDMA). This generation supports voice, text messaging (SMS) and circuit-switched data communication (1990's). The two most implemented examples of 2G networks are Global System for Mobile communication (GSM) and Interim Standard 95 (IS-95). Forward error correction (FEC) and encryption are also supported by this generation. Some additional standards were developed to increase the data speeds of 2G networks, including 2.5G General Packet Radio Service (GPRS) and 2.75G Enhanced Data Rates for GSM Evolution (EDGE).

3G : Provides data rates up to 2 Mbps using wideband modulation techniques with increased user capacity (2000's). Services like Internet, e-mail, multimedia stream-

ing, video telephony and instant messaging are all services available to 3G devices (this includes mobile handsets and computers). The two main 3G standards are Wideband Code Division Multiple Access (WCDMA) and Multi Carrier CDMA (MC-CDMA). Improvements on the 3G standard include 3.5G High Speed Downlink Packet Access (HSDPA) and 3.75G High Speed Uplink Packet Access (HSUPA).

4G : Higher data rates (up to 1 Gbps) will be supported through the use of an all-IP futuristic network (2010). The multiple access technology used by this generation will be OFDM (Orthogonal Frequency Division Multiplexing), e.g., LTE and WiMAX. The application layer services like mobile video will be provided. There will be seamless integration between all generation networks, which gives rise to interoperable heterogeneous wireless networking.

5G : A recent discussion and research trend on this topic, let imagine a software highly improvement of 4G networks. This generation will mainly base on how a resource should be used and in particular take benefit from the newly concept of virtual Multiple In Multiple Out (MIMO). The idea here is to consider all devices around as virtual antennas in order to improve the instantaneous throughput. Flexible radio called also Software Defined Networks is the most known 5G candidate.

The aim behind this successive generations is to facilitate our productivity, and even enhances our recreational capability. In the past we witnessed the PC revolution, the advent of the PDA, and the growth in the use of wireless LANs. Moreover, with the rapid growth in the number of wireless applications, services and devices, using a single wireless technology such as a second generation (2G) and third generation (3G) wireless system would not be efficient to deliver high speed data rate and Quality-of-Service (QoS) support to mobile users in a seamless way, see [75]. The next generation wireless systems (also sometimes referred to as Fourth generation (4G) systems) are being devised with the vision of heterogeneity in which a mobile user/device will be able to connect to multiple wireless networks (e.g., WLAN, cellular, WMAN, Mesh) simultaneously. For example, IP-based wireless broadband technology such as IEEE 802.16/WiMAX (i.e., 802.16a, 802.16d, 802.16e, 802.16g) and 802.20/MobileFi will be integrated with 3G mobile networks, 802.11-based WLANs, 802.15-based WPANs, and wireline networks to provide seamless broadband connectivity to mobile users in a transparent fashion.

Rather than being an inconvenience, these new heterogeneous networks must indeed be regarded as a new challenge/solution to offer to the users an efficient and ubiquitous radio access, by means of a coordinate use of the available Radio Access Technologies (RATs). In this way, not only the user can be served through the RAT that fits better to the terminal capabilities and service requirements, leading to the "connected everywhere, anytime, anyhow" experience, but also a more efficient use of the available radio resources can be achieved from the operator's point of view. Now, the heterogeneous network becomes transparent to the final user and the so-called Always Best Connected (ABC) paradigm [56, 75, 116], which claims for the connection to the

RAT that offers the most efficient radio access at each instant, can be achieved.

Heterogeneous wireless systems will achieve efficient wireless resource utilization, seamless handoff, global mobility with QoS support through load balancing and tight integration with services and applications in the higher layers. After all, in such a heterogeneous wireless access network, a mobile user should be able to connect to the Internet in a seamless manner. The wireless resources need to be managed efficiently from the service providers point of view for maximum capacity and improved return on investment. The convergence of the different wireless technologies is a contiguous process. With each passing day, the maturity level of the mobile user and the complexity level of the cellular network reach a new limit. Current networks are no longer "traditional" GSM networks, but a complex mixture of 2G, 2.5G, and 3G technologies. Furthermore, new technologies beyond 3G (e.g., HSPA, WiMAX, LTE) are being utilized in these cellular networks. Existence of all these technologies in one cellular network has brought the work of design and optimization of the networks to be viewed from a different perspective. We no longer need to plan GSM, GPRS, or WCDMA networks individually. The cellular network business is actually about dimensioning for new and advanced technologies, planning and optimizing 4G networks, while upgrading 3G/3.5G networks.

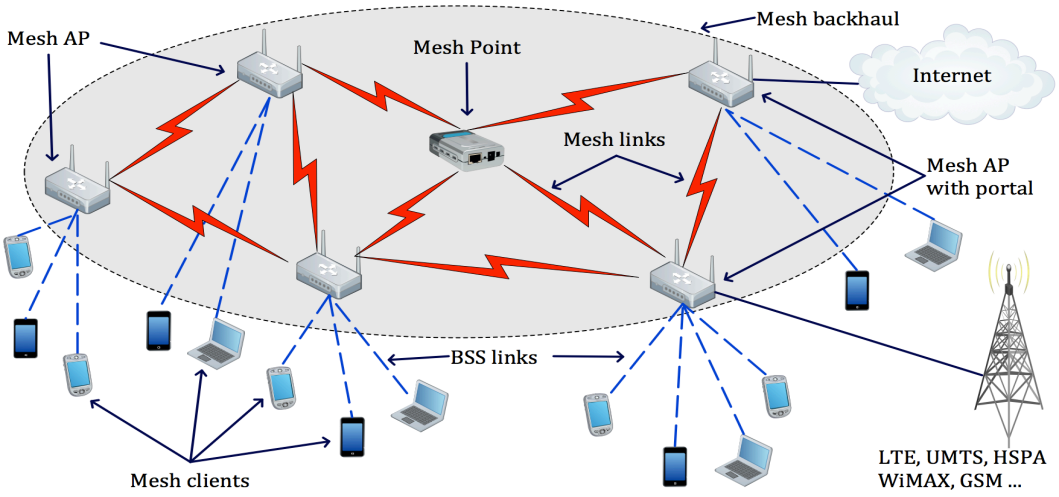


Figure 2: *Illustrative example of a heterogeneous Wireless Mesh Network.*

Protocol engineering and architecture design for broadband heterogeneous wireless access systems is an emerging research area. Load balancing and network selection, resource allocation and admission control, fast and efficient vertical handoff mechanisms, and provisioning of QoS on an end-to-end basis are some of the major research issues related to the development of heterogeneous wireless access networks. This thesis covers different aspects of analysis, design, deployment, and optimization of protocols and architectures for heterogeneous wireless access networks. In particular, the

topics include challenges and issues in distributed algorithms design and provisioning of a QoS framework for heterogeneous wireless access networks, architectures and protocols for optimal spectrum utilization in multihop ad hoc wireless networks, network selection in heterogeneous wireless access networks, modeling and performance analysis of heterogeneous mobile networks, quality-oriented multimedia streaming in heterogeneous/multihop ad hoc wireless networks, and extensive simulations.

Motivation and general overview

In the near future, multitude of wireless communication network based on a variety of radio access technologies and standards will emerge and coexist. The availability of multiple access alternatives offers the capability of increasing the overall transmission capacity, providing better service quality, dealing with health problems of wireless systems and reducing the deployment costs for wireless access. This way, practically all existing technologies will become simple RATs (e.g., GSM, UMTS, CDMA2000, HSPA, WLANs, WiMAX, LTE, etc.) to access the real network which is henceforth mandatory heterogeneous. In order to exploit this potential multiaccess gain, it is required that different RATs are managed in a cooperative fashion. In the design of such a cooperative network, the main challenge will be bridging between different networks technologies and hiding the network complexity and difference from both application developers and subscribers and provide the user seamless and QoS guaranteed services. The trend will also bring about a revolution in almost all fields of wireless communications, such as network architecture, protocol model, radio resource management, and user terminal as well. As wireless networks grew larger, it became evident that centralized control would be impractical for coordinating all elements of the network, and in particular end-user transmissions. The celebrated Aloha protocol was designed at the early 70's [2] as a distributed mechanism which can allow efficient media sharing. This protocol and its variants, such as slotted aloha [125] CSMA-CD and tree-algorithms [25], are cooperative in the sense that each user is committed to perform his part of the protocol.

Clearly, modern wireless network protocols are often based on Aloha-related concepts (for example, the 802.11 and 802.16 standards [36, 146, 147, 148]). The design of such protocols raises novel challenges and difficulties, as the wireless arena becomes more involved. The most studied issue is the under-utilization of aloha medium access method (18% using pure aloha and 37% using slotted aloha). The basic underlying assumption in legacy slotted aloha protocols is that any concurrent transmission of two or more users causes all transmitted packets to be lost [10, 13, 25, 125]. However, this model does not reflect the actual situation in many practical wireless networks where some information can be received correctly from a simultaneous transmission of several packets. Therefore, this assumption has been subjected to some improvements in literature. The first improvement is called the capture effect: The packet with the strongest power level can be received successfully (captured) in the presence of contending transmissions if its power level is sufficiently high. It occurs in networks with single packet reception capability where packets arrive at the common receiver with different power levels due to near-far effect, shadowing or fading. The effect of capture on Aloha [3, 18, 87, 89, 112, 132, 136, 169, 170] and on IEEE 802.11 protocol (Carrier Sense Multiple Access-Collision Avoidance (CSMA/CA)) [69, 100, 114] has been studied extensively in the literature and new MAC protocols for channels with capture have been proposed [37]. Despite of the bounty of works and efforts investigated in analyzing aloha-like protocols with capture effect our approach is different. We propose to consider a random choice of a power level and fine-tune the transmit probability in or-

der to maximize the objective function. Through the first part of this thesis we define a stochastic game where mobiles are assumed to be selfish and develop many distributed schemes with several power level setups. Moreover, the researchers community usually focuses on saturated users, i.e., users have all the time packets to be transmitted. Here, we analyze the non saturated case with an infinite buffer capacity. In contrast to saturated users with a fixed throughput demand, two equilibria may exist [104]. This issue is then to be addressed when relaxing the saturation assumption.

Hajek et al. [70] provided asymptotic results on the capture probability in the limit of infinite number of users, see also [34] and references therein for detailed overview. The other major improvement to the original protocol, known as multipacket reception capability, assumes that a subset of the collided packets can be received successfully. The impact of the multipacket reception capability on MAC protocols has received limited attention to date. Ghez et al. [63, 64] proposed a channel model for networks with multipacket reception capability and studied stability properties of slotted aloha in such a setting. Tong et al. have proposed MAC protocols based on multipacket reception capability [106, 107, 167, 168] using the channel model suggested in [63]. The protocols developed in [167, 168] maximize the normalized throughput by controlling the set of users who are allowed to transmit in each slot. However, these protocols require a centralized controller and hence are impractical for large distributed networks. [122] presents a pseudo PHY/MAC cross-layered approach for multipacket reception.

Despite its advantages, the cellular concept (or infrastructure-based wireless networks, to be more general), has its drawbacks: it is of relatively low bandwidth, similar in many ways to wired dial-up access, and it generally takes time and potentially high cost to set up the necessary infrastructure [65]. Moreover, even if costs were reduced and efficiency improved, infrastructure-based networks may not always be possible, scalable, appropriate, or even desirable in many of the envisaged scenarios for ubiquitous and pervasive computing/communications. To fulfill the vision of ubiquitous computing/communications, wireless networks need to evolve beyond the current infrastructure-type of networks. Fortunately, because of significant advances in hardware technology in the past decades - most notably in the areas of processing capability and storage capacity- it has now been possible to include more "intelligence" into smaller devices with significant reductions in power consumption and higher performances. As a consequence, the deployment of wireless networks without any preexistent infrastructure (the so-called wireless ad hoc networks) are now becoming possible. Although derived from military research into mobile networks[1], the emergence of multihop ad hoc and Mobile ad hoc networking (MANET, [66, 67]) has its greatest potential in the commercial marketplace and is the one of key points of this dissertation. Each node operates not only as a host but also as a router, forwarding packets on behalf of other nodes that may not be within direct wireless transmission range of their desti-

[1]Initial interest in ad hoc networks emerged within the military arena with the Packet Radio Network project funded by the Defense Advanced Research Projects Agency in 1972, followed by the Survivable Radio Networks project in 1983, and DARPA's Global Mobile Information Systems program in 1994 [58].

nations. In recent years, the research in ad hoc networks becomes more and more dedicated to specialized network applications like Wireless Mesh Networks (WMN) [5, 76], opportunistic networks, Wireless Sensor Networks (WSN) and vehicular ad hoc networks (VANET). This kind of ubiquitous networks are dynamically self-organized and self-configured, with the nodes in the network automatically establishing and maintaining mesh connectivity among themselves (creating, in effect, an ad hoc network). This feature brings many advantages such as low up-front costs, easy network maintenance, robustness, and reliable service coverage.

However, the main contribution of "general" ad hoc networks research consists on the understanding of the constraints and limitations of such wireless networks [39, 50]. The multi-hop nature of such networks and the broadcast nature of wireless channel are responsible of many challenges and hard issues. Since the current design of ad hoc networks is based on the standard OSI layered approach, significant research was occupied to propose new protocols at different layers independently to combat the network limitations. Despite the flexibility and the simplicity of layered model, it leads to poor performance and thus a cross-layer protocol design is needed. In a cross-layer design an information in a given layer is made available to protocols in different layers which enable resource optimization and therefore promises more efficient performance improvement. Moreover, carrier sensing is widely adopted in wireless communication to protect data transfers from possible collisions. e.g., distributed coordination function (DCF) in IEEE 802.11 standard renders a node to defer its communication if it senses the medium busy. However, even if the carrier signal is detected to be greater than the threshold, both ongoing and a new communication can be simultaneously successful depending on their relative positions in the network or equivalently, their mutual interference level. Therefore, supporting multiple concurrent communications is important in multihop ad hoc networks in order to maximize the network performance and improve the reuse factor of the common channel. However, it is largely ignored in DCF of the 802.11 standards because it is primarily targeted at single-hop wireless LANs [144]. This latter point justifies the need to adapt/design dedicated protocols for multihop communications instead of reusing infrastructure-based standards. In [21], the authors achieved a high throughput and low delay in ad hoc networks. El Gamal et al. [59] analyzed the optimal delay-throughput scaling for different wireless network topologies. In the static random network with n nodes, they obtained an optimal tradeoff between throughput and delay. A plenty of other trials concerning the stability, capacity and delay in ad hoc networks were performed, see e.g. [62, 91, 108, 162].

Our contributions

This dissertation focuses on the developing new version of slotted aloha like random access protocols and ad hoc networks performance using multi-layer parameters and forwarding capability. The performance metrics of interest are throughput, delay, stability of forwarding queues and energy consumption. The chapters are organized in three parts. The first part deals with the under utilization problem of wireless collision channels. In the second part, we based our study on performance evaluation of a WiMAX/ad hoc integrated network. We developed an analytical model using a new cross-layer approach with a weighted fair queueing scheduler. We are then able to differentiate between packets to be forwarded and new entering packets. Tracking traffic is now easy and very comprehensive. The last part is dedicated to extending the previous model to IEEE 802.11e DCF/EDCF. We indeed integrate APPLICATION, NETWORK, MAC and PHY layers in a unified cross-layered model. This way, the modeling framework focuses on the interactions between several layers, and on the impact that each node has on the dynamics of every other node in the network. A key feature of our model is that nodes can be modeled individually, i.e., it allows a per-node setup of many layer-specific parameters. Moreover, no spatial probability distribution or special arrangement of nodes is assumed; the model allows the computation of individual (per-node/per-path) performance metrics for any given network topology and radio channel model. This later feature gives very important insight to judiciously set each parameter taking into account the information coming from other layers. Our main contributions and the outline of the chapters content are the following

Chapter 1 : Slotted Aloha with Random Power Level Selection and Capture Effect. We consider the uplink case of a cellular system where m bufferless mobiles transmit over a common channel to a base station, using the slotted aloha medium access protocol. The novelty here is the consideration of several power differentiation schemes. Indeed, we consider a random set of selectable transmission powers and further study the impact of priorities given either to new arrival packets or to the backlogged ones. Later, we address a general capture model where a mobile transmits successfully a packet if its instantaneous SINR (signal to interferences plus noise ratio) is lager than some fixed threshold. Under this capture model, we analyze both the cooperative team in which a common goal is jointly optimized as well as the noncooperative game problem where mobiles seek to optimize their own objectives. Furthermore, we derive the throughput and the expected delay and use them as the objectives to optimize and provide a stability analysis as an alternative study. Exhaustive performance evaluations are performed, we showed that schemes with power differentiation improve significantly the individual and the global performances.They also could eliminate in some cases the bi-stable nature of slotted aloha.

Chapter 2 : Sustaining Partial Cooperation in Hierarchical Wireless Collision Channels. In a spirit similar to that of Chapter 1, we consider a wireless system composed of one central receiver and several selfish transmitters communicating via the

slotted aloha protocol. The set of users is split into two classes: Leaders and Followers. Then, we study the induced non-cooperative hierarchical game based on the Stackelberg equilibrium concept. Each user seeks to maximize his own throughput or to minimize the expected delay experienced by its backlogged packets. Those utility functions are clearly depending on the individual transmission probability and the transmission probabilities of all other concurrent users in the network. Using a 4D Markovian model, we compute the steady state of the system and derive the average throughput and the expected delay as well. We start by discussing the protocol design and propose a controlled slotted aloha using a virtual controller that make the channel lossy to reduce the channel access of concurrent users. Later, we investigate the impact of introducing hierarchy in such a random access protocol and discuss how to distribute leader/follower roles. Furthermore, exhaustive performance evaluations are carried out, we show that the global performance of the system is improved compared to standard slotted aloha system. However, a slight performances slow-down may be observed for the followers group small number of users.

Chapter 3 : Learning Constrained Nash Equilibrium in Wireless Collision Channels. We consider a finite number of users, with infinite buffer storage, sharing a single channel using the aloha medium access protocol. This is an interesting example of a non saturated collision channel. We investigate the uplink case of a cellular system where each user will select a desired throughput. The users then participate in a non cooperative game wherein they adjust their transmit rate to attain their desired throughput. We show that this game, in contrast to the saturated case where two equilibria may exist, either has no Nash Equilibrium or has infinitely many Nash Equilibria. Further, we show that the region of NE coincides with an appropriate 'stability region'. We also discuss the efficiency of the equilibria in term of energy consumption and congestion rate. Next, we propose two learning algorithms using a stochastic iterative procedure that converges to the best Nash equilibrium. For instance, the first one needs partial information (transmit rates of other users during the last slot) which can be estimated by observing enough the system behavior. The second is an information-less and fully distributed scheme. We approximate the control iterations by an equivalent ordinary differential equation in order to prove that the proposed stochastic learning algorithm converges to the desired Nash equilibrium even in the absence of any coordination or extra information. Extensive numerical examples and simulations are provided to validate our results.

Chapter 4 : Performance Evaluation of WiMAX and Ad hoc Integrated Networks. Current mobile users are often equipped with several network interfaces, which may be of different access technologies. Each access technology has specific characteristics in terms of coverage area and technical characteristics (bandwidth, QoS, etc.) and provides diverse commercial opportunities for the operators. It seems likely that these various technologies have to coexist and, from then, solutions of integration and interoperability will be necessary to deal with the technological diversity. We consider in this chapter the WiMAX and ad hoc integrated networks. A user who needs to estab-

lish a high reliable service prefers to use WiMAX network and another who needs to transmit Best Effort traffic with high rate but delay tolerant prefers to connect ad-hoc network that is assumed to have access to Internet. We aim through this combination to locally extend the coverage of the WiMAX cell and to study the stability of involved nodes, in particular, gateway nodes. Under stability condition, our main result is characterization of the end-to-end throughput and delay using the rate balance equation and a $G/G/1$ queueing model. Through numerical results, we demonstrate the utility and efficiency of our approach.

Chapter 5 : Asymptotic Delay in Wireless Ad hoc Networks with Asymmetric Users. In this chapter, we present an analytical model for an approximate calculation of the end-to-end delay performance in multihop wireless ad hoc networks. In contrast to literature that largely focuses on average delay, our work focuses on the distribution of end-to-end delay. Now, we assume that each source injects packets in the network, which traverse intermediate relay nodes until they reach the final destination. Firstly, we employ discrete-time queueing theory to derive the expressions for the queue length and the delay in terms of probability generating functions. Secondly, in order to improve the control routing and transmission scheduling, we adopt a new architecture that allows information sharing across different layers for efficient utilization of network resources, and meeting the end-to-end performance requirements of demanding applications. Thirdly, we propose a cross-layered packet admission control scheme based on delay timeout mechanism. This guarantees quality of service for multimedia applications such as voice and video streaming. Finally, we conduct extensive simulations in order to verify and assist our analytical results.

Chapter 6 : A Cross-layered Modeling of IEEE 802.11-Operated Ad hoc Networks. Performance of IEEE 802.11 in multi-hop wireless networks depends on the characteristics of the protocol itself, and on those of the upper layer routing protocol. We are interested here in modeling the IEEE 802.11e enhanced distributed coordination function (EDCF) networks. We investigate the intricate interactions between several PHY, MAC and Network layer parameters, including the carrier sense threshold, the contention window size, limit number of retransmissions, multi rates, routing protocols and the network topology. In fact, we are focusing to study the effect of cooperation and PHY/MAC parameters on the stability and the throughput of ad hoc networks. We extend the results of [166] to a multi-hop ad hoc network with asymmetric topology and asymmetric traffic loads using a cross-layer architecture. We develop an analytical model that predicts the throughput of each connection as well as the stability of forwarding queues of intermediate nodes. Performance of such a system is also evaluated via simulations. To the best of our knowledge, our work is the first to consider asymmetric topology of the network and asymmetric parameters in PHY/MAC layers. We show that the performance measures of MAC layer are affected by the intensity of traffic of a connection that an intermediate node forwards. More precisely, the attempt rate and the collision probability are now dependent on the traffic flows, topology and routing.

Chapter 7 : Fountain Code-based Fair IEEE 802.11-Operated Ad hoc Networks.

Our objective in this chapter is to study a class of transmission mechanisms that include coding in order to improve the probability of successful delivery within a multihop ad hoc network. We propose an analytical approach that allows to quantify tradeoffs between throughput and fairness. We study the effect of coding on the performance of the network while optimizing parameters that govern NETWORK, MAC and PHY layers. In other words, this chapter extends the multi-layered model discussed in the previous chapter. We present a solution based on coding packets at MAC layer before transmitting them over the common channel. To the best of our knowledge, we are the first to use Fountain codes at MAC layer level. Indeed, coding packets is usually studied at APPLICATION [51, 53, 95, 140] or TRANSPORT layers, see e.g., [86]. In standard IEEE 802.11, whenever a collision occurs the whole packet is lost and needs to be retransmitted after the random backlog period. Our solution uses an incremental combination of all previous copies of a backlogged packet to improve the decoding probability and therefore to correctly recover the original packet. As we expected, using Fountain code in our context increase the Jain's Fairness index over the network. It indeed improves the throughput on paths that suffer from high collision probability. However unfortunately, it may decrease the throughput on the other paths. This is due to the injected redundancy in coded packets and its negative effect that reduces the channel idle period seen by concurrent users.

Part I

On Improving Wireless Collision Channels Utilization

Chapter 1

Slotted Aloha with Random Power Level Selection and Capture Effect

Contents

1.1 Introduction

Aloha [2] and slotted aloha [125] have long been used as randomly distributed medium access protocols for radio channels. They are used in satellite networks and cellular telephone networks for the sporadic transfer of data packets. In these protocols, packets are transmitted sporadically by various users. If packets are sent simultaneously by more than one user then they collide. After a packet is transmitted, the transmitter receives the information on whether there has been a collision (and retransmission is needed) or whether it has been well received. All colliding packets are assumed to be corrupted which get backlogged and are retransmitted after some random time. We focus on the slotted aloha [25], in which time is divided into units. At each time unit a packet may be transmitted, and at the end of the time interval, the sources get the feedback on whether there was zero, one or more transmissions (collision) during the time slot. A packet that arrives at a source is immediately transmitted. Packets that are involved in a collision are backlogged and are scheduled for retransmission after a random time.

Interest has been growing in recent years in studying competition of networking in general, access to a common medium in particular, within the frame of non-cooperative game theory, see e.g. the survey paper [11]. Various game formulations of the standard Slotted Aloha (with a single power) have been derived and studied in [12, 9, 8, 81, 98, 99] for the non-cooperative choice of transmission probabilities. In [9, 12, 13, 44], authors consider slotted Aloha system as both cooperative (where a common objective is jointly optimized) and non cooperative game, where each users fine-tunes its transmit probability to maximize its payoff, with partial information and power diversity. In [81], the authors discuss the equilibrium of a non-cooperative game for Aloha protocols. In their game formulation, Users are heterogeneous and each one fine-tunes its transmit probability in order to guarantee its demand. Mackenzie et al. [99] discusse the stability of slotted Aloha with selfish users behavior and perfect information. The authors showed the existence of an equilibrium and characterized it. In [110] it is shown that the system capacity could be increased from 0.37 to 0.53 if one class of terminals always uses high power and the other always uses low power level. In [87], power diversity is studied with the capture model that we use as well as with another capture model based on signal to noise ratio. [133] studies power diversity under three types of power distribution between the power levels and provides also a detailed stability analysis. [96] proposes a model and evaluates the throughput that can be achieved in a system of N mobiles using generalized aloha like protocols where the mobiles transmit data using a two-state decision system. For cooperative systems, it gives the throughput bounds and explores the tradeoff between throughput and short-term fairness. But our proposal is different here, we address the effect of randomization in power levels for both cooperative and non-cooperative setups.

When multiple users share a common channel and contend for access, a typical conflict problem arises. Recently, the selfish behavior of users in MAC protocols has been widely analyzed using game theory with all its powerful solution concepts. It was shown in [8, 13, 44] and [115] that the users selfish behavior likely leads to a network collapse, where a typical prisoners dilemma situation occurs. This illustrates, in fact, that Nash equilibrium[1] (NE) is not efficient in some situations. This way, full system utilization requires coordination among users using explicit message exchanges or presence of an arbitration mechanism [10], which may be impractical given the distributed nature and arbitrary topology change (due to mobility, ended calls, new calls, environment, ...) of wireless networks. To achieve a better performance without coordination schemes, users need to sustain cooperation or priority. It is beneficial to design a set of users whose mission is to provide incentives for other users to behave cooperatively as well as respect the defined priority, this mechanism may limit the aggressiveness level (access to the channel) and resolve the contention/random problem. Another way to reduce the access contention is to introduce a transmission cost. Indeed, it was shown in [101] that costs have a stabilizing effect; being rational, users will defer packet transmissions when congestion develops and the cost for successful

[1] A Nash equilibrium of a game is a strategy profile such that no player can improve its reward by changing its strategy, assuming the complementary strategies of the other players stay the same. The strategies here are said to be "mutual best responses".

transmission becomes high. This way, users will drop packets when the total transmission costs are high which can cause a huge delay and then the analysis seems to be not applicable to delay-sensitive traffic. Authors in [128], showed that introducing users hierarchy improve the performances of the whole system and the enhancement depends on the offered load. At low load, this improvement is due to the fact that users (both of leaders and followers) retransmit with high probability at low load, hence the backoff time is reduced. At average and high loads, leaders turn to be generally less aggressive than Nash equilibrium case. This way, the collision probability is reduced and then the performance of the hierarchical scheme are enhanced. A pretty phenomena, when the number of followers becomes larger than the number of leaders, is that those latter users become more friendly and reduce their retransmission probability, whereas the followers become very aggressive and transmit at probability close to 1. Here, we extend the model first proposed in [8] and introduce new schemes in which multiple power levels are used for transmission. When several packets are sent simultaneously, one of them can often be successfully received due to the power capture effect. In this chapter, we consider a general capture model where a mobile transmits successfully a packet if his instantaneous SINR (signal to interferences plus noise ratio) is lager than the fixed threshold. In particular, we introduce the differentiation between new packets and backlogged packets allowing prioritization of one or the other in terms of transmit power. We study and compare the following schemes:

1. The first scheme considers the power diversity but defines no prioritization in transmission or retransmission. This scheme will mainly show the effect of power diversity on system performances;
2. In the second scheme, each new packet is immediately transmitted with the lowest power, whereas backlogged packets are sporadically retransmitted at a random power selected among $N-1$ larger distinct levels;
3. Here, new packets are transmitted with the highest power, and backlogged packets are retransmitted at a random power level picked from $N-1$ lower distinct levels;
4. This scheme Gives more priority to new arrivals. Indeed,backlogged packets are retransmitted with the lowest power level and a new packet is transmitted at a random power selected among $N-1$ larger distinct levels.
5. All previous schemes are compared with standard slotted aloha taken as reference, this allows to compare and analyze performances of each scheme.

The capture model used in [8] is not realistic, authors assume therein that when a unique mobile chooses the highest power, compared to other mobiles, its transmission is succeeded independently of the other mobiles and their respective choices. This assumption could not be always true. Indeed, the aggregate signal of other mobiles may jam the signal of the tagged mobile, i.e., whose power level is the highest, therefore no successful transmission exists. Now, a terminal succeeds its transmission if it chooses the most elevated power level compared to other mobiles and its instantaneous SINR

is greater than a given threshold γ_{th}. Under more general capture model, we study the team problem in which we optimize transmission probabilities for the various schemes so as to achieve the maximum throughput or to minimize the expected delay. We discover however that in heavy load, the optimality is obtained at the expense of huge expected delay of backlogged packets (EDBP). We therefore consider the alternative objective of minimizing the EDBP. We study both the throughput as well as the delay performance of the global optimal policy. We also solve the multi-criteria problem in which the objective is a convex combination of the throughput and the EDBP. This allows in particular, to compute the transmission probabilities that maximize the throughput under a constraint on EDBP, which could be quite useful for delay-sensitive applications (e.g., Video streaming, Voice, VoD, ...). We show that schemes with priority do not only improve the average performances considerably but they are also able in some cases to eliminate the bi-stable nature of the slotted aloha. Furthermore, we study the game problem in which each mobile chooses its transmission probability selfishly in order to optimize its own objective. This gives rise to a game theoretic model in which we study the equilibrium properties (Nash equilibrium). We show that the power diversity and the prioritization profit to mobiles also in this competitive scenario even if the advantage is less notorious than in the team's behavior.

1.2 Problem formulation

We consider a collision channel used by one central receiver and m mobiles without buffer, i.e., mobiles do not generate new packet till the current one is successfully transmitted. This assumption can have realistic application in the context of signaling. Indeed, it is natural to assume that a source does not start generating a new signaling packet (e.g., a new reservation) as long as the current signaling packet is not transmitted successfully. Here, the process of attempts to retransmit a new packet from a source after the previous packet has been successfully transmitted corresponds well with our bufferless model.

At the beginning of any slot, each mobile can transmit a packet using a power level among N available levels $\mathcal{P} = \{p_1, p_2, ..., p_N\}$. We consider a general capture model where a packet transmitted by a mobile i is received *successfully when and only when its instantaneous SINR is larger than some given threshold* γ_{th}. Let p_i be the transmit power chosen by mobile i in the current slot, and σ^2 be the spectral density of the background noise which is assumed to be AWGN and time-independent. Let us denote by g_i the gain experienced by mobile i over the channel. It depends on the distance between the mobile and the AP, but it is also impacted by the reflective paths. The instantaneous SINR of mobile i received by the central receiver is given by

$$\gamma_i = \frac{g_i p_i}{\sum_{j=1}^{k} g_j p_j I(j) + \sigma^2}, \tag{1.1}$$

where $I(j)$ is indicator of the event that at the current slot, user j is transmitting. We address here a random power selection fashion where transmit powers are selected according to a probability distribution $X = [x_1, x_2, ..., x_N]$, i.e., power level p_j is chosen with probability x_j. The impact of selection probability distribution will be discussed in Subsection 1.5. In particular, we propose to discuss the case of 1) uniform distribution; 2) high power levels are prioritized and; 3) the case where low power levels are prioritized.

We extend here the Markovian model first described by Altman et al. [8, 9, 13] in order to incorporate the instantaneous capture effect. The arrival probability of the packets to the source i follows a Bernoulli process with parameter q_a (i.e., at each time slot, there is a probability q_a of a new arrival at a source, and all arrivals are independent). As long as there is a packet at a source (i.e., as long as it is not successfully transmitted) new packets to that source are blocked and lost (because we consider sources without buffer). The arrival processes of different sources are assumed to be independent. Similarly, we consider that a backlogged packet at source i is retransmitted with probability q_r^i. We should restrict in our control and game problems to simple policies in which q_r^i does not change in time. We also will be interested in the case of symmetric sources, we should then find a symmetric optimal solution, i.e., a retransmission probability q_r which does not depend on i. Next, we consider as a state of the system the stochastic process representing the number of backlogged packets in the beginning of a slot, we denote it n. For any choice of values $q_r^i \in (0, 1]$, the state process is a Markov chain that contains a single ergodic chain (and possibly transient states as well). Define $\bar{\mathbf{q}}_r$ to be the vector of retransmission probabilities for all users (whose j^{th} entry is q_r^j). We note the transition matrix of the Markov chain by $P(\bar{\mathbf{q}}_r)$. Let $\bar{\pi}(\bar{\mathbf{q}}_r)$ be the corresponding vector of the steady-state probabilities where its n^{th} entry $\pi_n(\bar{\mathbf{q}}_r)$ denotes the probability of n backlogged mobiles. When all entries of $\bar{\mathbf{q}}_r$ are the same, say q, we shall write (with some abuse of notation) $\bar{\pi}(q)$ instead of $\bar{\pi}(\bar{\mathbf{q}}_r)$.

We introduce further notation, assume that there are n backlogged packets, and all use the same value q_r for retransmission rate. We denote by $Q_r(i, n)$ the probability that i mobiles out of the n backlogged packets retransmit at the current slot. Then

$$Q_r(i, n) = \binom{n}{i} (1 - q_r)^{n-i} (q_r)^i. \tag{1.2}$$

Let $Q_a(i, n)$ be the probability that i unbacklogged mobiles transmit packets in a given slot (i.e., that i arrivals occurred at mobiles without backlogged packets). We have

$$Q_a(i, n) = \binom{m-n}{i} (1 - q_a)^{m-n-i} (q_a)^i. \tag{1.3}$$

Let $Q_r(i, 0) = 0$ and $Q_a(i, m) = 0$.

For ease of reading, we summarize the assumptions of our contention model as follows

- A finite set of m bufferless mobiles interacts over a single collision channel.
- Time is divided into multiple equal and synchronized slots. Transmission feedback (success or collision) are received in the end of the current slot.
- Each user i is assumed to be non saturated, packets arrive from higher layers according to a Bernoulli process with parameter q_a.
- Each user i retransmits, in every slot, its packets with probability q_r^i.
- The average throughput or minus expected delay of backlogged packets are the objective functions to maximize.

Remark 1.2.1. *Quite frequently one uses the ALOHA protocol for sporadic transmissions of signaling packets such as packets for making reservation for a dedicated channel for other transmissions (that do not use ALOHA), see e.g. the description of the SPADE on demand transmission protocol for satellite communications in [135]. In the context of signaling, it is natural to assume that a source does not start generating a new signaling packet (e.g. a new reservation) as long as the current signaling packet is not transmitted. In that case, the process of attempts to retransmit a new packet from a source after the previous packet has been successfully transmitted coincides with our no buffer model.*

1.3 The Team Problem

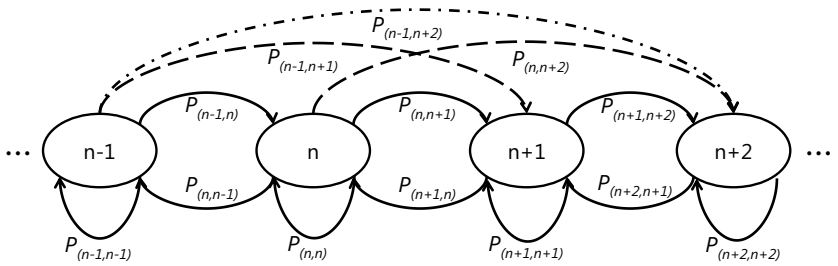


Figure 1.1: *Markov chain for the team problem. The state of the system is the backlog, it can decrease by at most one per transition but can increase by an arbitrary amount less than or equal $m - n$.*

We first study slotted aloha in a team problem point of view, i.e., all mobiles optimize the same objective function (maximize the system throughput or minimize the average delay). Here, we analyze the case when there exists a central entity (base station, dedicated device, ...) that computes the optimal strategy profile and broadcasts it to operating mobiles. This entity should know the total number of mobiles and their individual new arrivals intensity q_a. Next, we allow prioritization by incorporating a selective transmit power allocation mechanism. On the beginning of each slot, each mobile picks a power level from the N available power levels and decides to transmit/retransmit its packet. The corresponding Markov chain is depicted in figure 1.1. Based on power allocation fashion, we analyze the following four schemes.

1.3.1 Scheme 1 : Random power levels without priority

We assume that there is no preference between new packets or backlogged ones. A mobile chooses to transmit using randomly a power level among N available levels. In case all mobiles attempt the channel with the same probability q, the transition probabilities of the system are given by

$$P_{n,n+i} = \begin{cases} Q_a(m-n,n)\sum_{j=0}^{n} Q_r(j,n)(1-A_{j+m-n}),\ i=m-n, i\geq 2 \\ Q_a(i,n)\sum_{j=0}^{n} Q_r(j,n)(1-A_{j+i}) \\ \quad +Q_a(i+1,n)\sum_{j=0}^{n} Q_r(j,n)A_{j+i+1}, \qquad 2\leq i<m-n \\ Q_a(1,n)\sum_{j=1}^{n} Q_r(j,n)(1-A_{j+1}) \\ \quad +Q_a(2,n)\sum_{j=0}^{n} Q_r(j,n)A_{j+2}, \qquad i=1 \\ Q_a(0,n)[Q_r(0,n)+\sum_{j=2}^{n} Q_r(j,n)(1-A_j)] \\ \quad +Q_a(1,n)\sum_{j=0}^{n} Q_r(j,n)A_{j+1}, \qquad i=0 \\ Q_a(0,n)\sum_{j=1}^{n} Q_r(j,n)A_j, \qquad i=-1 \end{cases}$$

where A_k is the probability of a successful transmission among $k \geq 2$. Denote by a_i^k the event that transmission of some tagged mobile i is successful when having $k-1$ other simultaneous transmissions. It can be derived by the following decomposition $a_i^k = \sum_{t=2}^{N} P(\textit{mobile i transmits with power level } p_t \cap \textit{ other mobiles transmit with powers less than } p_t \cap \textit{ SINR of mobile i is greater than the threshold } \gamma_{th})$. Sine mobiles are assumed to be symmetric, then $A_k = \sum_{i=1}^{k} a_i^k = k.a_i^k$, it follows that

$$A_k = k\sum_{l=0}^{N-2}\sum_{k_1=0}^{k-1}\sum_{k_2=0}^{k-1}\cdots\sum_{k_{N-l-1}=0}^{k-1} x_1^{k_1}\cdot x_2^{k_2}\cdots x_{N-l-1}^{k_{N-l-1}}\cdot x_{N-l}^{1}\cdot\delta\left(k-1-\sum_{s=1}^{N-l-1}k_s\right) \cdot u\left(\frac{p_{N-l}}{\sum_{s=1}^{N-l-1}p_s k_s+\sigma^2/g}-\gamma_{th}\right), \tag{1.4}$$

with $A_0 = 0$ and $A_1 = 1$. x_s denotes the probability that a user (with new arrival or backlogged packet) (re)transmits using power level p_s. p_{N-l} is the power level chosen by the terminal whose transmission maybe potentially succeed, i.e., it's corresponding to the highest power selected in the current slot. Whereas k_s denotes the number of

terminals those choose the power level p_s. $\delta(t)$ (Dirac distribution) and $u(t)$ (unite echelon) are defined as following

$$\delta(t) = \begin{cases} 1 & if \quad t = 0 \\ 0 & else \end{cases} \qquad u(t) = \begin{cases} 1 & if \quad t \geq 0 \\ 0 & else \end{cases} \tag{1.5}$$

Computing the success probability is a very hard issue. The difficulty in formula (1.4) is to consider one single transmitting mobile at the highest power level and list all the cases where the $k-1$ remaining mobiles transmit at lower power. This corresponds exactly to the set of partitions [2] of the positive integer $k-1$ considering all possible permutations. Generating all the partitions of an integer is widely studied in the literature and several algorithms were proposed, e.g., see [16] and [124]. The computational complexity of such algorithms is very high and may takes long time to list the set of all partitions as well as their permutations. However, in our model the success probability depends on none of the following: the instantaneous backlog of the system n; the arrival probability q_a; and, the retransmission probability q_r. Henceforth, success probability matrix $\mathbf{A} = (A_k)$, $k = 1 \cdots m$ can be calculated once and reused to derive the transition matrix.

1.3.2 Scheme 2 : Retransmission with more power

Now, backlogged packets are assumed to have more priority; a tagged mobile having a backlogged packet retransmits it using a random power level among the N available, while a mobile with a new arrived packet uses always the lowest power level (p_1). A successful capture is occurred when a backlogged packets is transmitted at a power level larger than those chosen by all other transmitters and the AP experiences an SINR larger than the threshold γ_{th}, or a single new arrival occurs and there is no retransmission attempt by any backlogged users. The transition probabilities are then given

[2] A partition of a positive integer n is a way of writing n as a sum of positive integers.

by

$$P_{n,n+i} = \begin{cases} Q_a(m-n,n)\sum_{j=1}^{n} Q_r(j,n)(1-B_{j,m-n}), & i = m-n, i \geq 2 \\ Q_a(i,n)\sum_{j=0}^{n} Q_r(j,n)(1-B_{j,i}) & \\ \quad +Q_a(i+1,n)\sum_{j=1}^{n} Q_r(j,n)B_{j,i+1}, & 2 \leq i < m-n \\ Q_a(1,n)\sum_{j=1}^{n} Q_r(j,n)(1-B_{j,1}) & \\ \quad +Q_a(2,n)\sum_{j=1}^{n} Q_r(j,n)B_{j,2}, & i = 1 \\ Q_a(0,n)[Q_r(0,n)+\sum_{j=2}^{n} Q_r(j,n)(1-B_{j,0})] & \\ \quad +Q_a(1,n)\sum_{j=0}^{n} Q_r(j,n)B_{j,1}, & i = 0 \\ Q_a(0,n)\sum_{j=1}^{n} Q_r(j,n)B_{j,0}, & i = -1 \end{cases}$$

where the probability of a successful transmission among k retransmissions and k' new arrival packets ($k + k' \geq 2$) is given by

$$\begin{aligned} B_{k,k'} &= k \sum_{l=0}^{N-2} \sum_{k_1=0}^{k-1} \cdots \sum_{k_{N-l-1}=0}^{k-1} \prod_{i=1}^{N-l-1} x_i^{k_i} \cdot x_{N-l}^{1} \cdot \delta\left(k-1-\sum_{s=1}^{N-l-1} k_s\right) \\ &\quad \cdot u\left(\frac{p_{N-l}}{\sum_{s=1}^{N-l-1} p_s k_s + k' p_1 + \sigma^2/g} - \gamma_{th}\right), \end{aligned} \tag{1.6}$$

with $B_{0,0} = 0$, $B_{0,1} = 1$ and $B_{1,0} = 1$.

1.3.3 Scheme 3 : Retransmission with less power

New arrivals are now transmitted at the highest power level , i.e., p_N, whereas mobiles having backlogged packets attempt new retransmission using a random power picked

from $N-1$ distinct lower power levels. The transition matrix is summarized as

$$P_{n,n+i} = \begin{cases} Q_a(i,n), & 2 \leq i \\ Q_a(1,n) \sum_{j=1}^{n} Q_r(j,n)(1-C_{j,1}), & i = 1 \\ Q_a(0,n)[Q_r(0,n) + \sum_{j=2}^{n} Q_r(j,n)(1-C_{j,0})] & \\ \quad + Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n) C_{j,1}, & i = 0 \\ Q_a(0,n) \sum_{j=1}^{n} Q_r(j,n) C_{j,0}, & i = -1 \end{cases}$$

where the probability of a successful transmission when $k \geq 1$ mobiles attempt retransmissions is given by the following expression

$$\begin{aligned} C_{k,1} &= k \sum_{k_1=0}^{k-1} \sum_{k_2=0}^{k-1} \cdots \sum_{k_{N-1}=0}^{k-1} \prod_{i=1}^{N-1} x_i^{k_i} \cdot x_{N-l}^{1} \cdot \delta \left(k - \sum_{s=1}^{N-1} k_s \right) \\ &\quad \cdot u \left(\frac{p_N}{\sum_{s=1}^{N-1} p_s k_s + \sigma^2/g} - \gamma_{th} \right). \end{aligned} \tag{1.7}$$

Similarly, the probability of a successful retransmission among $k \geq 2$ simultaneous retransmissions is given by

$$\begin{aligned} C_{k,0} &= k \sum_{l=1}^{N-2} \sum_{k_1=0}^{k-1} \sum_{k_2=0}^{k-1} \cdots \sum_{k_{N-l-1}=0}^{k-1} \prod_{i=1}^{N-l-1} x_i^{k_i} \cdot x_{N-l}^{1} \cdot \delta \left(k - 1 - \sum_{s=1}^{N-l-1} k_s \right) \\ &\quad \cdot u \left(\frac{p_{N-l}}{\sum_{s=1}^{N-l-1} p_s k_s + \sigma^2/g} - \gamma_{th} \right), \end{aligned} \tag{1.8}$$

whereas $C_{k,k'} = 0$ if $K' \geq 2$, $C_{0,1} = 1$ and $C_{1,0} = 1$.

1.3.4 Scheme 4 : Retransmission with the lowest power

In this last proposal, a new transmitted packet uses a power among $N-1$ higher available power levels. Mobiles having backlogged packet retransmit with the lowest power

level, i.e., p_1. The transition matrix of the Markov chain is given by

$$P_{n,n+i} = \begin{cases} Q_a(m-n,n)\sum_{j=0}^{n} Q_r(j,n)(1-D_{j,m-n}), & i = m-n, i \geq 2 \\ Q_a(i,n)\sum_{j=0}^{n} Q_r(j,n)(1-D_{j,i}) & \\ \quad +Q_a(i+1,n)\sum_{j=0}^{n} Q_r(j,n)D_{j,i+1}, & 2 \leq i < m-n \\ Q_a(1,n)\sum_{j=1}^{n} Q_r(j,n)(1-D_{j,1}) & \\ \quad +Q_a(2,n)\sum_{j=0}^{n} Q_r(j,n)D_{j,2}, & i = 1 \\ Q_a(0,n)[Q_r(0,n)+\sum_{j=2}^{n} Q_r(j,n)(1-D_{j,0})] & \\ \quad +Q_a(1,n)\sum_{j=0}^{n} Q_r(j,n)D_{j,1}, & i = 0 \\ Q_a(0,n)Q_r(1,n), & i = -1 \end{cases}$$

where $D_{k,k'}$ represents the probability of a successful new transmission among k backlogged packets and k' new packets such that $k' + k \geq 2$. The value $D_{k,k'}$is given by

$$\begin{aligned} D_{k,k'} &= k' \sum_{l=0}^{N-2} \sum_{k'_1=0}^{k-1} \cdots \sum_{k'_{N-l-1}=0}^{k'-1} \prod_{i=1}^{N-l-1} x_i^{k'_i} \cdot x_{N-l}^{1} \cdot \delta\left(k' - 1 - \sum_{p_l=1}^{N-l-1} k'_{p_l}\right) \\ &\quad \cdot u\left(\frac{p_{N-l}}{\sum_{p_l=1}^{N-l-1} p_l k'_{p_l} + kP_1 + \sigma^2/g} - \gamma_{th}\right), \end{aligned} \tag{1.9}$$

where $D_{0,0} = 0$, $D_{0,1} = 1$ and $D_{1,0} = 1$

1.3.5 Performance metrics

We now turn to present the performance measures (average throughput and expected delay) of interest for optimization as a function of the steady-state probabilities of the Markov chain. Let us denote by $\pi_n(q)$ the equilibrium probability that the network is in state n (number of backlogged packets at the beginning of a slot). Hence the equilibrium state equations are

$$\begin{cases} \bar{\pi}(q) = \bar{\pi}(q) \cdot P(q) \\ \sum_{n=0}^{m} \pi_n(q) = 1 \\ \pi_n(q) \geq 0, \qquad n = 0, 1, ..., m \end{cases} \tag{1.10}$$

System (1.10) yields the equilibrium probabilities vector. The average number of backlogged packets can then simply be calculated by

$$S(q) = \sum_{n=0}^{m} \pi_n(q) \cdot n. \tag{1.11}$$

Similarly, the system throughput (defined as the sample average of the number of packets that are successfully transmitted) is given almost surely by the constant

$$\begin{cases} \sum_{n=0}^{m} \sum_{i=0}^{m-n} \sum_{j=0}^{n} \pi_n(q) Q_a(i,n) Q_r(j,n) A_{j+i}, & \text{Scheme 1} \\ \sum_{n=0}^{m} \pi_n(q) \left[\sum_{i=0}^{m-n} \sum_{j=1}^{n} Q_a(i,n) Q_r(j,n) B_{j,i} + Q_a(1,n) Q_r(0,n) \right], & \text{Scheme 2} \\ \sum_{n=0}^{m} \pi_n(q) \left[Q_a(0,n) \sum_{j=1}^{n} Q_r(j,n) C_{j,0} + Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n) C_{j,1} \right], & \text{Scheme 3} \\ \sum_{n=0}^{m} \pi_n(q) \left[\sum_{i=1}^{m-n} \sum_{j=0}^{n} Q_a(i,n) Q_r(j,n) D_{j,i} + Q_a(0,n) Q_r(1,n) \right], & \text{Scheme 4} \\ \sum_{n=0}^{m} \pi_n(q) \left[Q_a(0,n) Q_r(1,n) + Q_a(1,n) Q_r(0,n) \right]. & \text{Same power} \end{cases}$$

Using the rate balance equation (i.e., input=output) at the steady state, the throughput satisfies (and thus can be computed more easily through)

$$thp(q) = q_a \sum_{n=0}^{m} \pi_n(q)(m-n) = q_a(m - S(q)). \tag{1.12}$$

Indeed, the throughput is the expected number of arrivals at a time slot (which actually enter the system), and this is expressed in the equation for $thp(q)$ by conditioning on n. The throughput should be equal to the expected number of departures (and thus the throughput) at stationary regime, which is expressed in (1.12). The expected delay of transmitted packets denoted by $D(q)$, is defined as the average time, in slots, that a packet takes from its source to the receiver. It can easily be derived by Little's result (see [93] for detailed overview and theorem proof), namely

$$D(q) = 1 + \frac{S(q)}{thp(q)} = 1 + \frac{S(q)}{q_a(m - S(q))}. \tag{1.13}$$

Analyzing equations (1.12) and (1.13), it is easy to show that maximizing the global throughput is equivalent to minimizing the average delay of transmitted packets. We shall therefore restrict in our numerical investigation to maximize average the throughput. However, we shall consider the delay of backlogged packets as another objective to minimize. This latter metric has a great interest, in particular while addressing the support of real-time or more generally delay sensitive applications.

1.3.6 Performance measures for backlogged packets

Let us denote by Δ the throughput of new arrivals, i.e., the amount of arrivals whose the first transmission attempt is successful. Therefore, the throughput of backlogged packets for each scheme is given almost surely by: $thp^c(q) = thp(q) - \Delta$, where Δ is calculated from Markov chain and given by

$$\begin{cases} \sum_{n=0}^{m} \sum_{i=1}^{m-n} \sum_{j=0}^{n} \frac{i}{i+j} \pi_n(q) Q_a(i,n) Q_r(j,n) A_{i+j}, & \text{Scheme 1} \\ \sum_{n=0}^{m} \pi_n(q) Q_a(1,n) Q_r(0,n), & \text{Scheme 2} \\ Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n) C_{j,1}, & \text{Scheme 3} \\ \sum_{n=0}^{m} \sum_{i=1}^{m-n} \sum_{j=0}^{n} \pi_n(q) Q_a(i,n) Q_r(j,n) D_{j,i}, & \text{Scheme 4} \\ \sum_{n=0}^{m} \pi_n(q) Q_a(1,n) Q_r(0,n). & \text{Same power} \end{cases} \tag{1.14}$$

The expected delay of backlogged packets D^c, which is defined as the average time, in slots, that a backlogged packet takes to go from the source to the receiver, can also be calculated by applying Little's result. Hence

$$D^c(q) = 1 + \frac{S(q)}{thp^c(q)}. \tag{1.15}$$

Team problem resolution : The optimal solution of the team problem is obtained by resolving the following optimization problem:

$$\max_q \ objective(q) \ \ s.t. \begin{cases} \bar{\pi}(q) = \bar{\pi}(q) \cdot P(q) \\ \sum_{n=0}^{m} \pi_n(q) = 1 \\ \pi_n(q) \geq 0, \qquad n = 0, 1, ..., m \end{cases} \tag{1.16}$$

where $objective(q)$ is replaced by the average throughput or minus expected delay. We note that the solution can be obtained by computing recursively the steady-state probabilities, as Problem 4.1 in [25].

Singularity at q = 0 : The only point where the Markov chain P does not have a single stationary distribution is at $q = 0$, where it has two absorbing states: $n = m$ and $n = m - 1$. All remaining states are transient (for any $q_a > 0$), and the probability to end at one of the absorbing states depends on the initial distribution of the Markov chain. We note that if the state $m - 1$ is reached then the throughput is q_a w.p. 1. However, if the state m is reached then the throughput is certainly 0, which means that it is a deadlock state. For $q_a > 0$ and $q_r = 0$, the deadlock state is reached with positive probability

from any initial state other than the absorbing state $m-1$, We shall therefore exclude the case of $q_r = 0$ and optimize only on the range $\epsilon < q_r \leq 1$.

Existence of a solution : The steady-state probabilities $\bar{\pi}(q)$ are continuous over $0 < q \leq 1$. This is not a closed interval, therefore a solution needs not to exist. However, as we restrict to the closed interval $[\epsilon, 1]$, where $\epsilon > 0$, an optimal solution indeed exists. Therefore for any $\delta > 0$, there exists some $q^* > 0$ which is δ-optimal. $q > 0$ is said to be δ-optimal for the throughput maximization if it satisfies $thp(q^*) \geq thp(q) - \delta$ for all $q \in [\epsilon, 1]$. A similar definition holds for any objective function (e.g., delay minimization).

1.3.7 Stability

Another qualitative way to compare schemes is the stability characteristics of the protocol. Slotted aloha is known to have a bi-stable behavior, we hence shall check whether this is also the case in our four schemes, if the answer is yes, under which conditions ?

Let us denote p_n^{succ} the expected number of successful transmissions in the current slot having n backlogged packets. Based on the derived Markov chains, the probability of a successful transmission is given by

$$p_n^{succ}(q) = \begin{cases} \sum_{i=0}^{m-n} \sum_{j=0}^{n} Q_a(i,n) Q_r(j,n) A_{j+i}, & \text{Scheme 1} \\ \sum_{i=0}^{m-n} \sum_{j=1}^{n} Q_a(i,n) Q_r(j,n) B_{j,i} + Q_a(1,n) Q_r(0,n), & \text{Scheme 2} \\ Q_a(0,n) \sum_{j=1}^{n} Q_r(j,n) C_{j,0} + Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n) C_{j,1}, & \text{Scheme 3} \\ \sum_{i=1}^{m-n} \sum_{j=0}^{n} Q_a(i,n) Q_r(j,n) D_{j,i} + Q_a(0,n) Q_r(1,n), & \text{Scheme 4} \\ Q_a(0,n) Q_r(1,n) + Q_a(1,n) Q_r(0,n). & \text{Slotted aloha} \end{cases}$$

Define now the *drift* D_n in state n, as the expected change in backlog from one slot to the next slot, which is the expected number of arrivals, i.e., $q_a(m-n)$, less the expected number of successful departures p_n^{succ}, that is

$$D_n = q_a(m-n) - p_n^{succ}. \tag{1.17}$$

For standard slotted aloha it has been shown that three equilibria may exist. System equilibrium points occur where the curves p_n^{succ} and the straight line $q_a(m-n)$ intersect. When the drift, which represents the difference between the straight line and the curve, is positive the system state tends to increase, because the system input rate becomes greater than its output rate. Whereas it decreases when the drift is negative. This explains why the middle equilibrium point is definitely unstable and the other two are stable. A bi-stable situation as in the standard aloha is hence undesirable since it means,

in practice, that the system spends long time in each of the stable equilibria, including the one with large backlog n corresponding to a congestion situation (low throughput and large delay).

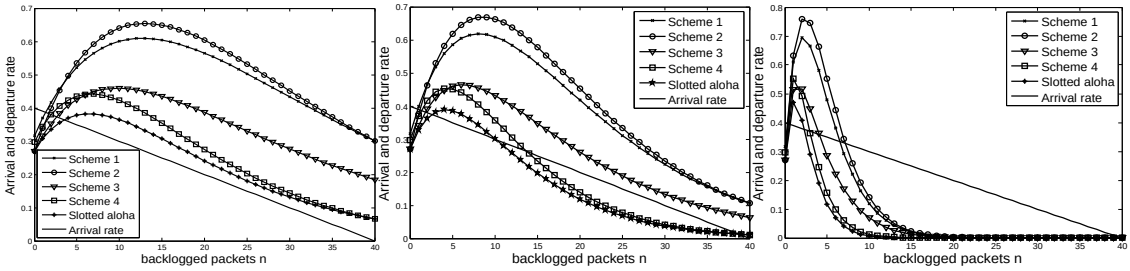


Figure 1.2: *Stability of slotted aloha with random power selection algorithms: Probability of success transmission versus backlogged packets n for all the schemes. We consider $\gamma_{th} = 10$ dB, transmit powers $\mathcal{P} = [1, 5, 25, 125, 625]$ mW, arrival probability $q_a = 0.01$ and different retransmission probabilities (aggressiveness levels) $q_r = 0.1$ (sub-figure a), $q_r = 0.25$ (sub-figure b) and $q_r = 0.5$ (sub-figure c).*

Let us now examine the stability behavior of slotted aloha and our new schemes for $m = 40$ mobiles, $\gamma_{th} = 10$dB, arrival probability $q_a = 0.01$ and $N = 5$ selectable powers, see figure 1.2. We note that no scheme suffers from the bi-stability problem under low aggressiveness $q_r = 0.1$, where the departure rate of our schemes is, all the time, greater than the arrival rate; It follows that slotted aloha under our algorithms is stable and the average number of backlogged packets is very low (which decreases significantly the expected delay). Under $q_r = 0.25$ slotted aloha and scheme 4 become unstable whereas other schemes with random power keep stability whatever the average number of backlogged packets. In contrast to standard slotted aloha and through simple computation of equilibria, the expected number of backlogged packets for schemes 1-4 can be approximated by the desired stable equilibrium which provides a very interesting feature. That means that in the bi-stability case for schemes 1-4, the system spends most of the time at that desired equilibrium. Next, we note that success probability P_n^{succ} decreases with n and vanishes for all schemes for $q_r > 0.5$, where all schemes acquire a bi-stable behavior; It follows that the stability region is tightening with transmission rate q_r. Indeed slotted aloha and all other schemes become instable when mobiles retransmit aggressively. Then the collision probability is close to 1 and the departure rate becomes less than the arrival rate which causes absorption of the system by the undesired equilibrium point. Under this situation the average number of backlogged packets can be approximated by the non-desired equilibrium. Here, the system spends most of the time at that equilibrium.

The observed bipolar behavior of slotted aloha as well as in our schemes can be avoided by decreasing the probability of retransmission. Yet, decreasing q_r expands the departure rate curve which removes all intersections with arrival straight line but the desired stability point. The drawback is that expected delay increases since a packet will

wait longer at a backlogged node before a successful transmission and the maximum throughput decreases slightly.

1.4 Game problem

Implementation of a centralized system is a real issue since it needs a high technology, advanced software performances, extra signaling protocol and full information about mobiles and their instantaneous QoS (Quality of Service) requirements. This way, a high amount of bandwidth should be reserved for signaling. Moreover, mobile users are not forced to cooperate and may act in a selfish way. We derive in this section an alternative framework which is decentralized. We shall then formulate a distributed model using game theory tools. Yet, the decentralized model is more powerful, more appropriate for slotted aloha and has high interest for analyzing random access games as well. Naturally, the Nash equilibrium concept will replace the optimality concept used in the team problem. It possesses a robustness property: At equilibrium and assuming rationality of mobiles, no mobile has incentive to deviate. The elements of our contention game are listed bellow

- A finite set of $m+1$ bufferless users interacts over a single collision channel.
- Each user i retransmits, in every slot, its packets with probability q_r^i. The open interval $q_r^i \in (0,1]$ corresponds to the set of all possible actions of tagged user i.
- There is no cost of (re)transmitting packets.
- The individual throughput or minus expected delay of backlogged packets are the utility functions to maximize.

For any instance of the game, we denote the policy vector of retransmission probabilities of all users by $\bar{\mathbf{q}}_r$ whose j^{th} entry is q_r^j. Define (q_r^{-i}, q_r^i) to be a retransmission policy at a slot, where user i retransmits with probability q_r^i and any other user j retransmits with probability q_r^j for all $j \neq i$. Each user i seeks to maximize his own $objective_i(\bar{\mathbf{q}}_r)$, either each mobile maximizes its own throughput or maximizes minus expected delay. We are interested here to find a symmetric equilibrium policy $\bar{\mathbf{q}}_r^* = (q_r, q_r, ..., q_r)$ such that for any user i and any retransmission probability q_r^i for that user,

$$objective_i\left(\bar{\mathbf{q}}_r^*\right) \geq objective_i\left(\left[\bar{\mathbf{q}}_r^*\right]^{-i}, q_r^i\right). \tag{1.18}$$

Without any loss of generality, we restrict to symmetric policy $\bar{\mathbf{q}}_r^*$ where all mobiles are balanced-payoff. We shall also identify it (with some abuse of notation) with the actual transmission probability which is the same for all users. We first note that due to symmetry, to see whether $\bar{\mathbf{q}}_r^*$ is an equilibrium it suffices to check (1.18) for a single player. We shall thus assume that m users retransmit with a given probability $\left[\bar{\mathbf{q}}_r\right]^{-(m+1)} = (q_r^o, q_r^o, ..., q_r^o)$ and the user $m+1$ retransmits with probability q_r^{m+1}. Define

the set $\mathcal{Q}^{m+1}(\bar{\mathbf{q}}_r^o)$ as the set of best response strategies of user $m+1$, it can be written as

$$\mathcal{Q}^{m+1}(\bar{\mathbf{q}}_r^o) = \underset{q_r^{m+1}\in[\epsilon,1]}{\text{argmax}} \left(objective_{m+1} \left([\bar{\mathbf{q}}_r^o]^{-(m+1)}, q_r^{m+1} \right) \right), \tag{1.19}$$

where $\bar{\mathbf{q}}_r^o$ denotes the policy where all users retransmit with probability q_r^0 and the maximization is taken with respect to q_r^{m+1}. Then $\bar{\mathbf{q}}_r^*$ is a symmetric equilibrium if

$$\bar{\mathbf{q}}_r^* \in \mathcal{Q}^{m+1}\left(\bar{\mathbf{q}}_r^*\right). \tag{1.20}$$

To compute the performance measures of interest, we introduce again a Markov chain with a two dimensional state, see figure 1.3. The first state component corresponds to the number of backlogged packets among users $1, ..., m$, and the second component is the number of backlogged packets (either 1 or 0) of tagged user $m+1$.

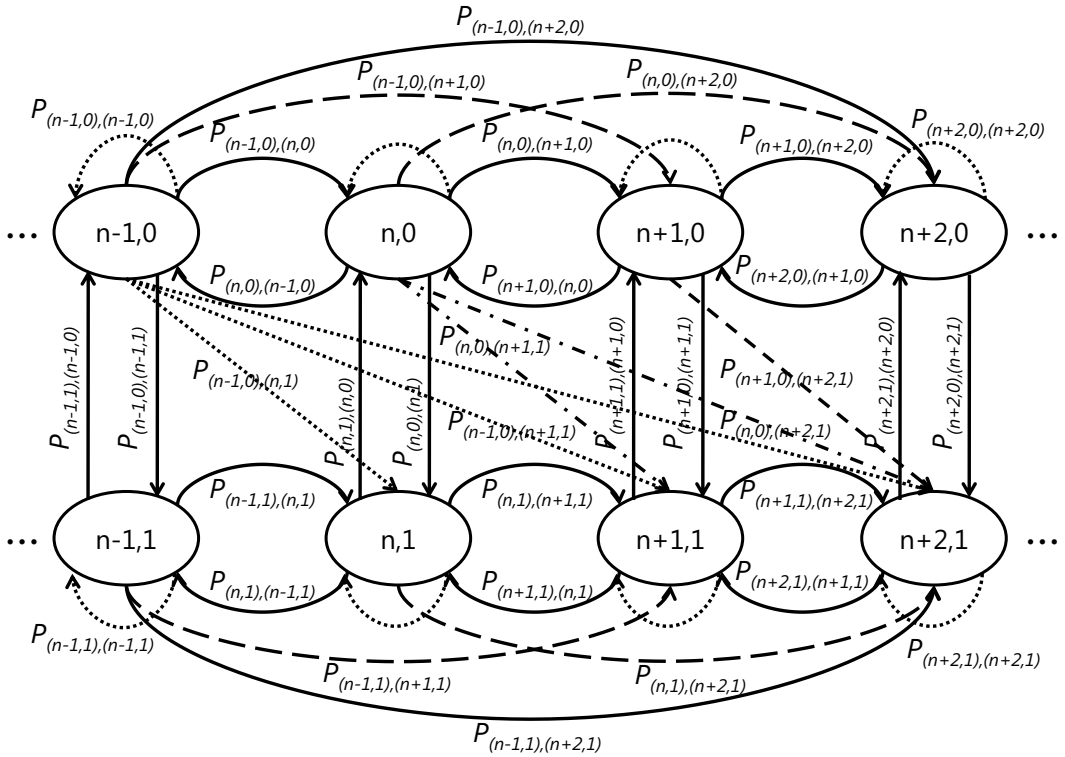


Figure 1.3: *Bi-dimensional Markov chain for the game setting. The state of the system is the backlog vector; the first component corresponds to the number of backlogged packets for m first mobiles whereas the second component indicates the number of backlogged packets of the tagged mobile $m+1$ (either 0 or 1).*

1.4.1 Scheme 1 : Random power without priority

We first consider the setting in which packets are transmitted/retransmitted with a random power selected randomly with the probability vector $X = [x_1, x_2, ..., x_N]$. In

this scheme, there is no priority between classes of packets. The transition probabilities when a tagged mobile uses q_r^{m+1} and the m other mobiles use q_r^0 are given in Appendix F.

1.4.2 Scheme 2 : Retransmission with more power

We consider now, the proposal in which each backlogged packet is retransmitted with a random power picked from $N-1$ high levels. A new arrival packet is always transmitted with the lowest power level, i.e., p_1. The transition probabilities when m mobiles use q_r^0 and the tagged mobile $m+1$ transmits at a rate q_r^{m+1} are detailed in Appendix G.

1.4.3 Scheme 3 : Retransmission with less power

In this scheme, we assume that new arrivals always transmit over the channel using the highest power level p_N. Whereas backlogged packets attempt retransmission with a random power picked from remaining $N-1$ lower levels. The corresponding transition matrix whose elements are $p_{(n,a),(n+i,b)}$ is summarized in Appendix H.

1.4.4 Scheme 4 : Retransmission with the lowest power

We finally consider the non-cooperative scenario where backlogged packets are retransmitted with the lowest power level p_1. Here, new arrival packets are transmitted with a power selected among $N-1$ distinct higher levels. The transition matrix of this scheme is given in Appendix I.

1.4.5 Performance metrics

Let $\pi_{n,a}$ be the steady-state of the Markov chain where n is the actual number of backlogged packets among the m first mobiles and a is the binary-valued number of backlogged packets of user $m+1$. The average number of backlogged packets of source $m+1$ is written as

$$S_{m+1}([\bar{\mathbf{q}}_r^o]^{-m+1}, q_r^{m+1}) = \sum_{n=0}^{m} \pi_{n,1}([\bar{\mathbf{q}}_r^o]^{-(m+1)}, q_r^{m+1}), \tag{1.21}$$

and the average throughput of user $m+1$ is almost surely given by

$$thp_{m+1}([\bar{\mathbf{q}}_r^o]^{-(m+1)}, q_r^{m+1}) = q_a \sum_{n=0}^{m} \pi_{n,0}([\bar{\mathbf{q}}_r^o]^{-(m+1)}, q_r^{m+1}). \tag{1.22}$$

Thus, the expected delay of transmitted packets at source $m+1$ for all schemes verifies (Little's result)

$$D_{m+1}([\bar{\mathbf{q}}_r^o]^{-(m+1)}, q_r^{m+1}) = 1 + \frac{S_{m+1}([\bar{\mathbf{q}}_r^o]^{-(m+1)}, q_r^{m+1})}{thp_{m+1}([\bar{\mathbf{q}}_r^o]^{-(m+1)}, q_r^{m+1})}. \tag{1.23}$$

Performance measures for backlogged packets. Let us denote the throughput of backlogged packets (i.e., packets that arrive and become backlogged) at source $m+1$ by

$$thp^c_{m+1}(\mathbf{q}^{m+1}) = \sum_{n=0}^{m}\sum_{n'=0}^{m} p_{(n,0),(n',1)}(\mathbf{q}^{m+1})\pi_{n,0}(\mathbf{q}^{m+1}). \tag{1.24}$$

Thus, the expected delay of backlogged packets at source $m+1$ is

$$D_{m+1}(\mathbf{q}^{m+1}) = 1 + \frac{S_{m+1}(\mathbf{q}^{m+1})}{thp^c_{m+1}(\mathbf{q}^{m+1})}. \tag{1.25}$$

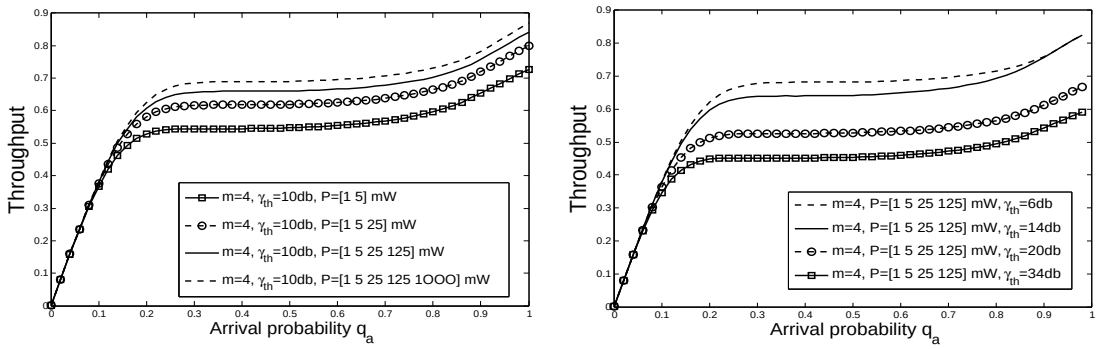
1.5 Numerical investigation

We conduct here an extensive numerical investigation of the discussed cooperative and non-cooperative frameworks of slotted aloha under power diversity, packets priority and capture effect. We fix throughout this section $\epsilon = 10^{-4}$, i.e., the solution q^* will be searched in the closed interval $[\epsilon, 1]$ instead of $]0, 1]$, see Subsection 1.3.6 for detailed explanation.

1.5.1 Impact of system parameters

In the following we investigate the impact of each parameter on the protocol performance. For illustrative purpose, we use scheme 1 and yield similar results for other schemes. We also focus on the throughput as a measurement metric, other metrics such as delay provide similar trends. We plot the throughput as a function of arrival probability when changing the number of available power levels, figure 1.4(a). We note that the throughput is improved when increasing the number of power levels. This is quite intuitive since mobiles will have larger choices and then high chance to decode correctly the received signal by the central receiver.

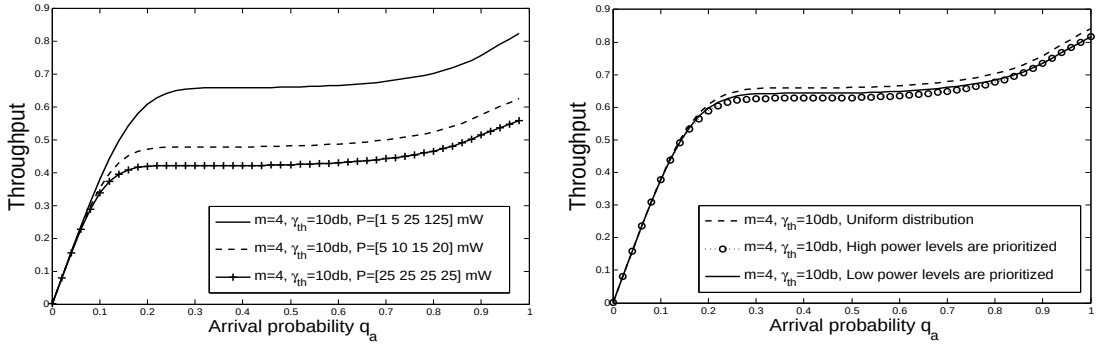
Next we depict the average throughput for different values of threshold SINR γ_{th}, see figure 1.4(b). The system performances are deteriorated while increasing γ_{th}. When the access point requires a high signal quality (i.e., high value of γ_{th}) to decode correctly the received signals, the loss probability becomes very high and therefore the throughput decreases. Furthermore, we note that the best performances are obtained when the available power levels follow a geometric progression (see figure 1.5(a)) whereas same power scenario (the case of standard slotted aloha) provides the lowest performance. This can be explained by the distance between power levels and its strong impact on the instantaneous SINR value. For a tagged mobile i, with geometric or more rapid progression, only successive power levels may interfere significantly with the chosen power p^i_l, whereas all other power levels have high interfering capability on transmitted packet when using arithmetic progression. It is also interesting to show that performances are improved when increasing the geometric step. Another parameter which influences



(a) (b)

Figure 1.4: *Impact of the number of available power levels (a) and threshold SINR γ_{th} (b) on the global throughput under scheme 1.*

on the whole performances is the probability distribution $X = [x_1, x_2, \cdots, x_N]$ to select transmit powers. Figure 1.5 (b) shows that uniform distribution (i.e., $x_i = 1/N$) is the best, whereas prioritizing high power levels provides the lowest performances. Yet, prioritizing some power levels leads the system to behave as an equivalent system with less number of available power levels, which explains the observed decrease of performance. When available power levels are uniformly distributed, they all have the same chance to be used and henceforth that fashion outperforms other distributions.



(a) (b)

Figure 1.5: *This figure shows the impact of power levels (a) and selection probabilities distribution (b) on the global throughput under scheme 1.*

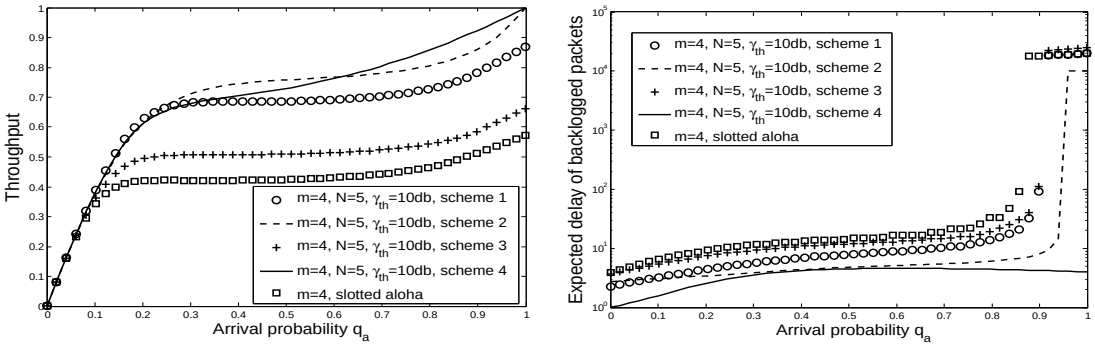
1.5.2 Team problem

Aggregate throughput maximization

In the following, we are interested in the symmetric solution that maximizes the global throughput. We depict in figure 1.6-1.8 the throughput, expected delay of backlogged packets (EDBP) and optimal retransmission probability for all addressed schemes.

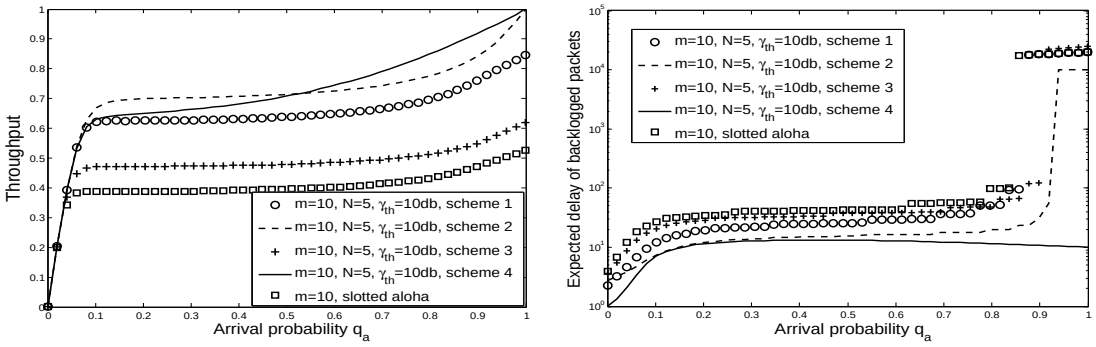
We fix the threshold SINR to $\gamma_{th} = 10$dB and consider five selectable power levels $\mathcal{P} = [1, 5, 25, 125, 625]$ mW for all schemes. Note that slotted aloha can be obtained from scheme 1 using same power policy and infinite value of γ_{th}.

First, we evaluate the system performance in terms of aggregate throughput for 4 mobiles. In figure 1.6, we plot the global throughput as a function of arrival probability q_a. At very low load ($q_a < 0.1$), all schemes have likely same performance which can be approached by a linear function in q_a. For low load ($0.1 < q_a < 0.24$), scheme 2 seems to perform slightly better than other schemes. When arrival rate is average (between 0.24 and 0.6), scheme 2 performs better and provides higher throughput. This is due to the fact that scheme 2 prioritizes the retransmission of backlogged packets exploiting the fact that there is few new arrivals. But at high load the throughput of scheme 4 becomes the highest because it prioritizes new arrivals. Clearly, new arrival packets have an extended choice of power levels and therefore benefit from prioritization and power diversity.



(a) (b)

Figure 1.6: *Aggregate throughput and expected delay of backlogged packets for 4 mobiles under the team problem and throughput maximization.*

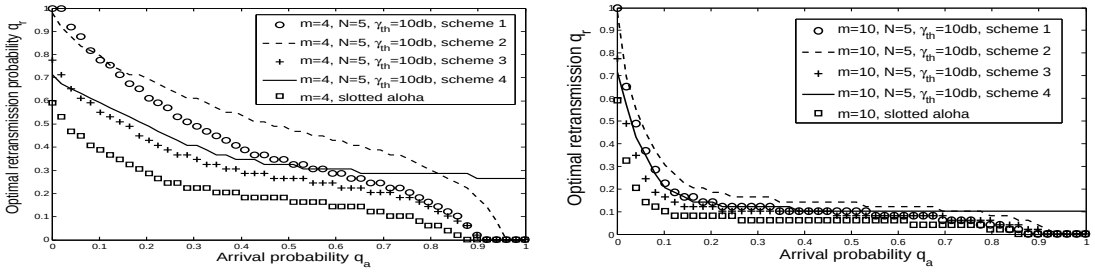


(a) (b)

Figure 1.7: *Aggregate throughput and expected delay of backlogged packets for 10 mobiles under the team problem and throughput maximization.*

We remark that scheme 3 (retransmission with less power) which is the same as the one first proposed in [8] presents the least performance compared to other schemes. This is due to the negative effect of large choice of power levels reserved to backlogged packets and in particular to the penalizing capture effect. This way, the instantaneous SINR of new arrival becomes strongly noised by backlogged packets. However, we note that all schemes with random power selections and capture effect outperform standard slotted aloha.

In terms of expected delay of backlogged packets, at low and average load and for all schemes (see figure1.6(b)), we obtain a slightly increasing function of the arrival probability, it can be approached by a semi-constant value. For instance, the average backoff duration of schemes 1-4 and slotted aloha can be approximated by $d_1 \simeq 7$, $d_2 \simeq 4$, $d_3 \simeq 10$, $d_4 \simeq 4$ and $d_0 \simeq 11$ slots respectively. It is clear that schemes 3 and 4 are the best either in terms of throughput and delay. At heavy and very heavy loads ($0.75 < q_a$), delay of schemes 1-3 and slotted aloha increases exponentially. Whereas scheme 4 holds a semi-constant value for expected delay. Under scheme 4, backlogged packets have extended choices (all power levels indexed as $p_1, p_2, ..., p_{N-1}$), this gives advantage to retransmissions when there is no new arrival. Hence,latter fashion provides a very good amount of successful transmission under a low delay tradeoff. The dramatic huge of delay of other schemes decreases significantly the system reliability by causing very large backoff stage. Then schemes 1-3 and slotted aloha are not recommended for the delay sensitive applications such real time services (e.g., voice, streaming ...) elsewhere the system reliability and QoS guaranteeing become a hard issue. Whereas scheme 4 seems to be perfectly adapted to support these classes of services since its respective backoff stage is strongly reduced.



(c) (e)

Figure 1.8: *Optimal retransmission probability for schemes 1-4 and slotted aloha versus arrival rate q_a for both 4 mobiles and 10 mobiles when throughput is maximized.*

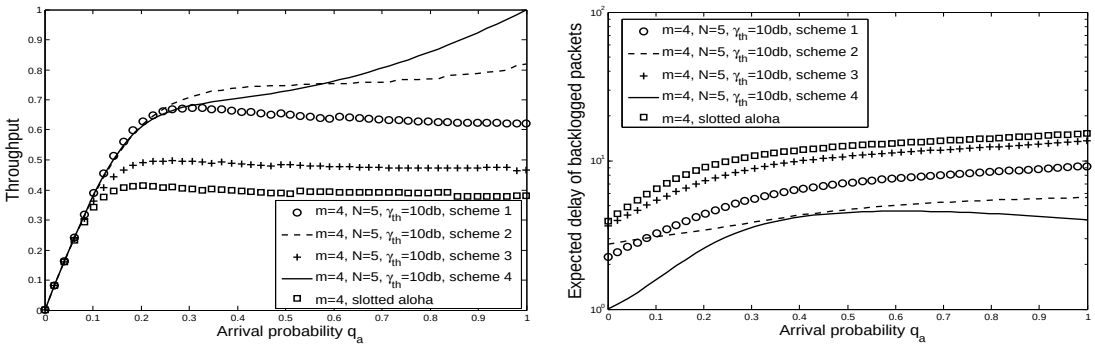
Next we plot the optimal retransmission probability versus offered load. We remark that for $m = 4$, see figure 1.8(a), all schemes optimal retransmission probabilities q_r are decreasing with q_a until to be semi-annulled ($q_r \simeq 10^{-4} = \epsilon$ because we only consider solutions in $[\epsilon, 1]$) for schemes 1-3 and slotted aloha. This explain the huge EDBP seen for these schemes: each backlogged packet stays a long time in the system. Whereas retransmission rate keeps a constant value (around 0.3) for scheme 4 (for q_a over 0.5)

because it prioritizes new packets and then it hurts not from backlogged packets. This later scheme seems to be the most fair since both new arrivals as well as backlogged packets take advantage from a high throughput, low delay and then may meet a good channel utilization.

For $m = 10$, see figure 1.7(a-b) and figure 1.8(b), we observe similar trends in terms of throughput and delay for all schemes. In fact even if the number of mobiles becomes large, a good whole performance is handled by decreasing retransmission rates so as to avoid/reduce potential collisions. We remark that at heavy load the base station asks mobiles to decrease their retransmission probabilities to avoid collisions, therefore the system keeps a very good amount of successful departure. Henceforth an optimal value of throughput is achieved and is much better compared to slotted aloha.

Delay minimization

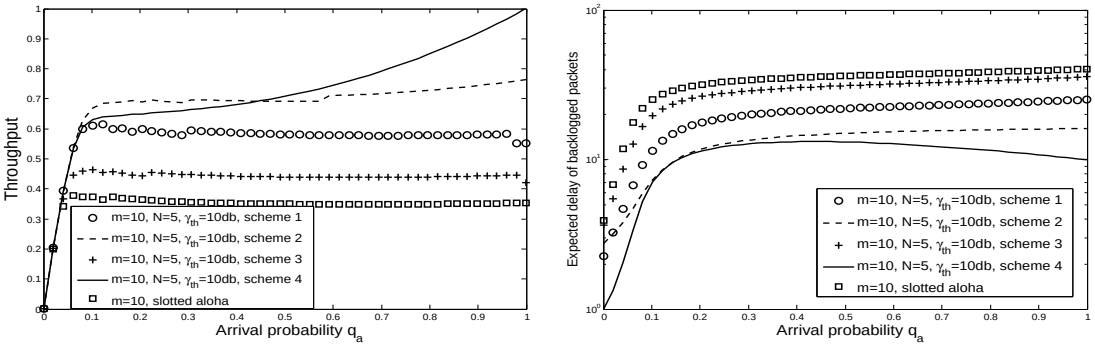
When maximizing the global throughput, we observed a huge EDBP for all schemes 1-3 and slotted aloha, in particular at heavy load, whereas scheme 4 keeps a constant low delay. This may be very harmful for many applications which are delay-sensitive (real-time applications). Now we shall investigate the problem of minimizing EDBP and the impact of this optimization (figure 1.9(a) and (b)) on the throughput performance. We note in particular that throughput performance in the four schemes improves considerably with respect to slotted aloha. Scheme 1 is slightly better in terms of throughput only at light load, scheme 2 is almost better at medium load whereas scheme 4 outperforms remarkably all other schemes at high ($0.55 < q_a$) and very high loads. The case of 10 mobiles provides similar trends, (figure 1.10.



(a) (b)

Figure 1.9: *Aggregate throughput and expected delay of backlogged packets for 4 mobiles under the team problem. The objective is to minimize the expected delay of backlogged packets.*

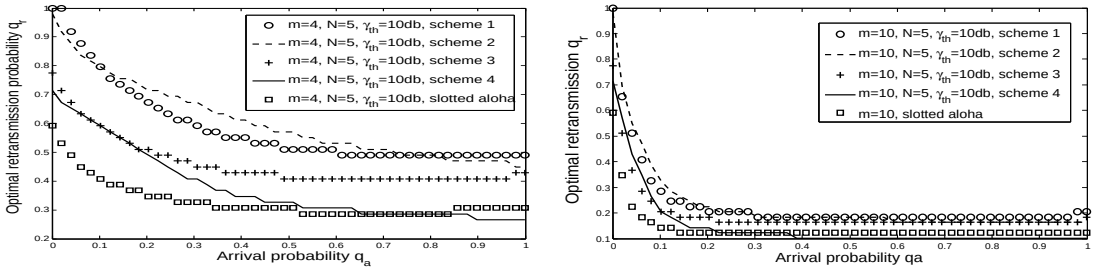
When EDBP is minimized, for $m = 4$ and $N = 5$, retransmission probability decreases with q_a, so standard aloha and scheme 4 have optimal retransmission probability of around 0.3 at heavy load whereas algorithms 1-3 have much higher retransmission



(a) (b)

Figure 1.10: *Aggregate throughput and expected delay of backlogged packets for 10 mobiles under the team problem. The objective is to minimize the expected delay of backlogged packets.*

probabilities (figure 1.11(a)). For larger mobile population $m = 10$ (figure 1.11(b)), we observe that optimal retransmission rate falls down rapidly for all schemes. Numerically, when traffic is high, slotted aloha mobiles retransmit with probability around 0.13, around 0.1 for scheme 4, and around 0.19 for schemes 1-3 at average and heavy load. In the team problem we note that the optimal retransmission probability q_r decreases when increasing arrival probability q_a and viceversa. This means that mobiles have to cooperate (adapt their retransmission probabilities according to the load) to reach the best performance either when maximizing global throughput or minimizing EDBP.



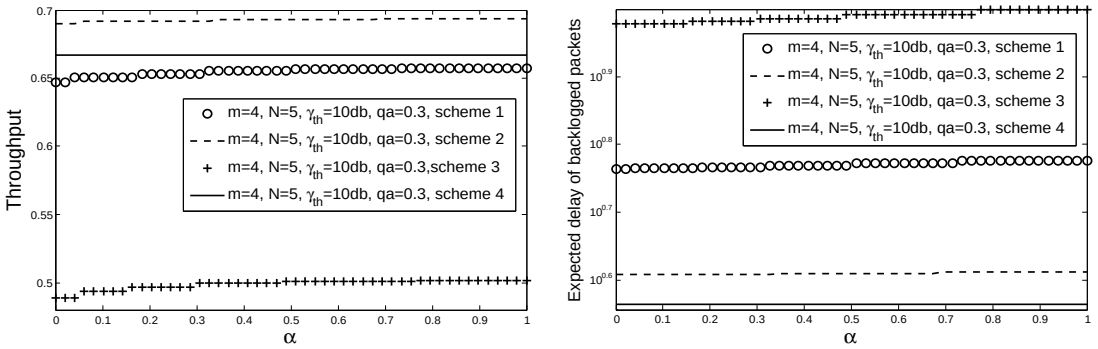
(a) (b)

Figure 1.11: *The optimal retransmission probability for schemes 1-4 and slotted aloha as function of arrival rate q_a for both 4 mobiles and 10 mobiles when throughput is maximized.*

Discussion and remarks

In previews simulations we considered the extreme cases of maximizing independently the throughput and minimizing the EDBP. Radunovic and Le Boudec [117] suggest that considering the total throughput as a performance objective may not be a good objective, this yields also for delay. In practice it may be more interesting to have a multi-criteria optimization in which a convex combination of both the throughput and EDBP is optimized. We consider the following objective $\alpha thp(q) + (1-\alpha)/D^c(q), 0 \leq$

$\alpha \leq 1$. This allows in particular handling QoS constraints: By varying α one can find appropriate tradeoff between the throughput and the expected delay, so that the throughput be maximized while keeping the EDBP bounded by some constant. This improves considerably the system reliability and makes the system able to support several kind of services with different QoS requirements.

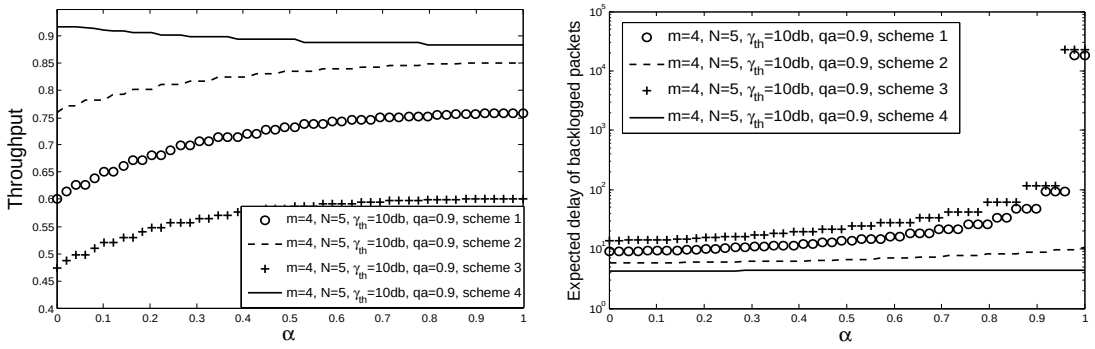


(a) (b)

Figure 1.12: *Joint throughput and delay convex optimization. This figures shows the impact of multi-criteria factor α on aggregate throughput and expected delay of backlogged packets at low load $q_a = 0.3$.*

At low load ($q_a = 0.3$) and under schemes 2 and 4, see figure 1.12(a and b), the optimal throughput and EDBP are slightly constant, so insensitive to the values of the weight α under different loads. Because optimal retransmission probabilities under separated objectives (maximizing throughput or minimizing delay) are so close. This means that for schemes 2 and 4, when throughput is maximized, the EDBP is also optimized which corroborates previous result in figure 1.9 and figure 1.10. At high load, when we give more weight to the throughput (by increasing α), throughput and EDBP for scheme 1 (without prioritization) and scheme 3 (retransmission with less power) increase slightly. In figure 1.13(a and b) we plot the performance when $N = 5, m = 4$ and $\gamma_{th} = 10$dB at high load ($q_a = 0.9$), we note that, for schemes 1 and 3, the throughput improves when increasing α whereas it decreases a little for scheme 4. Although this hybrid optimization, a huge delay is seen for schemes 1 and 3 while schemes 2 and 4 keep constant value because of prioritization and large power randomization given to backlogged packets.

We note, through simulations, that our algorithms perform much better than the standard slotted aloha, and remark that schemes behavior is changing with the offered load. Yet, scheme 2 is generally more efficient at average load whereas scheme 4 seems to be the most performing at high and very high loads. These remarks motivate us to



(a) (b)

Figure 1.13: *Impact of multi-criteria factor α on aggregate throughput and expected delay of backlogged packets at heavy load $q_a = 0.9$.*

propose a load adaptive algorithm described as follows

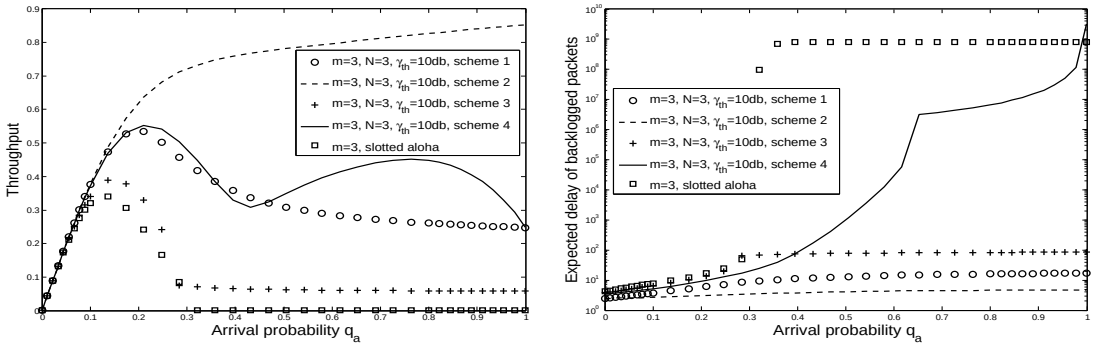
$$q^* = \max_{\text{scheme}=1\cdots4} \left(\max_{q} \; objective(q) \;\; s.t. \left\{ \begin{array}{l} \bar{\pi}(q) = \bar{\pi}(q) \cdot P(q), \\ \sum_{n=0}^{m} \pi_n(q) = 1, \\ \pi_n(q) \geq 0, \quad n = 0, 1, \cdots, m. \end{array} \right. \right). \qquad (1.26)$$

Our mixed algorithm can be implemented in practice as: The SP (Service Provider) is continuously monitoring the state of its radio system and then can estimate the instantaneous load as well as the number of communicating mobiles and the average number of backlogged packets. Here, one can use history statistics or any adapted algorithm such the Pseudo-Bayesian Algorithm described in Problem 4.2 in [25]. Having this information, the SP will be capable to judiciously switch its system to the best scheme and therefore take advantage from collected data. However, the base station should integrate a powerful software coupled with a high technology measurement devices.

1.5.3 Game problem

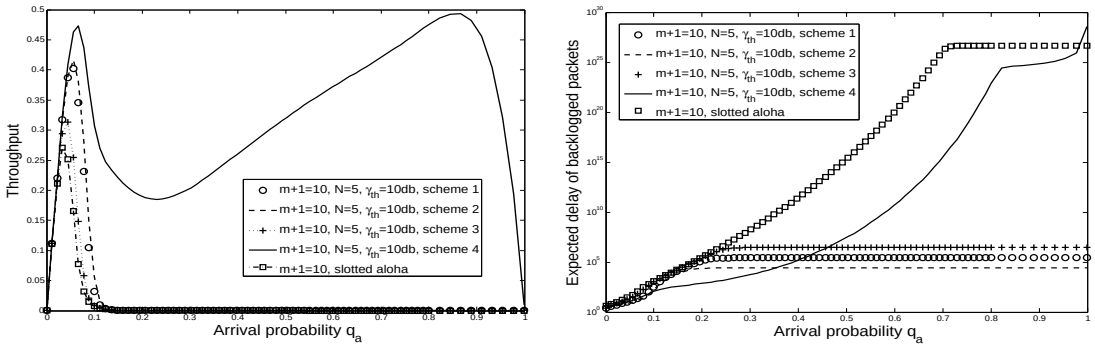
Throughput maximization

Figures 1.14 (a) and (b) show the variation of Nash equilibrium throughput and expected delay of backlogged packets versus the offered load for 4 mobiles. The throughput of slotted aloha increases with the arrival rate till achieving a maximum throughput of $thp^{max} = 0.34$ at $q_a \simeq 0.14$, then it decreases till getting annulled for $q_a \geq 0.32$. This throughput collapse causes a huge delay for slotted aloha. Scheme 1 and 3 have similar behavior but keep a constant value of throughput although the arrival rate continues to increase. We have respectively $thp_1^{max} = 0.38$ at $q_a \simeq 0.15$ and $thp_3^{max} = 0.53$ at $q_a \simeq 0.23$. Under these parameters values, an interesting feature occurs. In contrast to [8] where scheme 3 seems to achieve a performance close to the team problem, the throughput obtained for scheme 2 in the game setting matches closely the throughput



(a) (b)

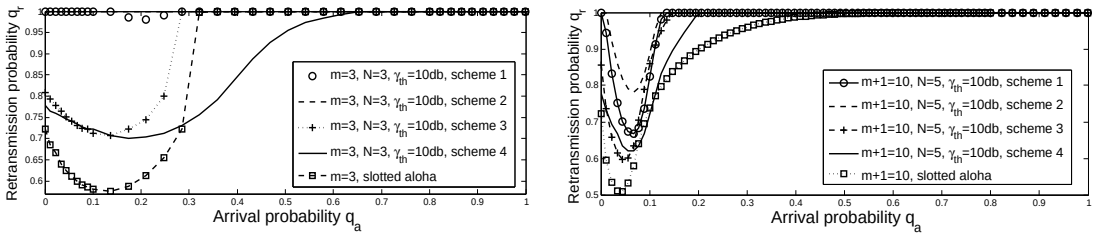
Figure 1.14: *Nash equilibrium throughput and EDBP for 4 mobiles when the payoff function is the individual throughput.*



(a) (b)

Figure 1.15: *Nash equilibrium throughput and EDBP for 10 mobiles when the payoff function is the individual throughput.*

of the team setting. Here, scheme 3 turns to perform less better than our schemes 1, 2 and 4. This shows that prioritizing backlogged packets rather than new arrival can be a good policy and constitutes a tradeoff to support services with different QoS requirements.



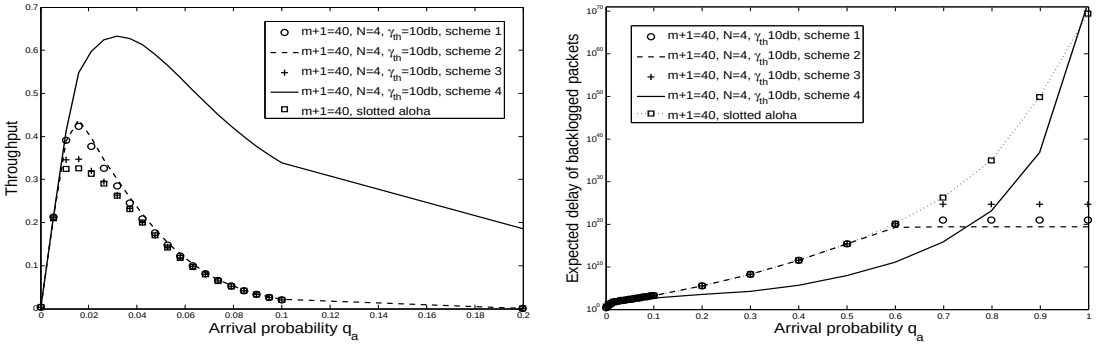
(a) (b)

Figure 1.16: *Retransmission rate (Nash equilibrium strategy) when maximizing individual throughput for 4 and 10 mobiles.*

Behavior under large mobile population - When the population size becomes large, we note that the throughput is only improved at low load for schemes 1-3 compared to slotted aloha, figure 1.15 (a) and figure 1.17. An interesting result is that scheme 4 outperforms all other schemes and does not suffer from throughput collapse in average and high loads, the throughput falls exponentially only at very high load. This behavior can be explained by the fact that new arrival are prioritized and can succeed their transmissions even if backlogged mobiles become very aggressive (retransmission probability close to 1). In term of EDBP, at low load scheme 4 seems to outperform all other schemes. Whereas it performs less better than schemes 1-3 when the load becomes average or high (figure 1.15 (b) and figure 1.17 (b)), indeed schemes 1-3 keep a constant value of EDBP (scheme 2 that prioritizes backlogged packets is the best). Plotting the retransmission probability in figure 1.16 shows why the throughput vanishes and the expected delay becomes huge. Contrary to the small population size case where mobiles become aggressive only at heavy load, here, users are very aggressive and transmit at probability close to 1 at average and high loads. One can note that schemes 1 and 2 are similar and provide same average throughput, however a slight difference is seen in terms of EDBP where scheme 2 performs a bit better. Scheme 3 seems to perform bad at large population but still slightly better than slotted aloha, in particular by providing a smaller expected delay.

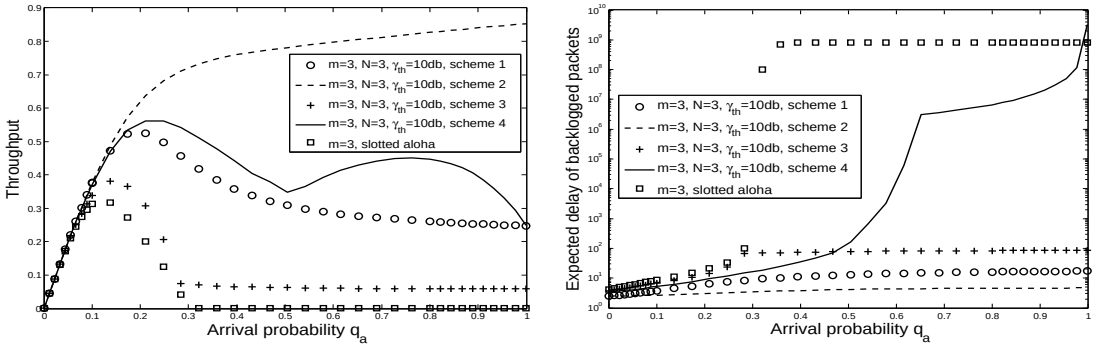
Delay minimization

When the expected delay of backlogged packets is minimized, see figure 1.18 and figure 1.19, we obtain similar behavior as when maximizing the individual throughput. We note that in some cases the throughput obtained when minimizing the delay is better than the corresponding one when throughput is maximized. This situation is similar to the prisoners dilemma, it shows in fact that Nash equilibrium is not efficient in some situations (very known result in the literature).



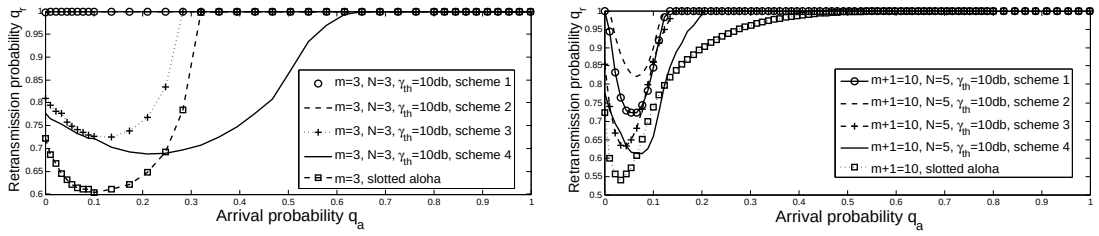
(a) (b)

Figure 1.17: *Nash equilibrium throughput and EDBP for 40 mobiles when maximizing individual throughput.*



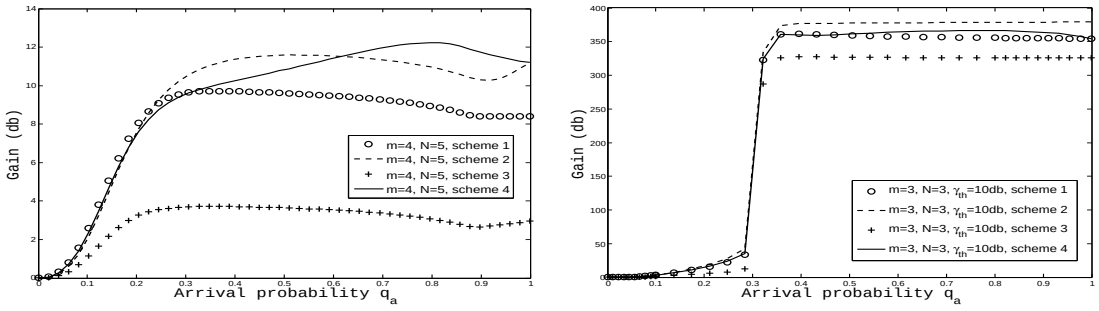
(a) (b)

Figure 1.18: *Throughput and delay when minimizing EDBP for 4 mobiles under the game problem.*



(a) (b)

Figure 1.19: *Retransmission probability when minimizing EDBP for 4 and 10 mobiles under the game problem.*



(a) (b)

Figure 1.20: *The throughput gain when the payoff function is the individual throughput for 4 mobiles, under both optimization (a) and game settings (b).*

1.6 Concluding remarks

We define the throughput gain by $G_i = 20\log(thp_i/thp_s)$, $i = 1 \cdots 4$, where thp_i is the throughput of scheme i and thp_s is the corresponding throughput of slotted aloha. Figures 1.20 (a) and (b) show that priority and power control improve considerably the system performance both in terms of throughput and delay. This gain becomes more important, in particular, at high load where congestion situation may be efficiently avoided or attenuated using our new schemes. The gain function looks like an S-shaped function and can be divided into two regions. The first region corresponds to low load where throughput is slightly improved and the average and high loads where throughput is considerably improved (throughput vanishes for slotted aloha under game problem) as well as expected delay.

Our schemes improve the channel utilization and allows the system to support several services requiring different QoS. The team formulation examines the case when coordination between users and control of their retransmission probabilities are possible. This needs some cognition capabilities and therefore requires to upgrade the already existent access points to support power control and coordination mechanisms. This solution consumes valuable bandwidth and requires advanced devices with complex computing capability, this is why we studied the game problem. Latter setting requires no coordination between the central access point and mobile users and represents well

the nature of aloha system; each mobile optimizes its own objective function according to the other users strategies and its collected information such as estimated value of the instantaneous backlog of the system. We found that performance indicators obtained at Nash equilibrium are, in general, less than those of the team setting. The experienced SINR is increasing as the number of available power levels is large as well as the distance between two consecutive power levels is larger. We also showed that introducing power differentiation and priority may improve the stability of the protocol.

However unfortunately, for large mobiles population and such slotted aloha, our schemes may also suffer from bi-stability and experience a throughput break down causing thereby a huge delay as the offered load becomes heavy. This is visible in particular under game setting. This is due to the fact that contention becomes very important and then mobiles are more aggressive. Controlling the retransmission probability q_r as in team setting seems to be an efficient way to stabilize the slotted aloha system and reach high performance. The idea behind this is to reduce the collision probability by decreasing the aggressiveness of communicating mobile users. Here, each mobile has to estimate the instantaneous backlog state n based on its feedback; There are many learning solutions, e.g. see [71] and [105], based on the following statement: increase q_r when experiencing an idle slot and decrease q_r when a collision occurs. One can also reduce the expected delay by limiting the number of retransmissions per packet, note that each packet is then definitively dropped after K retransmissions. The problem of throughput collapse in overload condition can also be resolved using a dynamic retransmission control which fine-tunes the retransmission limit according to the instantaneous backlog state. By means of game theory, one can also shrink mobiles aggressiveness by adding a (re)transmission cost to reduce access contention and then improve the stability of the system. Finally, it should be noted that aloha's characteristics are still not much different from those experienced today by CSMA/Wi-Fi, and similar contention-based systems that have no carrier sense capability. Indeed, there is a certain amount of inherent inefficiency in this family of systems. For instance 802.11b sees about a 2-4 Mbit/s effective throughput with a few stations talking, versus its theoretical maximum of 11 Mbit/s. These remarks show that our presented result can be extended to other contention systems.

In the next chapter, we reconsider a similar model, i.e. modeling slotted aloha as a stochastic game with partial information. However and without any loss of generality, we restrict to the case where all user transmit at the same transmit power. We introduce a priority/preference scheme by using defining two classes of users, then we derive a hierarchical model and analyze it by the mean of the Stackelberg solution concept.

Chapter 2

Sustaining Partial Cooperation in Hierarchical Wireless Collision Channels

Contents

2.1 Introduction

In the OSI hierarchical layers model, there is no doubt that the Medium Access Control layer plays the most important role in any wireless communication chain. ALOHA protocol [2] is the first random access protocol described in the literature for MAC layer. When a mobile terminal has a newly arrived packet, it transmits it sporadically on the common channel. If there is no other mobile accessing the channel at the same time, the packet is correctly received, otherwise, all packets sent simultaneously will collide and become backlogged. A backlogged packet needs to be retransmitted after some random time. This protocol is pretty simple to implement, but suffers from high amount of collisions. This results in low channel utilization (at most $1/2e$ of the available bandwidth with infinite users population and Poisson arrivals). An improved version of aloha is slotted aloha [125], here the time is divided into equal time units, say slots. At the beginning of each time slot, a packet may be transmitted, and at the end of the time

interval, the sources get the feedback on whether there was zero, one or more transmissions (collision) during the time slot. Packets that are involved in a collision are backlogged and are scheduled for retransmission after a random number of time slots. The achievable throughput is then seen to be increased to $1/e$ [25]. One note that due to synchronization constraint, packets overlap is now either completely or not at all. Aloha-like protocols are still nowadays, used in cellular telephone systems (e.g. slotted aloha is used in UMTS for dedicated channel reservation) and satellite networks for the sporadic transfer of data packets.

When multiple users share a common channel and contend for access, a typical conflict problem arises. Recently, the selfish behavior of users in MAC protocols has been widely analyzed using game theory with all its powerful solution concepts. It was shown in [13], [8], [44] and [115] that the users selfish behavior likely leads to a network collapse, where a typical prisoners dilemma situation occurs. This illustrates, in fact, that Nash equilibrium is not efficient in some situations and more appropriate solution concepts need to be considered. This way, full system utilization requires coordination among users using explicit message exchanges or presence of an arbitration mechanism [10], which may be impractical given the distributed nature and arbitrary topology changes of wireless collision channels. To achieve a better performance without coordination schemes, users need to sustain cooperation. It is promising to introduce a set of special users whose mission is to provide incentives for other users to behave cooperatively, this mechanism may limit the aggressiveness level (access to the channel).

Despite of the bounty of works and efforts investigated in analyzing aloha-like protocols, Aloha is still an ideal tool to understand wireless behavior and users selfishness. This will lead to design more robust protocols in the future. The main objective of this work is to improve the performance of slotted aloha-like protocols by introducing a hierarchy mechanism between users. The main principle of our new scheme is the following: users are split into two different groups, the leaders and the followers. The game is played sequentially such that the followers play knowing the decision of the leaders. Moreover, the leaders choose their strategy knowing that the followers will play depending on their chosen strategy (best response). This class of bi-level games is called a Stackelberg game [11]. It is widely used in economy, and recently used for studying performance of communication systems [28]. Specifically, the application of that kind of mechanism in a power control game has proved many promising improvements compared to the Nash game [72]. Authors in [115] analyzed collision channel access game and proposed a methodology that transforms the noncooperative game into a Stackelberg game, this allows to overcome the deficiency of the Nash equilibria of the original game. The authors associate to an additional user, called "Manager" or also leader, an administrating role based on the intervention function. This function is simply the effective transmission rate of the manager. It is used to regulate the access probabilities of competitors users and make them transmitting at some target transmission rate vector defined by the manager.

Conceptually, the system is split into two non-cooperative and coupled sub-games : The first game concerns the set of leaders and the second implies the followers group. But, the two sub-games are implicitly interdependent; The followers game is played after the leaders game, and its outcome depends on the strategies profile of that latter. In our setting, the base station has an important role to play. Indeed, the first role of the base station is to decide the set of users that should be leaders, this decision may be taken based on the QoS of users (e.g., users requiring high rate or having hard constraints on throughput/delay would be selected as leaders). Then, the base station computes the optimal decision vector of the leaders and communicates it to the followers group. Using this hierarchical mechanism, we observe an improvement of the global throughput of the slotted aloha protocol. It is also possible to improve the performance by introducing an admission control mechanism. Yet, we observe that the more the system is saturated (the probability of new arrival is important), the more users become aggressive, and then the performance of the system is reduced.

2.2 Model and problem formulation

We consider a collision channel used by one central receiver (e.g., base station) and m users without buffer, similar to that one user in Chapter 1. The arrival flow of packets to source i follows a Bernoulli process with parameter q_a (i.e., at each time slot, there is a probability q_a of a new arrival at a source, and all arrivals are independent). As long as there is a packet at a source (i.e., as long as it is not successfully transmitted) new packets to that source are blocked and lost[1]. The arrival processes are all independent for all sources. Whereas a backlogged packet at source i is retransmitted with probability q_r^i. We shall restrict in our game problem to simple policies in which q_r^i does not change in time. Since sources are symmetric, we shall further restrict to finding a symmetric Nash equilibrium, that is retransmission probabilities q_r^i that do not depend on i . Indeed, slotted aloha protocol is usually built for medium access or sporadic data transfer, the assumption of symmetric sources is then justified

2.3 Overview on the non-cooperative game

We formulate here the distributed scheme of slotted aloha using a game theoretic formulation. For a given policy vector $\mathbf{q_r}$ of retransmission probabilities for all users (whose jth entry is q_r^j). Define $([\mathbf{q_r}]^{-i}, \hat{q}_r^i)$ to be a retransmission policy where user j retransmits at the current slot with probability q_r^j for all $j \neq i$ and where user i retransmits with probability $\hat{q}_r^i$. Each user i seeks to maximize his own throughput T_i

[1]Considering the number of packets in the system, this assumption is equivalent to say that a source does not generate new packets as long as the previous packet is not successfully transmitted.

considered as the objective function. We summarize the assumptions of our contention model as

- A fixed set of m bufferless users interact over a single collision channel.
- Time is divided into multiple equal slots, and slots are synchronized. Transmission feedback (success or collision) are received in the end of the current slot.
- Saturation assumption is relaxed; Packets arrive from higher layers of source node i following a Bernoulli process with parameter q_a.
- Each user i retransmits its packets with probability q_r in every slot.
- (Re)Transmissions are cost free.
- The individual throughput (or alternatively the expected delay) is considered as the objective function to maximize.

The problem we address is then to find a symmetric equilibrium policy $\mathbf{q_r^*} = (q_r, q_r, .., q_r)$ such that for any user i and any retransmission probability q_r^i for that user, we obtain

$$T_i\left(\mathbf{q_r^*}\right) \geq T_i\left(\left[\mathbf{q_r^*}\right]^{-i}, q_r^i\right) \qquad \text{for all } q_r^i \in [\epsilon,\, 1]. \tag{2.1}$$

The point $\mathbf{q_r^*}$ is the Nash equilibrium of the non-cooperative game. We consider the interval $[\epsilon,\, 1]$ instead of the non closed interval $(0,\, 1]$ to guarantee the existence of a solution. Since we restrict to the symmetric case $\mathbf{q_r^*}$, we shall also identify it (with some abuse of notation) with the actual transmission probability (which is the same for all users). Next we show how to obtain an equilibrium policy. We first note that due to symmetry, to see whether $\mathbf{q_r^*}$ is an equilibrium it suffices to check (2.1) for a single player. For ease of notation, we assume that the first $m-1$ users retransmit with a given probability $\mathbf{q_r}^{-(m)} = (q_r^o, .., q_r^o)$ and user m retransmits with probability $q_r^{(m)}$. Define the set

$$\mathcal{Q}^m\left(\mathbf{q_r^o}\right) = \underset{q_r^{(m)} \in [\epsilon,1]}{\operatorname{argmax}}\, T_m\left(\left[\mathbf{q_r^o}\right]^{-(m)}, q_r^{(m)}\right),$$

where $\mathbf{q_r^o}$ denotes (with some abuse of notation) the policy where all users retransmit with probability q_r^o, and where the maximization is taken with respect to $q_r^{(m)}$. Then q_r^* is a symmetric equilibrium if

$$q_r^* \in \mathcal{Q}_r^m\left(q_r^*\right).$$

To compute the throughput of tagged user m noted by $T_m([\mathbf{q_r^o}]^{-(m)}, q_r^{(m)})$, we introduce a Markov chain with a two dimensional state. The first state component corresponds to the number of backlogged packets among the users 1,...,$m-1$, and the second component is the number of backlogged packets (either 1 or 0) of user m. The transition probabilities are given by

$$P_{(n,i),(n+k,j)}\left(q_r^o, q_r^{(m)}\right) = \begin{cases} Q_a(k,n)\,(1-q_a)^{1-j}\,(q_a)^{j-i}, & i \le j \;\Big\}\, 2 \le k \le m-n-1 \\ Q_a(1,n)\left[1 - Q_r(0,n)\left(1-q_r^{(m)}\right)\right], & i=j=1 \\ Q_a(1,n)[1-Q_r(0,n)](1-q_a), & i=j=0 \\ Q_a(1,n)q_a, & i=0, j=1 \end{cases} \Big\}\; k=1 \\ \begin{cases} (1-q_a)Z + q_a Q_a(0,n)Q_r(0,n), & i=j=0 \\ q_a Q_a(0,n)[1-Q_r(0,n)], & i=0, j=1 \\ q_r^{(m)} Q_a(0,n)Q_r(0.n), & i=1, j=0 \\ q_r(1-Q_r(0,n))Q_a(0,n) + \left(1-q_r^{(m)}\right)Z, & i=j=1 \end{cases} \Big\}\; k=0 \\ \begin{cases} Q_a(0,n)Q_r(1,n)\left(1-q_r^{(m)}\right), & i=j=1 \\ Q_a(0,n)Q_r(1,n)(1-q_a), & i=j=0 \end{cases} \Big\}\; k=-1 \\ 0, \qquad \text{otherwise.} \end{cases}$$

where $Z = Q_a(1,n)Q_r(0,n) + Q_a(0,n)\,[1-Q_r(1,n)]$ and where Q_a and Q_r are given by (1.3) and (1.2), respectively (with q_r^o replacing q_r). Let $\pi_{n,n'}\left([q_r^o]^{-(m)}, q_r^{(m)}\right)$ be the steady state of the system. The throughput of the m-th user, defined as the sample average number of packets successfully transmitted, is given by

$$T_m\left([\mathbf{q_r^o}]^{-(m)}, q_r^{(m)}\right) = q_a \sum_{n=0}^{m-1} \pi_{n,0}\left([q_r^o]^{-(m)}, q_r^{(m)}\right). \tag{2.2}$$

As mentioned in [13] and [44], the Nash equilibrium (NE) in such games can be inefficient and may provide bad performance. Indeed, we see that the users become more and more aggressive as the arrival probability increases which explains the dramatic decrease in the system's aggregate throughput. Moreover, the equilibrium retransmission quickly increases to 1 when the number of users increases. We note that a similar aggressive behavior at equilibrium has been observed in [42] in the context of flow control by several competing users that share a common drop tail buffer. However in that context, the most aggressive behavior (of transmission at maximum rate) is the "equilibrium" solution for *any arrival rate*, and not just at high rates as in our case. We may thus wonder why retransmission probabilities of 1 are not an equilibrium in our slotted Aloha problem (in the case of light traffic). An intuitive reason could be that if a mobile deviates and retransmits with probability one, (while other continue to retransmit with the equilibrium probability $q^* < 1$) the congestion level in the system (i.e., the number of backlogged users) increases; this provokes more retransmissions from other users which then causes sufficiently more collisions of packets from the deviating mobile so as to cause a decrease in its throughput.

2.4 Virtual controller and protocol design

The Stackelberg formulation arises naturally in many context of practical interests such as service differentiation or traffic priority. In order to reduce the congestion of slotted aloha operated networks, we propose a simple mechanism based on introducing some kind of hierarchy/differentiation between users. Then, we develop a methodology that transforms the original non-cooperative game into a hierarchical Stackelberg game. We start in this section by describing the Stackelberg game with one single leader and analyze the general case where several leaders may exist in the next section. Clearly, to apply this approach to the medium access problem, signals need to be conveyed from a mediator to all users to distribute leader/follower roles as well as making communication between the two subgames (leaders game and followers game), and users need to know the correct meanings of the observed signals as well. We focus now on the protocol design and consider a single leader with several followers. Then, we deal with the engineering issue of our new protocol and discuss how it can be exploited in real systems.

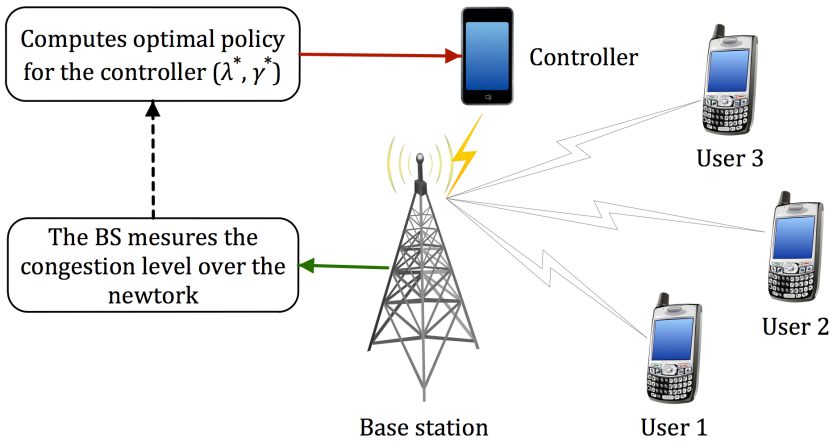


Figure 2.1: *Hierarchical slotted aloha protocol design.*

For that aim, we introduce a special user called *Controller* that will play the role of the leader. It is a fictitious/virtual user that can be any additional user or the base station itself. The main role of the controller is to control the retransmission probabilities of m concurrent users also considered as followers. With judicious choice of control decision, the controller can shape the incentives of the followers in such a way that their selfish behavior results in nearly cooperative outcomes. Naturally and based on transmissions history, the base station can measure the aggregate success rate over the network, see figure 2.1. This information is used to decide if the followers should keep their actual retransmission policy or should change it according to the congestion level of the network. The controller knows the response of the follower to their decisions correctly, i.e., he knows that a Nash equilibrium between followers will be played. Let λ and γ denote, respectively, the arrival probability and the retransmission probability of the

controller, theses couple of parameters will be referred as the intervention level of the controller. Therefore, having the network congestion level, the base station computes the optimal intervention level (λ^*, γ^*) of the controller. Then the variation of (λ and γ) will induce an immediate reaction by the group of followers and then a variation on their retransmission probabilities in order to decrease the system congestion and then to improve the average throughput. In other words, the transmissions of the controller only makes the channel lossy but it does not provide incentives for users not to choose the maximum possible transmission probability, i.e., transmission w.p. 1. Hence, in order to provide an incentive to choose a smaller transmission probability, the controller needs to vary his intervention level depending on the transmission probabilities of the followers. Let $\mathbf{q}(\lambda, \gamma)$ be the response policy of the followers when the controller parameters are (λ, γ). We define the utility function of the controller as the aggregate throughput over the network, i.e., $T(\lambda, \gamma, \mathbf{q}(\lambda, \gamma)) = \sum_{i=1}^{m} T_i(\lambda, \gamma, \mathbf{q}(\lambda, \gamma))$. The controller problem can be expressed as the following

Controller problem:

$$(\lambda^*, \gamma^*) \in \underset{\lambda,\gamma}{\operatorname{argmax}}\, T(\lambda, \gamma, \mathbf{q}(\lambda, \gamma)). \tag{2.3}$$

Since we restrict to symmetric equilibria, let us define $\bar{q} = (q, q, \cdots, q)$ as the transmission policy where the followers tagged from 1 to $m-1$ transmit with same probability q. Then the problem of the tagged follower m is :

Follower problem:

$$q_r^* \in \underset{p^f}{\operatorname{argmax}}\, T_i\left(\lambda^*, \gamma^*, q, p^f\right). \tag{2.4}$$

It follows that $(\lambda^*, \gamma^*, q^*)$ is a symmetric Stackelberg equilibrium iff
1) $(\lambda^*, \gamma^*) \in \underset{\lambda,\gamma}{\operatorname{argmax}}\, T(\lambda, \gamma, q^*)$, and
2) $q^* \in \underset{p^f}{\operatorname{argmax}}\, T_i\left(\lambda^*, \gamma^*, q^*, p^f\right)$.

For performance metrics derivation, we introduce a 3-dimensional Markov chain. The state of the system is a vector (n, a, i) where n is the number of backlogged packets among first $m-1$ sources, a is the number of backlogged packets of the m-th user (either 0 or 1) and i is the number of backlogged packets of the controller (either 0 or 1). Let us denote by $P_{(n,a,i),(n+k,b,j)}\left(q, p^f, \gamma, \lambda\right)$ the transition probability from state (n, a, i) to state $(n+k, b, j)$, see Appendix I for detailed derivation of the transition matrix. Let $\pi_{n,a,i}\left(\gamma, \lambda, q, p^f\right)$ denotes the steady probability that the Markovian system is in state (n, a, i). The average throughput of source m is given by

$$T_m\left(\gamma, \lambda, q, p^f\right) = q_a \sum_{n=0}^{m-1} \sum_{i=0}^{1} \pi_{n,0,i}\left(\gamma, \lambda, q, p^f\right). \tag{2.5}$$

Similarly, the average number of backlogged packets at source m is

$$S_m(\gamma, \lambda, q, p^f) = \sum_{n=0}^{m-1} \sum_{i=0}^{1} \pi_{n,1,i}(\gamma, \lambda, q, p^f). \tag{2.6}$$

The average number of packets in the system, at the beginning of any slot, is exactly $T_m(\cdot) + S_m(\cdot)$. It follows by Little's result that the expected delay of transmitted packets ca be written as

$$EDTP_m = 1 + \frac{S_f(\gamma, \lambda, q, p^f)}{T_f(\gamma, \lambda, q, p^f)}. \tag{2.7}$$

Altruistic behavior of the virtual controller : As common in some wireless settings, e.g., tactical mobile ad hoc networks (MANETs), network nodes may engage in cooperative, coalitional, or simply altruistic behavior with respect to some of their peers. Altruistic action can be for the purposed of routing, medium access, etc. According to the objective function, we note that the virtual controller maximizes the utility of concurrent users in the network, which is clearly a purely altruistic behavior. If we consider that the controller is built on the base station itself, then the maximization of aggregate throughput is plausible. Otherwise, it will be more suitable for the controller to maximize a linear combination of its own utility $T_c(\cdot)$ and the aggregate utility of other users $T(\cdot)$. We define η as the altruism degree, then the controller problem becomes

$$(\lambda^*, \gamma^*) \in \underset{\lambda,\gamma}{\operatorname{argmax}} \left[(1-\eta)\, T_c\left(\lambda, \gamma, \mathbf{q}\left(\lambda, \gamma\right)\right) + \eta T\left(\lambda, \gamma, \mathbf{q}\left(\lambda, \gamma\right)\right) \right]. \tag{2.8}$$

Note that $\eta = 0$ corresponds to a purely selfish controller (this case will be addressed in next section) and $\eta = 1$ corresponds to a purely altruistic controller which is analyzed above. Furthermore, one can fine-tune the degree of altruism for better performances and potential tradeoff between individual and overall performances.

2.5 Hierarchical game formulation of slotted aloha

2.5.1 Markov chain and Stackelberg solution

We extend in this section our approach to the general case where several leaders may exist. It is plausible to consider that the base station only broadcasts common messages and the corresponding amount of additional signaling is reasonable. We propose a Stackelberg formulation of the slotted aloha where m^l users are chosen as the leaders whereas m^f others act as the followers. We label each leader with an integer number $i = 1, .., m^l$. Similarly, each follower is labeled with an integer number $j = 1, .., m^f$. Conceptually, the followers play a noncooperative game with each other, given the decision of the leaders group.

Let $\mathcal{N}^*(\bar{\mathbf{q}}^l)$ be the set of Nash equilibria of the followers subgame given the leaders strategy $\bar{\mathbf{q}}^l = (q_1^l, .., q_{m^l}^l)$, where q_i^l is the retransmission probability of the leader i. In other words, the leaders maximize their utility function which depends on the Nash equilibrium $\mathbf{q}^f \in \mathcal{N}^*(\bar{\mathbf{q}}^l)$ of the followers. The definition of a Stackelberg equilibrium in this context is the following.

Definition 2.5.1.1 (Stackelberg equilibrium)**.** *A policies vector* $\mathbf{q} = (\mathbf{q^l}, \mathbf{q^f})$ *is called a Stackelberg equilibrium (SE) if* $\mathbf{q}^f \in \mathcal{N}^*(\mathbf{q}^l)$ *and the strategy profile* $\mathbf{q}^l$ *is the Nash equilibrium for the group of leaders, i.e,* q_i^l *maximizes the utility function of user i for* $i = 1, .., m^l$*. Namely*

$$q_i^l \in \underset{\tilde{q}_i}{\operatorname{argmax}}\, T_i\left(\tilde{q}_i, \left(\mathbf{q}^l\right)^{(-i)}, \mathbf{q}^f\left(\tilde{q}_i, (\mathbf{q}^l)^{(-i)}\right)\right).$$

Existence of a solution – The only point where the associated Markov chain does not have a single stationary distribution is when the retransmissions vector is such that $(\bar{\mathbf{q}}^l, \bar{\mathbf{q}}^f) = 0$. The steady state probabilities are continuous over $0 < q^l \leq 1$ and $0 < q^f \leq 1$. Since this is not a closed domain, a solution needs not exist. However, as we restrict to the closed domain $(q^l, q^f) \in [\epsilon, 1]^2$ where $\epsilon > 0$, an optimal solution indeed exists. Therefore for any $\delta > 0$, there exists some $(q^{l*}, q^{f*}) > 0$ which is δ-optimal. $(q^l, q^f) > 0$ is said to be δ-optimal for the throughput maximization if it satisfies $T(q^{l*}, q^{f*}) \geq T\left(q^l, q^f\right) + \delta$ for all $(q^l, q^f) \in [\epsilon, 1]^2$. A similar definition holds for any objective function (e.g., delay minimization). For detailed discussion, this optimality concept is discussed in [13] and [8].

Since we restrict to symmetric $(\mathbf{q}^l, \mathbf{q}^f)$, we shall also identify $\mathbf{q}^l$ (resp. $\mathbf{q}^f$) with the actual transmission probability which is the same for all leaders (resp. followers). Next we show how to obtain an equilibrium policy. We first note that due to symmetry, to see whether a retransmission policy is a Stackelberg equilibrium or not, it suffices to check for a single user among followers and a single user among leaders. We shall thus assume that the first $m^f - 1$ followers retransmit with a given probability $(\mathbf{q}^f)^{-(m^f)} = (q^f, .., q^f)$ whereas tagged user m^f retransmits with probability p^f. Similarly, we assume that the first $m^l - 1$ leaders retransmit with a given probability $(\mathbf{q}^l)^{-(m^l)} = (q^l, .., q^l)$ whereas tagged leader m^l retransmits with probability p^l.

Define the following sets (best responses of tagged leader m^l and follower m^f)

$$\mathcal{Q}_f^{(m^f)}\left(q^f, q^l, p^l\right) = \underset{\tilde{p}^f \in [\epsilon,1]}{\operatorname{argmax}}\, T_f\left(\left[q^f, q^l, p^l, \tilde{p}^f\right]\right), \tag{2.9}$$

$$\mathcal{Q}_l^{(m^l)}(q^f, q^l, p^f) = \underset{\tilde{p}^l \in [\epsilon,1]}{\operatorname{argmax}}\, T_l\left(\left[q^f, q^l, p^f, \tilde{p}^l\right]\right). \tag{2.10}$$

Then the policy profile $(\bar{\mathbf{q}}^l, \bar{\mathbf{q}}^f)$ is a symmetric Stackelberg equilibrium iff

$$\bar{q}^f \in \mathcal{Q}_f^{(m^f)}\left(\bar{q}^f, \bar{q}^l, \bar{q}^l\right) \text{ and } \bar{q}^l \in \mathcal{Q}_l^{(m^l)}\left(\bar{q}^f, \bar{q}^l, \bar{q}^f\right). \tag{2.11}$$

To compute the throughput of leaders and followers, we introduce a Markov chain P with four dimensions state. The first state component corresponds to the number of backlogged packets among the $m^l - 1$ leaders, and the second component is the number of backlogged packets among the $m^f - 1$ followers, the third component is the number of backlogged packets (either 1 or 0) of the tagged leader m^f and the last component is the number of backlogged packets (either 1 or 0) of the tagged follower m^f. Replacing q_a and q_r in expressions (1.3) and (1.2) (respectively) by (q_a^l, q_r^l) and (q_a^f, q_r^f), we get the arrival and retransmission flows for both leaders and followers groups. The transition matrix M is an $m^l \times m^f \times 2 \times 2$ square matrix. It is composed of $m^l \times m^l$ matrix blocks of $4m^f \times 4m^f$ each one and has the following form

$$M = \begin{pmatrix} A_{0,0} & \cdots & A_{0,m^l-1} \\ \vdots & \ddots & \vdots \\ A_{m^l-1,0} & \cdots & A_{m^l-1,m^l-1} \end{pmatrix},$$

where each matrix block $A_{n,n+k}$ is a $m^f \times m^f$-dimensional matrix of blocks given by

$$A_{n,n+k} = \begin{pmatrix} B_{0,0} & \cdots & B_{0,m^f-1} \\ \vdots & \ddots & \vdots \\ B_{(m^-1,0} & \cdots & B_{m^f-1,m^f-1} \end{pmatrix}.$$

A matrix block $B_{(n,n')(n+k,n'+k')}$ is a 2×2-dimensional matrix of blocks, it is expressed by

$$B_{(n,n')(n+k,n'+k')} = \begin{pmatrix} C_{0,0} & \cdots & C_{0,1} \\ C_{(1,0} & \cdots & C_{1,1} \end{pmatrix}.$$

The fourth abstraction level of the matrix M is a 2×2-dimensional matrix. Block $C_{i,j}$ is of the form

$$C_{(n,n',i)(n+k,n'+k',j)} = \begin{pmatrix} P_{(n,n',i,0),(n+k,n'+k',j,0)} & P_{(n,n',i,0),(n+k,n'+k',j,1)} \\ P_{(n,n',i,0),(n+k,n'+k',j,1)} & P_{(n,n',i,1),(n+k,n'+k',j,1)} \end{pmatrix},$$

where $P_{(n,n',i,a),(n+k,n'+k',j,b)}$ is the probability that the system transits from the actual state vector (n, n', i, a) to the state $(n + k, n' + k', j, b)$. The transition probabilities are provided in appendix E.

2.5.2 Individual performance metrics

We now look at the calculation of the average throughput and the expected delay of backlogged packets for both the tagged leader m^l and the tagged follower m^f. Let $\pi_{n,n',i,a}(\bar{q}^l, \bar{q}^f, p^l, p^f)$ denotes the equilibrium probability vector when the system is in state (n, n', i, a). We define the throughput of a node as the sample average of the number of packets that are successfully transmitted by this node. Based on the balance rate

equation, i.e., at equilibrium the arrival and departure rates of each node are balanced, the average throughput of the tagged leader m^l is given by

$$T_l(\bar{q}^l, p^l, \bar{q}^f, p^f) = q_a^l \sum_{n=0}^{m^l-1} \sum_{n'=0}^{m^f-1} \sum_{a=0}^{1} \pi_{n,n',0,a}(\bar{q}^l, p^l, \bar{q}^f, p^f). \tag{2.12}$$

Similarly, the average throughput of the tagged follower m^f can be derived as

$$T_f(\bar{q}^l, p^l, \bar{q}^f, p^f) = q_a^f \sum_{n=0}^{m^l-1} \sum_{n'=0}^{m^f-1} \sum_{i=0}^{1} \pi_{n,n',i,0}(\bar{q}^l, p^l, \bar{q}^f, p^f). \tag{2.13}$$

The average number of backlogged packets for the tagged leader node m^l and the tagged follower node m^f are, respectively, given by

$$\begin{cases} S_l(\bar{q}^l, p^l, \bar{q}^f, p^f) = \sum_{n=0}^{m^l-1} \sum_{n'=0}^{m^f-1} \sum_{a=0}^{1} \pi_{n,n',1,a}(\bar{q}^l, p^l, \bar{q}^f, p^f), \\ S_f(\bar{q}^l, p^l, \bar{q}^f, p^f) = \sum_{n=0}^{m^l-1} \sum_{n'=0}^{m^f-1} \sum_{i=0}^{1} \pi_{n,n',i,1}(\bar{q}^l, p^l, \bar{q}^f, p^f). \end{cases} \tag{2.14}$$

By Little's formula, we easily obtain the corresponding expected delay of transmitted packets (at tagged nodes m^l and m^f respectively) given by

$$EDTP_l = 1 + \frac{S_l(\bar{q}^l, p^l, \bar{q}^f, p^f)}{T_l(\bar{q}^l, p^l, \bar{q}^f, p^f)} \quad \text{and} \quad EDTP_f = 1 + \frac{S_f(\bar{q}^l, p^l, \bar{q}^f, p^f)}{T_f(\bar{q}^l, p^l, \bar{q}^f, p^f)}. \tag{2.15}$$

Another important objective function is to evaluate the performance of our new hierarchical scheme using only backlogged packets, i.e., packets which become backlogged due to simultaneous transmissions. This gives a promising insight to evaluate the behavior of the system to serve backlogged packets. In particular, this will show the ability to support delay-constrained services (Voice, VoD, Streaming, ...). The average throughput of a backlogged packet at the tagged leader source, respectively at the tagged follower source, is given by equations (2.16) and (2.17).

Therefore, the expected delay of backlogged packets (the average time from the instant the packet reaches the MAC layer of the source to the instant it is successfully transmitted) at the tagged leader source m^l, respectively at the tagged follower source m^f, is

$$EDBP_l = 1 + \frac{S_l\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right)}{T_l^b\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right)} \quad \text{and} \quad EDBP_f = 1 + \frac{S_f\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right)}{T_f^b\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right)}. \tag{2.18}$$

$$T_l^b(\bar{q}^l, p^l, \bar{q}^f, p^f) = \sum_{n_1=0}^{m^l-1}\sum_{n_2=0}^{m^l-1}\sum_{n_1'=0}^{m^f-1}\sum_{n_2'=0}^{m^f-1}\sum_{i=0}^{1}\sum_{j=0}^{1} P_{(n_1,n_1',0,a),(n_2,n_2',1,b)}\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right)\pi_{n_1,n_1',0,a}\left(\bar{q}^l, \bar{q}^f, p^l, p^f\right) \tag{2.16}$$

$$T_f^b(\bar{q}^l, p^l, \bar{q}^f, p^f) = \sum_{n_1=0}^{m^l-1}\sum_{n_2=0}^{m^l-1}\sum_{n_1'=0}^{m^f-1}\sum_{n_2'=0}^{m^f-1}\sum_{i=0}^{1}\sum_{j=0}^{1} P_{(n_1,n_1',i,0),(n_2,n_2',j,1)}\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right)\pi_{n_1,n_1',i,0}\left(\bar{q}^l, \bar{q}^f, p^l, p^f\right) \tag{2.17}$$

2.6 Numerical investigation

We turn now to investigate the performance of our proposed hierarchical slotted aloha in term of average throughput, retransmission probabilities (equilibrium policy) and expected delay of backlogged packets. We illustrate the impact of introducing a virtual controller in the network and depict the performance metrics. Later we discuss how the leader/follower role should be distributed and what is the optimal size of each group of users. We consider the throughput and delay as objective functions, we obtain similar trends with better performances of our scheme compared to standard slotted aloha medium access method.

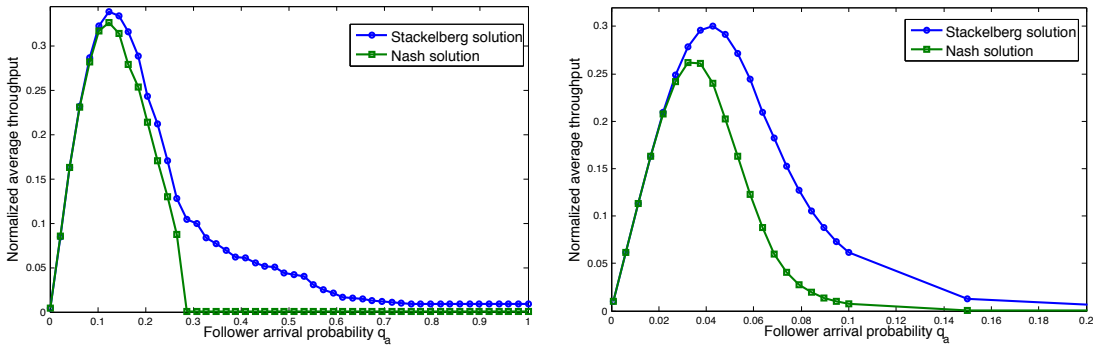
2.6.1 Symmetric Stackelberg solution

We recall that we restrict in this paper to the symmetric Nash and Stackleberg solutions. Indeed, if users are asymmetric then the computation of the system steady state becomes intractable and consequently the evaluation of throughput and delay will not be possible. A symmetric Stackelberg equilibrium is computed in three steps as follows: 1) The set of leaders deicide to select a retransmission profile. 2) The base station broadcasts the transmission policy of the leaders to the followers group. The best response, which results naturally in a symmetric Nash equilibrium, of the followers is then computed. 3) Getting feedback of followers decision, the leaders group checks their selected profile; if it is not a symmetric Nash equilibrium, they should update their strategies profile by unilateral deviation till getting absorbed by a symmetric equilibrium.

Remark 2.6.1.1. *Studying an asymmetric network numerically requires one to consider all possible combinations of the network parameters. Since the degree of freedom (the parameters to choose) are usually very large in asymmetric networks, such a numerical study is not carried out generally.*

2.6.2 Sustaining cooperation by introducing a virtual controller

We perform here simulations with $m = 4$ and 10 users, and a central receiver which itself plays the role of a virtual controller. Each user maximizes its own throughput whereas the virtual controller is purely altruistic and maximizes the aggregate throughput. We depict in figure 2.2 and figure 2.3, respectively, the aggregate throughput and the expected delay of hierarchical slotted aloha taking slotted aloha as reference. We note that either in terms of average throughput or expected delay of backlogged packets, hierarchical scheme outperforms the standard slotted aloha for all loads. An important result is that the expected delay is efficiently reduced and tends to remain constant for all loads. Another important result is the resolution of the throughput collapse. At Stackelberg equilibrium, we note that users are less aggressive. Indeed, at low load, our scheme behaves as slotted aloha, but at average and high loads, the controller make the channel lossy and retransmits w.p. 1, see figure 2.4. A surprising behavior is that when the users experience a bad channel (mainly due to regulating transmissions of the controller), they reduce their retransmission probabilities to 0.5 independently of the number of concurrent users. This has a positive impact and explains, first the throughput improvement, and second the efficient delay limitation. This is not the case for standard slotted aloha where an exponential increase of average delay is obtained due to bad collisions resolution. The optimal arrival probability of the controller is depicted in figure 2.5, it is clear that a saturated controller ($\lambda = 1$) is not optimal. Indeed, we note that when users tends to become aggressive and then transmit w.p. 1, the controller turns to transmit w.p. 1. which results in a zero throughput and infinite delay.

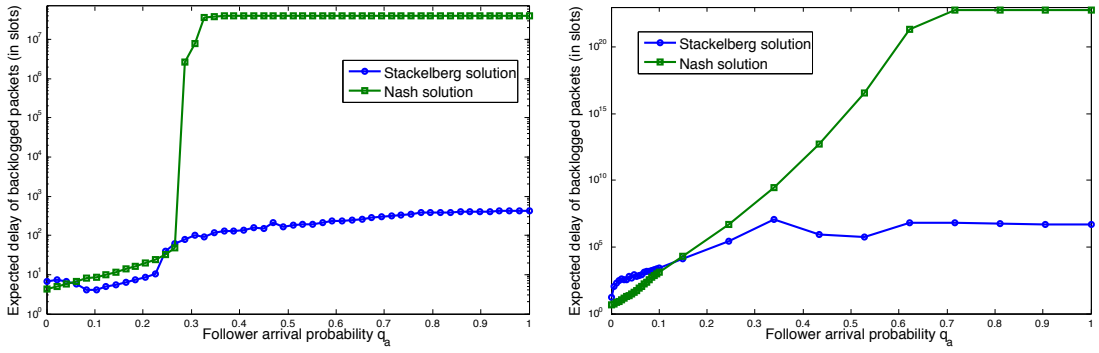


(a) (b)

Figure 2.2: *Aggregate throughput versus the arrival probability q_a of the followers, under intervention of a virtual controller for 4 (a) and 10 (b) users.*

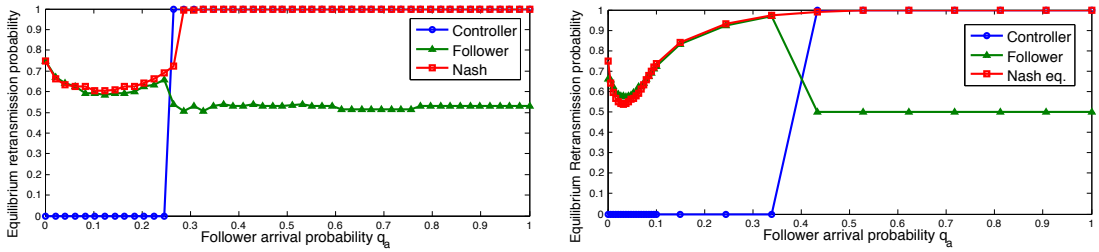
2.6.3 Stackelberg equilibrium Vs Nash equilibrium

We plot in figure 2.7 the individual throughput of a leader (figure a) and a follower (figure b) for different repartition of the users in each group. We note that many properties are fulfilled.



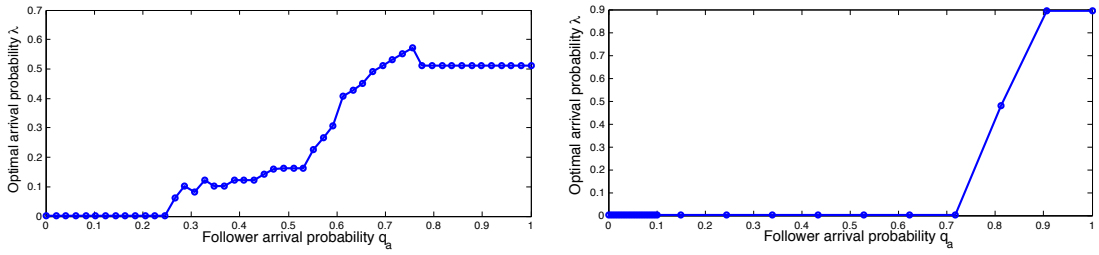
(a) (b)

Figure 2.3: *Expected delay of backlogged packets versus the arrival probability q_a of the followers, under intervention of a virtual controller for 4 (a) and 10 (b) users.*



(a) (b)

Figure 2.4: *Retransmission probability at Stackelberg equilibrium versus the followers arrival probability q_a of the followers for 4 (a) and 10 (b) users.*



(a) (b)

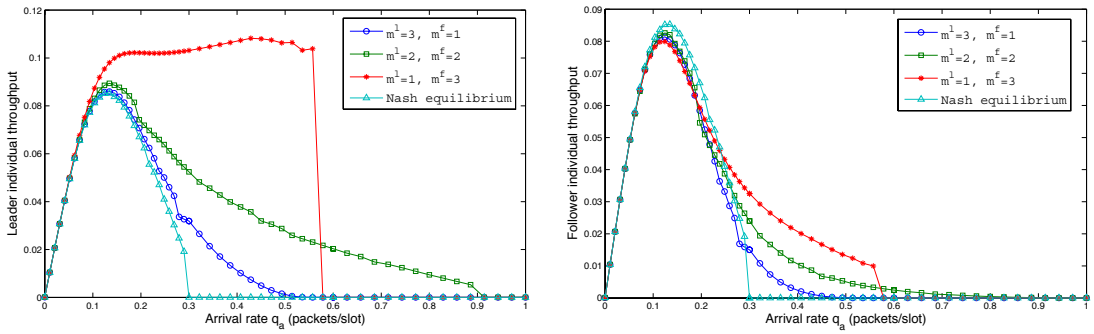
Figure 2.5: *Optimal arrival probability of the controller as a function of the followers arrival probability q_a of the followers for 4 (a) and 10 (b) users.*

- A first result is that the throughput of any player, leader or follower at the Stackelberg equilibrium, is larger than the throughput of this player at the Nash equilibrium taken as reference. This result is particularly interesting for the follower case, indeed it is somehow surprising that a follower improves his utility in a context of a Stackelberg game compared to a Nash game (this result is obviously true for a leader). We observe also on figure 2.9 that any users (leader or follower) experience less expected delay than the Nash game.
- A second result is that the throughput of any leader is always higher than any follower for every repartition scheme. We observe the same behavior of the throughput for the Nash game and for the Stackelberg game depending on the arrival rate.
- The competition between the leaders has an impact on the performance of the system. Indeed, we observe that the throughput of any player is decreasing function depending on the number of leaders. We depict in figure 2.11 (a) and figure 2.11 (b) the retransmission probabilities of any player at the Stackelberg equilibrium. This metric shows the level of aggressiveness and we observe two different behaviors for a leader depending if the number of leaders is higher than the number of followers. The hierarchy mechanism induces that any follower becomes quickly totally aggressive (probability of retransmission is close to 1) even if the group of followers is small. This aggressiveness is due to our free-cost model. Indeed, the followers have to compete between them and also with the leaders. For the leaders, the competition between followers is taking into account for their optimal decision and we obtain that they should be non-aggressive at the equilibrium (when the number of leaders is less than the number of followers) which is not the case for the Nash game. This explains the lower collision rate and the enhancement of the leader's throughput.

In terms of EDBP, we note that at low, average and high loads , for every repartition schemes, the hierarchy mechanism outperforms slotted aloha (see figure 2.9). At very high load ($0.85 \leq q_a$), the EDBP of a leader or a follower can be larger than that one of Nash equilibrium. This heavy load situation leads to very bad performance of the system, the individual throughput is null. One solution to solve this problem is to introduce an admission control such that the load of the system can not exceed some threshold.

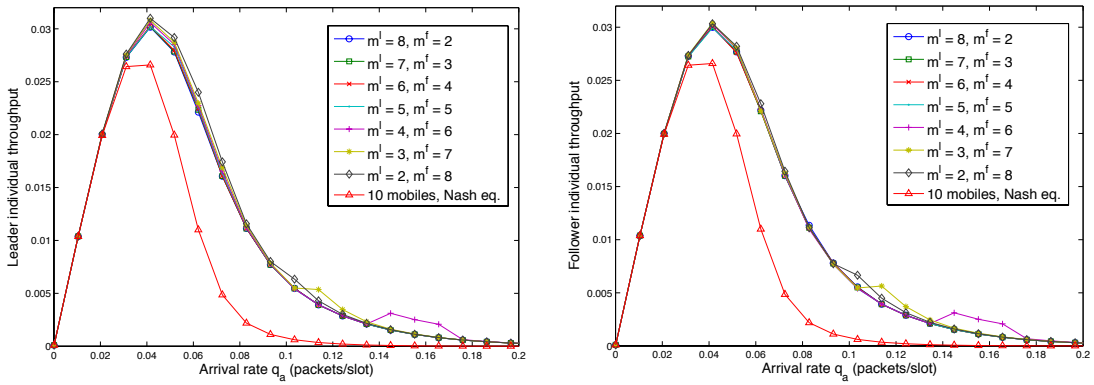
2.6.4 Optimal repartition

Through numerical examples, we come out that performances (in term of throughput and delay) are generally improved with a lower number of leaders. Hence the case with only one leader outperforms all other schemes. One can intuitively predict this result since the scheme with only one leader has the following behavior: Followers are competing over the shared resource whereas the leader is a maximizer of the game, i.e., the decision of the latter is the retransmission probability that maximizes its utility.



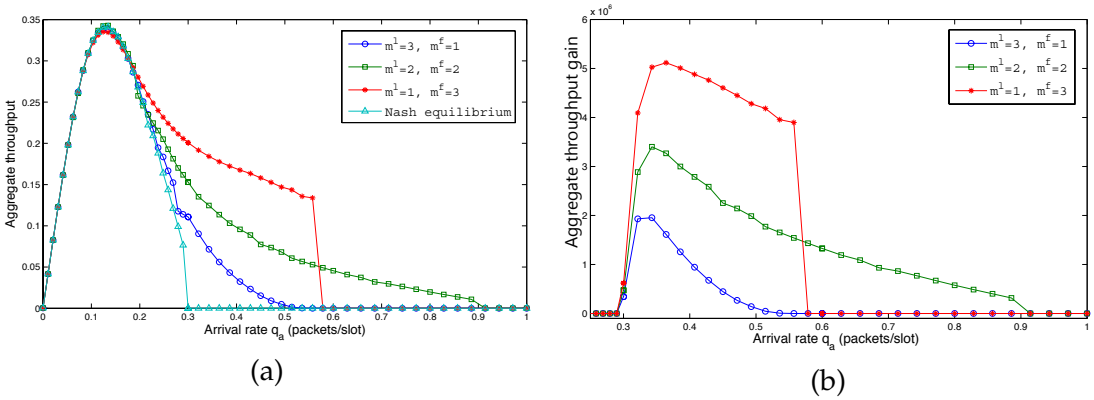
(a) (b)

Figure 2.6: *Individual throughput of a leader (figure a) and a follower (figure b) versus the arrival probability q_a, under different number of leaders. The total number of users is 4.*



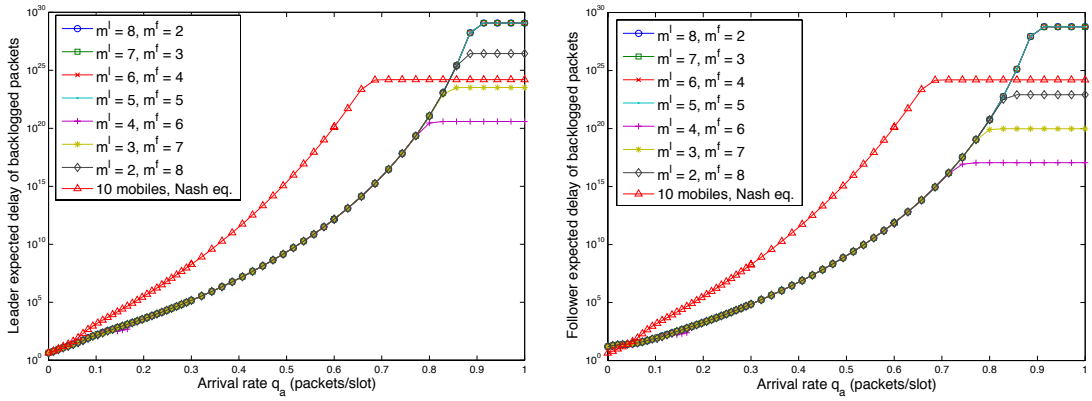
(a) (b)

Figure 2.7: *Individual throughput of a leader (figure a) and a follower (figure b) versus the arrival probability q_a, under different number of leaders. The total number of users is 10.*



(a) (b)

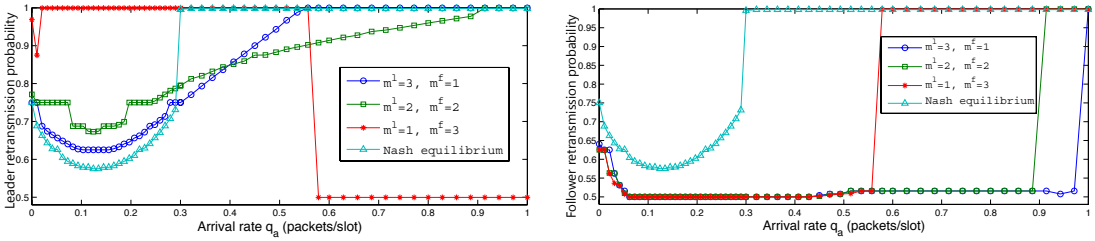
Figure 2.8: *Subfigure (a) shows the total throughput of the hierarchical slotted aloha taking Nash equilibrium point as reference. Bellow $q_a = 0.3$, the two systems has equivalent performances in term of throughput. For the moderate and heavy loads, the hierarchical scheme outperforms slotted aloha for all repartitions of leaders and followers. Subfigure (b) depicts the aggregate throughput gain versus the arrival probability q_a for 4 users. The gain is notable at average and high loads, we also see that the hierarchical system is performing better as the number of leaders is small.*



(a) (b)

Figure 2.9: *Expected delay of backlogged packets of a leader (figure a) and a follower (figure b) versus the arrival probability q_a for 10 users, when changing the number of leaders in the network.*

We observe also that the optimal repartition for maximizing the leader's throughput depends on the arrival rate. When the arrival rate increases, the optimal repartition is to have more and more leaders. The competition between the leaders takes up the load of the system.



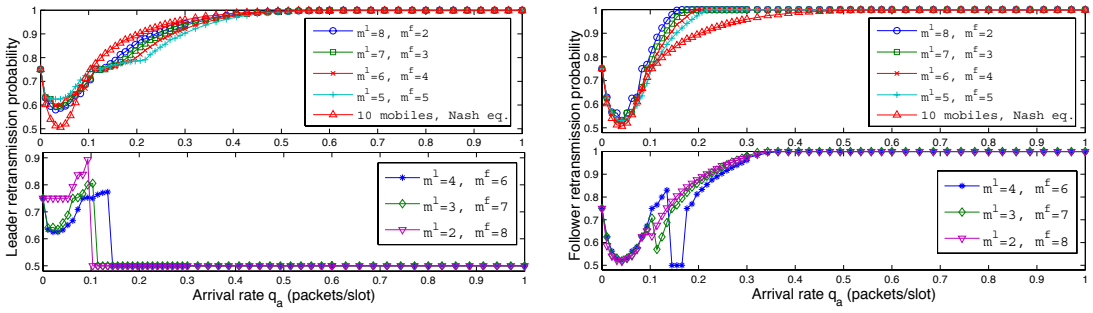
(a) (b)

Figure 2.10: *Optimal retransmission probability of a leader (figure a) and a follower (figure b) versus the arrival probability q_a for 4 users.*

2.6.5 Stability of hierarchical aloha

Finally, we observe through the simulations that the average throughput (total throughput divided by the number of users) is improved whatever the load of the system. This implies that the success probability is also enhanced. Therefore we can deduce that the stability region of our hierarchical scheme is larger than that one of standard slotted aloha. Depending on the arrival rate and retransmission policies, one can also have some cases where the protocol does not suffer from the bi-stability[2] problem, i.e., the

[2]System equilibrium points occur where the curve of success probability and the straight line representing arrival rate $(m-n)q_a$ intersect. m is the total number of users and n is the number of backlogged packets (see Chapter 1 for more details).



(a) (b)

Figure 2.11: *Optimal retransmission probability of a leader (figure a) and a follower (figure b) versus the arrival probability q_a for 10 users.*

new protocol will only have one equilibrium which is the desired one.

2.7 Concluding remarks

We have proposed a hierarchical scheme for slotted aloha based on a Stackelberg game. The performance, either in terms of throughput and delay of any player is improved compared to a system without hierarchy (Nash game). The gain depends on the load of the system. At low load, one can explain this improvement by the fact that users (either leaders and followers) are more aggressive at low load, hence the backoff time is reduced. At average and high loads, leaders turn to be generally less aggressive than Nash equilibrium case. This reduce the collision probability and then increases the performance of the hierarchical scheme. We have seen a pretty phenomena when the number of followers becomes larger than the number of leaders. Those latter players become more friendly and reduce their retransmission rate, whereas the followers become very aggressive and transmit at probability close to 1. This way, the throughput vanishes drastically and a huge of delay is observed. In order to support delay sensitive services, instead of maximizing individual throughput, one can consider the average delay as an alternative objective function. Moreover, a more sophisticated objective such as $\alpha T + (1-\alpha)/EDBP$ (see Chapter 1) where α is a weight that can define the preference level to either maximize the throughput or minimize the delay, can be adopted.

We also discussed the protocol design and how our scheme can be implemented in real systems. We proposed a hierarchy mechanism based on the virtual controller concept. According to the congestion level of the channel, the virtual controller introduces some noisy transmissions to make the channel lossy. This result in a decrease of retransmission probabilities of the concurrent users and then a distributed sustain of partial cooperation. An efficient delay limitation is then obtained (the delay is almost constant compared to exponential increase under slotted aloha) as well as avoidance of

the well known throughput collapse, in particular at average and high loads.

In the next chapter, we revisit the system based on slotted aloha protocol. Each user has some QoS requirement and the transmit with some probability to achieve this later. We analyze the constrained Nash equilibrium which seems to be the most appropriate solution concept for this kind of problems. Furthermore, we discuss the uniqueness issue as well as the stability region of the protocol. We also propose two algorithms to converge to the desired equilibrium point. The first one needs a partial information, but the second algorithm is fully distributed and then it doesn't needs any external knowledge.

Chapter 3

Learning Constrained Nash Equilibrium in Wireless Collision Channels

Contents

3.1 Introduction

Modern wireless networks protocols are often based on Aloha-related concept. For our analysis, we use the standard slotted Aloha model. Slotted Aloha and its unslotted version (pure Aloha) has been central to the understanding of random access networks. These two protocols have over the years evolved into a rich family of medium access control schemes, most notably CSMA/CD, the Ethernet standard, and CSMA/CA which is the basis of the IEEE 802.11 protocol. All results presented in this chapter are easily extended to CSMA and CSMA/CD via the concept of virtual slot concept that we will detail in following chapters. A major challenge in designing such protocol is how to provide quality of service (QoS) guarantees to various multimedia applications. Quality of Service (QoS) is defined as the ability to provide a level of assurance for data delivery over the network. Hence the required throughput of a node may be dictated

by its application (such as video or voice).

We reconsider in this chapter a shared uplink in the form of a collision channel, where a users's transmission can be successful only if no other user attempts transmission simultaneously. Packets those are involved in a collision are backlogged and are scheduled for retransmission after a random time. The determination of the above random time can be considered as a stochastic control problem. We study this control problem in a noncooperative framework: each user has a fixed throughput demand and it dynamically adapts its transmission probability in order to obtain its required demand. For such game, we need to reuse the important and robust concept of Nash equilibrium. Authors in [103] studied the propriety of Nash equilibrium (NE) under saturated users, i.e., each user has always packets ready to send. Unfortunately, the saturation assumption is unlikely to be valid in most real networks. Data traffic such as web browsing and email are typically bursty in nature while streaming traffic such as voice operates at relatively low rates and often in an on-off manner. Hence, for most real traffic the demanded transmission rate is variable with significant idle periods, e.g., CSMA and its variants. In [10, 159], authors assumed that users are independently active with some given probability. Our first main result is then to show the inaccuracy of the independence assumption used by those latter works. Our second aim is to derive a mathematical model that allows us to study an important example of non saturated case, the slotted aloha system with infinite buffer queues. More precisely, a user has in general a limited information to transmit, hence it will stop using network when it succeeds all its transmissions. Here, we analyze the system equilibria without this assumption. We consider a more realistic model in which users transmit only during their activity period. Yet the activity duration depends on the volume of information the user needs to send and the required transmit rate. We show that this new approach enlarges the existence condition of the Nash equilibrium. In contrast to [103] for this non saturated case we establish the possible existence of infinitely many NE.

In addition to discussing some properties of the Nash equilibrium, we also propose two discrete stochastic learning transmission control algorithms which converge to the efficient Nash equilibrium. The first one is based on the Best Response algorithm in which all users iteratively updates their transmission probability through a given rule. But in this algorithm, all users should be able to obtain good estimates of the idle probability. We finally propose another learning algorithm using a stochastic iterative procedure. We approximate the control iterations by an equivalent Ordinary Differential Equation (ODE) to prove that the proposed stochastic learning algorithm converges to a Nash equilibrium even in the absence of any extra information.

A stochastic learning technique has been successfully used in wireless network [19, 134, 164]. In [164], the authors propose a stochastic learning algorithms for distributed discrete power control game in wireless network. At each iteration, the only information needed to update the power strategies for individual terminal users is the feedback (payoff) from the base station. The convergence and stability of the learning algorithm

are theoretically studied in detail for a two-user two-power-level case.

3.2 Model description and Problem formulation

We consider a finite user population m users those transmit to a common base station over a shared channel. We identify each one by a unique i.d. number between 1 and m. We assume that time is slotted and all packets have the same length. Since we are interested in collision channels (such Aloha-like systems), a successful transmission occurs when only one user transmits in the current slot. We note that packets those are involved in a collision are backlogged and are scheduled for retransmission after a random time. Each user i handles a buffer Q_i (see figure 3.1) that carries packets arriving from high layers. Assume that packets arrive to the buffer Q_i according to a Bernoulli process with fixed parameter λ_i. Hence λ_i represents the normalized throughput demand (in number of packet per slot) for the user i needed to hold the service reliability. Until we contraindicate, we assume that the buffers Q_i have infinite capacity of storage, hence the loss probability due to buffer overflow is null. In the rest of the chapter we refer to the vector of throughput demands by $\underline{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_m)$.

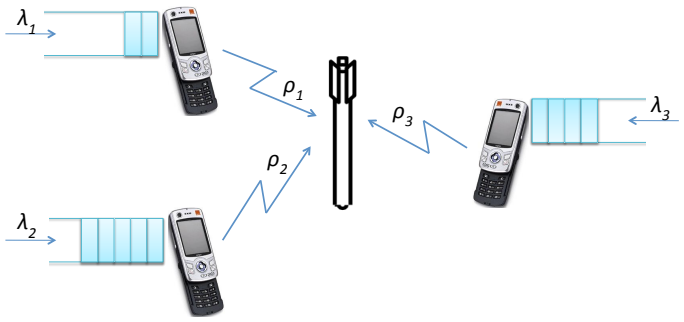


Figure 3.1: *One central receiver and multiple transmitters. Each transmitter i handles an infinite-capacity buffer and requires some minimum QoS λ_i.*

Remark 3.2.1. *As mentioned already, this kind of systems are not analyzed in game theoretic framework before and we are trying to consider one important example of a non saturated system here. Indeed, a slotted aloha system with infinite buffer is non saturated as the users may not be busy in the beginning of each slot.*

The underlying assumption of our user model is that users are selfish and do not cooperate in any manner in order to obtain their required throughput demands. We note clearly that the transmission rate of each users affects the throughput of other users. Each user i fine-tunes its transmit probability so as to maximize its throughput (cannot exceed λ_i); A non cooperative game is then established. A user is said to be *active* if it has, in its buffer, at least one packet ready to be transmitted. Let q_i denotes the probability that a user i transmits on a given time slot and denote by $\underline{q} = (q_1, q_2, ..., q_m)$ the transmit probabilities vector. We characterize the state of the system by an m-dimensional

vector. Let $\mathcal{M} = \{0,1\}^m$ represents the set of all 2^m subsets of $1,2...,m$. At each time slot, a subset $\mathcal{Z}$ of users is assumed to be active. The instantaneous number of active users is given by $|\mathcal{Z}(t)| = \sum_{i=1}^{m} \mathcal{Z}_i(t)$ which is the *Hamming weight* of $\mathcal{Z}$ [10]. The average throughput of user i is then given by:

$$\rho_i(\underline{q}) = q_i \zeta(e_i) \sum_{\mathcal{Z} \in \mathcal{N} \setminus \{i\}} \zeta(\mathcal{Z}) \prod_{j \in \mathcal{Z}} (1 - q_j), \tag{3.1}$$

e_i is the vector whose all entries equal zero but the ith which equals one and $\zeta(e_i) = \pi_i$ is the probability that user i will be active. The equation (3.1) generalizes the throughput formula of collision channels where at most one successful transmission can occur per slot. Indeed, the activity probabilities of users present in the system are correlated, interdependent and still depend on time. Hence we shall write $\pi_i(\underline{q})$ instead of π_i.

The analysis of this system is quite complicated because of the complex nature of the formula (3.1). Even in symmetric user case this formula does not simplify to a good extent. In some papers, authors try to approximate this success probability under independence assumption. Under this assumption, the equation (3.1) would have simplified to [159]

$$q_i \pi_i(\underline{q}) \Pi_{j \neq i} (1 - q_j \pi_j(\underline{q})). \tag{3.2}$$

However unfortunately, this approximation is not good even for symmetric cases (please see figure 3.2). Hence the analysis of this system can not be simplified. This makes this study interesting and we carry out the analysis without independence assumption to obtain the required accurate analysis. We will start this job with first understanding the stability behavior of the system.

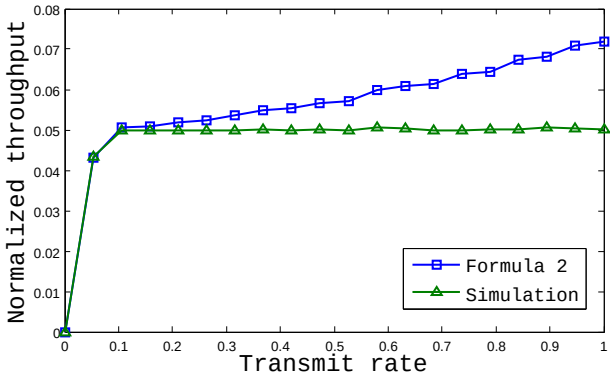


Figure 3.2: *Figure illustrating the error while formula (3.2) based on the independence assumption is approximating the average throughput (success probability) of a symmetric Aloha system. Here, each user is assumed to have a fixed demand $\lambda = 0.05$ (packets per slot). Transmit rate of 3 users is fixed to 0.5 whereas it is varying for the 4th user.*

3.3 Stability region and rate balance equation

Let Aloha($\underline{\lambda}$, $\underline{q}$) represent the slotted aloha system with arrival rates $\underline{\lambda}$ and transmitting probabilities $\underline{q}$. An Aloha($\underline{\lambda}$, $\underline{q}$) system is an example of finite number of interacting queues. An Aloha system is stable whenever parameters related to it like, Buffer sizes etc., does not grow with probability one (please see [151] for exact definition). The stability analysis of such a system has been carried out in literature to a good extent.

Previously some studies have obtained explicit conditions for stability of Aloha $(\underline{\lambda}, \underline{q})$ for given values of $(\underline{\lambda}, \underline{q})$. Some of the interesting results in this direction are : the system is stable if it satisfies the condition (3.5) mentioned in the next section. In most of the other cases ([120], [151]), authors study the stability region, defined for a given transmit probability vector $\underline{q}$, as the set of arrival rate vectors $\underline{\lambda}$ for which Aloha $(\underline{\lambda}, \underline{q})$ system is stable. While in some cases (for example [15]), they define the stability region as the set of those arrival rates $\underline{\lambda}$ for which there exists a transmit vector $\underline{q}$ with Aloha $(\underline{\lambda}, \underline{q})$ system being stable.

We are interested in this chapter in a different concept of stability region defined for a given arrival rates $\underline{\lambda}$ to be :

$$\mathcal{Q}(\underline{\lambda}) := \left\{ \underline{q} \in [0,1]^m : \text{ Aloha } (\underline{\lambda}, \underline{q}) \text{ is stable} \right\}.$$

The reason for this kind of stability region becomes evident in the next section. In the following we obtain an alternative characterization of this stability region.

If an infinite buffer queue is stable it satisfies flow balance equations (as there can be no loss of packets): The input flow rate must be equal to the output flow rate. Here the arrival rate is the rate at which packets from higher layers arrives, i.e., λ_i whereas the departure rate is exactly the success probability of user i denoted by P_i^{succ}. Success probability P_i^{succ} is same as $\rho_i(\underline{q})$ of equation (3.1) and these two terms are used to refer the same quantity. Hence our infinite buffer slotted aloha system is stable whenever

$$\lambda_i = P_i^{succ}(\underline{q}) \text{ for all } i. \tag{3.3}$$

and hence we have:
Proposition 3.3.1. *The stability region is then given by:*

$$\mathcal{Q}(\underline{\lambda}) = \left\{ \underline{q} \in [0,1]^m : \lambda_i = P_i^{succ}(\underline{q}) \textit{ for all } i \right\}.$$

3.4 Nash equilibrium analysis

As mentioned before, we study this problem using game theoretic frame work. It is interesting to note here that we will actually have a constrained game (will be more clear in the following paras). As a first step, we define the strategies of the underlying non cooperative game.

3.4.1 Feasible strategy

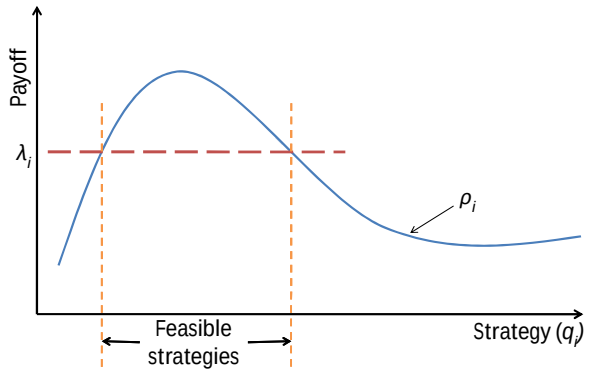


Figure 3.3: *Feasible strategies of user i under throughput demand λ_i and average throughput ρ_i.*

Each user i has a demand in the form of throughput λ_i. Then, it fine-tunes its instantaneous transmit rate q_i in order to fulfill this demand. Let q_{-i} be the strategies vector of other users. From illustrative example in figure 3.3, one can see that user i has, in general, several strategies that can guarantee the demand λ_i. Let $\Gamma_i(q_{-i})$ be the set of all feasible strategies of user i, i.e., all strategies those provide a throughput greater than or equal λ_i. This requirement of the constraint on the strategies results then in a constrained game. We recall however that the obtained throughput of some given user i cannot exceed its arrival rate λ_i

Definition 3.4.1.1 (Definition 1). **.** *A Nash equilibrium is a profile $\underline{q}^*$ of feasible strategies such that no user can improve its utility by deviating unilaterally from its strategy to another chosen from the set of feasible actions $\Gamma_i(q^*_{-i})$. Namely, (q^*_i, q^*_{-i}) is a NE if and only if*

$$\rho_i(q^*_i, q^*_{-i}) \geq \rho_i(q_i, q^*_{-i}),\ \forall q_i \in \Gamma_i(q^*_{-i})\ et\ \forall i \tag{3.4}$$

3.4.2 Constrained Nash Equilibrium (CNE)

We have a m-player constrained noncooperative game with throughputs as the utilities. Let $\mathcal{N}(\underline{\lambda})$ represent the region of all Constrained Nash equilibria.

The constrained game has a NE $\underline{q}^*$ if and only if satisfies (3.3). Hence for the infinite buffer Slotted Aloha system, the region of CNE coincides with stability region, i.e.,

$$\mathcal{Q}(\underline{\lambda}) = \mathcal{N}(\underline{\lambda}).$$

3.4.3 Existence of CNE

As mentioned in the previous section, for a given arrival rates $\underline{\lambda}$, the transmit rate vector $\underline{q}$ is a CNE if and only if Aloha $(\underline{\lambda}, \underline{q})$ is stable. There are many results ([155], [120], [151] etc.) which establish the stability conditions, which thereby give the conditions for

existence of NE. One such condition is (Theorem 1, [120]) : An Aloha ($\underline{\lambda}$, $\underline{q}$) system is stable if for all i,

$$\lambda_i < q_i \Pi_{j \neq i}(1 - q_j). \tag{3.5}$$

Below we establish the existence of infinitely many NE.

Lemma 3.4.3.1. *If there exists a NE $\underline{q}^*$, then there exists a ball $B(\underline{q}^*, \epsilon)$ such that all the vectors in the ball are NE, i.e., $B(\underline{q}^*, \epsilon) \subset \mathcal{N}$.*

Proof : At NE (if exists, by flow balance equations of the stable systems),

$$\lambda_i = P_i^{succ}(\underline{q}^*), \text{ the success probability.}$$

Let π_i^* represent the activity probability of the system at NE $\underline{q}^*$ for all i. A stable system is non saturated, i.e., the activity probability $\pi_i^* < 1$ for all i. Hence, by upper bounding using independence upper bound for all i,

$$\begin{aligned} \lambda_i &= P_i^{succ}(\underline{q}^*) \\ &\overset{a}{\leq} \pi_i^* q_i^* \Pi_{j \neq i}(1 - \pi_j^* q_j^*) \\ &< q_i^* \Pi_{j \neq i}(1 - q_j^*). \end{aligned}$$

In the above the inequality a is obtained using the following reasoning : Consider a modified system where the packets in queue i are generated with probability π_i^* independently across time as well as mobiles in every slot. No packet is carried over to the next slot, however a successful transmission is declared only if there is no collision. Clearly the probability of successful transmission at mobile i, in this system equals $\pi_i^* q_i^* \Pi_{j \neq i}(1 - \pi_j^* q_j^*)$. The probability of two or more buffers being active simultaneously is higher in the original Aloha system in comparison with the changed system. Hence the probability of collisions in the original system is higher and hence we obtain the upper bound a.

By continuity of the map

$$\underline{q} \mapsto \begin{bmatrix} q_1 \Pi_{j>1}(1 - q_j) \\ q_2 \Pi_{j \neq 2}(1 - q_j) \\ \vdots \\ q_N \Pi_{j<N}(1 - q_j) \end{bmatrix}$$

we have a ball $B(\underline{q}^*, \epsilon)$ such that for all $\underline{q} \in B(\underline{q}^*, \epsilon)$ and for all i,

$$\lambda_i < q_i \Pi_{j \neq i}(1 - q_j).$$

Hence Slotted Aloha ($\underline{\lambda}$, $\underline{q}$) system is stable for each $\underline{q} \in B(\underline{q}^*, \epsilon)$. Hence $B(\underline{q}^*, \epsilon) \subset \mathcal{N}$. □

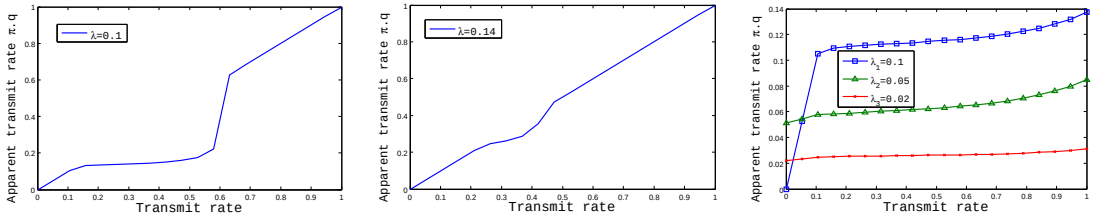
The above lemma establishes that the NE region $\mathcal{N}(\underline{\lambda})$ is either empty or is a open set. Further, using the intermediate results in the lemma and condition (3.5), one can obtain the following alternative characterization of the NE region (when non empty):

$$\mathcal{N}(\underline{\lambda}) = \left\{ \underline{q} \in [0,1]^m : \lambda_i < q_i \Pi_{j \neq i}(1 - q_j) \text{ for all } i \right\}.$$

3.4.4 Energy Efficient Nash Equilibrium

As seen in the previous section whenever CNE exists, there in fact exist infinitely many of them. From Lemma 3.4.3.1, in fact, the set of CNE, $\mathcal{N}$, is an open set. In this section, we are interested in choosing a vector from $\mathcal{N}$ which will be energy efficient. Assume that $\mathcal{E}$ is the energy consumed for one single transmission. Hence the expected energy consumption of user i at any given slot is

$$E_i = q_i \pi_i \mathcal{E} \tag{3.6}$$



(a) (b) (c)

Figure 3.4: *figure -a and Fig-b show the apparent transmit rate as a function of the effective transmit rate for 3 symmetric users under different throughput demands. Figure (c) shows the apparent transmit rate for 3 users with different throughput demands.*

There can be many ways in obtaining such a Energy Efficient NE. Empirically, the quantity $\pi_i q_i$, representing the apparent transmit rate, is increasing with q_i (see figure 3.4). We note that when the demand is high the apparent transmit rate tends to be equal to the effective transmit rate q_i since users become saturated. In the following we define EEE (Energy Efficient Nash Equilibrium) as the the point where the total energy consumption is minimized. It is given using the most simple objective function possible:

$$\begin{aligned} EEE &\propto \operatorname*{argmin}_{\underline{q}\in\mathcal{N}} \prod_{i=1}^{N} E_i \\ &= \operatorname*{argmin}_{\underline{q}\in\mathcal{N}} \prod_{i=1}^{N} q_i \pi_i \mathcal{E} \\ &= \operatorname*{argmin}_{\underline{q}\in\mathcal{N}} \prod_{i=1}^{N} q_i \end{aligned} \tag{3.7}$$

Whenever the transmit probabilities are less (for all users), the amount of collisions will be less and hence minimizing the above objective function gives us an energy efficient equilibrium.

The set $\mathcal{N}$ is bounded, by continuity arguments one can obtain an optimal point in the closure of the set $\mathcal{N}$. Even if the optimal point lies in the boundary of $\mathcal{N}$, one can chose a vector in $\mathcal{N}$ close to the optimal point so that its energy efficiency is arbitrarily close to that of the optimal point. Hence, there exists a $\underline{q}^* \in \mathcal{N}$ which will be (nearly) energy efficient.

3.5 Stochastic 'CNE' Learning Algorithms

In this section we will describe two stochastic iterative algorithms which converge to the *efficient equilibrium* point. Mobiles learn what strategy to adopt in order to obtain their required throughput. The first one is a semi distributed algorithm while the second one is a completely distributed and information less type.

3.5.1 Best Response-based Distributed Algorithm (BRDA)

In [81] and [103], the authors proposed a best response-based algorithm that allows users to learn the efficient equilibrium point for saturated case. One can extend the same iterative algorithm for non saturated (infinite buffer) case as follows.

Let q_i^t, π_i^t respectively represent the transmit probability and activity probability of user i at time slot i. Let

$$\begin{aligned} x_i^t(\underline{q}, \underline{\pi}) &:= \sum_{\mathcal{Z} \in \mathcal{N} \setminus \{i\}} \zeta(\mathcal{Z}) \prod_{j \in \mathcal{Z}} (1 - q_j^t) \\ &\approx \prod_{j \neq i} (1 - \pi_j^t q_j^t) \end{aligned}$$

be the idle probability of all users but the i^{th} one at time slot t, which is approximated using independence assumption. Also, let ϵ_i^t be the update (learning) step at iteration t. At the beginning of each slot, we assume that the base station broadcasts the information $\{x_i^t, i \leq m\}$. The BRA version for the non saturated users is as below.

Algorithm 1 : BRA

1: **for** $i = 1, 2, ..., m$. **do**
2: $\pi_i^{t+1} := \pi_i^t + \epsilon_i^{t+1} \left(1_{\{i \text{ active}\}}^{t+1} - \pi_i^t \right)$,
3: $q_i^{t+1} := q_i^t + \epsilon_i^{t+1} \left(\frac{\lambda_i}{x_i^t} - \pi_i^{t+1} q_i^t \right)$,
4: $1_{\{i \text{ active}\}}^t$ is indicator of the event that at t^{th} time slot, user i has a packet to transmit.
5: **end for**

However the above extension of BRA algorithm has two important problems. It uses independence assumption in calculating idle probability (at BS), which as shown in Figure 3.5 may not be a good approximation. If one were to estimate this quantity accurately, one needs to estimate $\zeta(\mathcal{Z})$, for each possible subset $\mathcal{Z}$, also iteratively as in the case of π_i. But this will complicate the algorithm considerably. Further, BRA is not a completely distributed algorithm and it requires the transmit rate information of all the users at every time step. This consumes valuable bandwidth and is also hard to implement. Further, this calls in for cooperation among users to share their private (transmit rate) information.

3.5.2 Fully Distributed Throughput Predicting Algorithm (FDTPA)

We now turn to develop a new algorithm that is completely distributed. We design an algorithm in which the users does not need any extra information except their own demand.

The key idea of the algorithm (2) is the following :

- The users can observe the sucess/failure of their own attempt to grab the collision channel.
- They can learn the effective throughput achieved by themselves by using the above observations.
- The users can estimate the deviation of their (current) throughput from their own demand (λ_i) and adjusts their attempt/transmit rate to decrease this error to zero.
- On convergence, each user's effective throughput equals their respective demand. Thus the limit point achieves the demand (which is an important property of the CNE).

Algorithm 2 : FDTPA

1: **for** each node $i = 1, 2, ..., m$. **do**
2: $\rho_i^{t+1} := \rho_i^t + \epsilon_i^{t+1}(1^t_{\{success\ i\}} - \rho_i^t)$,
3: $q_i^{t+1} := q_i^t + \epsilon_i^{t+1}(\lambda_i - \rho_i^{t+1})$,
4: $1^t_{\{i\ success\}}$ is indicator of the event that at t^{th} time slot, user i has transmitted the packet successfully.
5: **end for**

Analysis

We use ODE analysis to study the proposed FDTP algorithm. As a first step, we note that one can probably approximate the trajectory of the algorithm with the solution of following ODE[1] system :

$$\begin{cases} \dot{\rho}_i\ (t) = P_i^{succ}(\underline{q}(t)) - \rho_i(t), \\ \dot{q}_i\ (t) = \lambda_i - \rho_i(t) \text{ for all } i. \end{cases} \tag{3.8}$$

We recall that $P_i^{succ}(\underline{q})$ represents the stationary probability of successful transmission at queue i in a slotted Aloha ($\underline{\lambda}$, $\underline{q}$).

[1]We are currently working towards obtaining the ODE approximation theorem.

Attractors : Since the ODE is approximating the trajectory of the FDTP algorithm, its attractors give the limit points[2] of the algorithm. From (3.8), any attractor $(\underline{q}^*, \underline{\rho}^*)$ of the ODE satisfies:

$$P_i^{succ}(\underline{q}^*) = \rho_i^* = \lambda_i, \text{ for all } i.$$

Hence any attractor of the ODE (3.8) is a CNE and hence the limit point of FDTPA is also a CNE.

Initialization

The ODE can have multiple attractors. It is already shown in the above that there exists infinitely many CNEs (Lemma 3.4.3.1). Hence the limit point of the algorithm depends mainly on the initial point. The algorithm converges to an attractor whenever it is initialized with a point close to it (a point in the region of attraction of the corresponding attractor). We are actually interested in converging to an EEE. The desirable EEE has lower value of q_i^* for all users i in comparison with the other attractors. Thus we will converge (with high probability) to the desired EEE if we initialize each q_i^0 with smallest possible value. Further from (3.1) at every NE, $q_i^* \geq \lambda_i$. Hence the appropriate initialization values for the FDTPA algorithm are $q_i^0 = \lambda_i$ and $\rho_i^0 = \lambda_i$ for all i.

Projected FDTPA

Sometimes FDTPA can diverge because of the second iteration in the algorithm involving the updating of q_i's. The updated value of q_i^{t+1} can go out of the window $[0, 1]$ and this will cause the FDTPA to diverge. This can be taken care by the following projected FDTPA:

Algorithm 3 : Projected FDTPA

1: **for** each node $i = 1, 2, ..., m$. **do**
2: $\quad \rho_i^{t+1} := \rho_i^t + \epsilon_i^{t+1}(1_{\{i \text{ success}\}}^t - \rho_i^t)$,
3: $\quad q_i^{t+1} := \max\left\{\min\left\{1, \left[q_i^t + \epsilon_i^{t+1}\left(\lambda_i - \rho_i^{t+1}\right)\right]\right\}, 0\right\}$.
4: **end for**

3.6 Numerical investigation

In the previous sections, we analyzed the Nash throughput for non saturated buffers and characterized the stability region. In the following, we present some analytical as

[2]This statement actually needs an asymptotic result for stochastic approximation. However it is widely understood that even without this result the above statement is true to a good extent.

well as some simulation results. We use among this section a discrete time simulator with Bernoulli process for packet generating.

3.6.1 Stability and Nash Equilibrium Region

We depict in figure 3.5 the individual throughput for 3 symmetric users when each one has a strict demand $\lambda = 0.1$ (in packet per slot) and infinite buffer. The first main result is the validation, through simulation, of the existence of infinite number of Nash equilibria for this non saturated case. In this example, we restrict all three users to use same transmit probability q. In figure 3.5 we plot the simulated value of P_1^{succ} and the approximating formula (3.2) versus the common transmit probability q. Indeed, we note (curve with triangle marks) that the throughput is increasing with the transmit rate. When the average throughput achieves the demand λ, we note that it becomes constant even if users continue to transmit more aggressively; This shows the existence of several Nash equilibria, potentially a continuum of NE. Over $q = 0.6$, the throughput turns to decrease and vanishes when users become very aggressive (transmit at probability 1). This situation is similar to the prisoners dilemma, it shows in fact that Nash equilibrium is not efficient in some situations. Curve corresponding to the plot of average throughput using equation (3.2) seems to provide an accurate approximation only when users are transmitting at low or at high rate. This can be explained as it follows: On one hand, at low load the wireless network is not congested and then the collision probability is negligible. Therefore, the average throughput is only function of transmit rate and the approximation becomes accurate. On the other hand, when users become very aggressive, the collisions increase, the activity probability increases towards 1, the system moves towards saturated case and henceforth the accuracy of the formula improves. In contrast to simulation, the approximation does not show existence of infinity NE points. To summarize, this kind of approximation is doubtful and has no interest since it seems to be so inaccurate that it can give wrong characterization of Nash equilibria points (equivalently stability region).

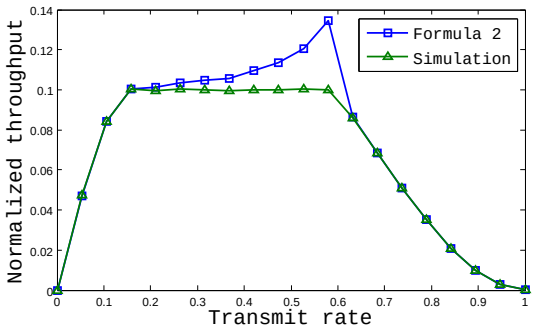
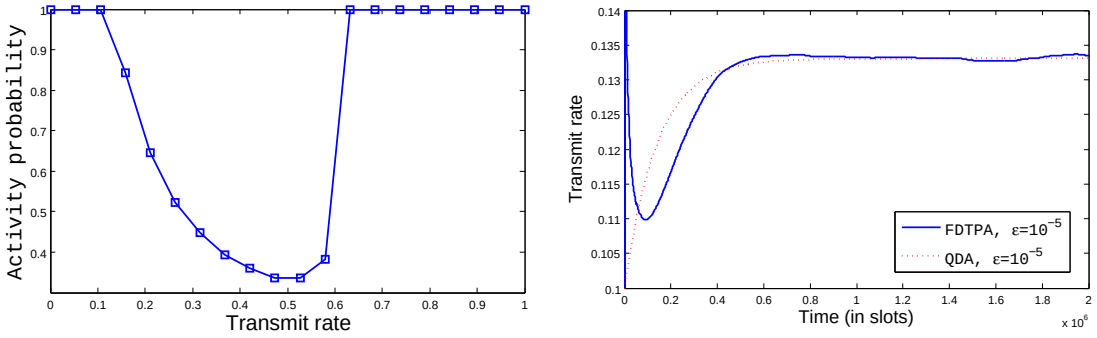


Figure 3.5: *In this figure, we show the individual throughput for 3 symmetric users whose demand is fixed to $\lambda = 0.1$, from both simulation and the approximation (using eq. 3.2).*

Later we plot in figure 3.6-a the variation of the activity probability (busy probabil-

ity) of some tagged user for the example of figure 3.5. When user does not transmit at a rate that can guarantee its nominal demand, we note that $\pi = 1$; This is due to the fact that the arrival rate is still greater than the departure rate. When the tagged user is transmitting at any rate in $[0.15, 0.6]$, the activity probability is less than one; This interval corresponds well to the stability region. Through figure 3.5 and figure 3.6-a, we check easily that the region of Nash equilibria corresponds well to the stability region of the buffers. Further from the NE region characterization of Section 3.4, one can easily calculate that $\mathcal{N} \cap \{\underline{q} \in [0,1]^3 : q_1 = q_2 = q_3\} = [0.135, 0.59]$, which is very close to the stability interval $[0.15, 0.6]$ obtained from the simulations.



(a) (b)

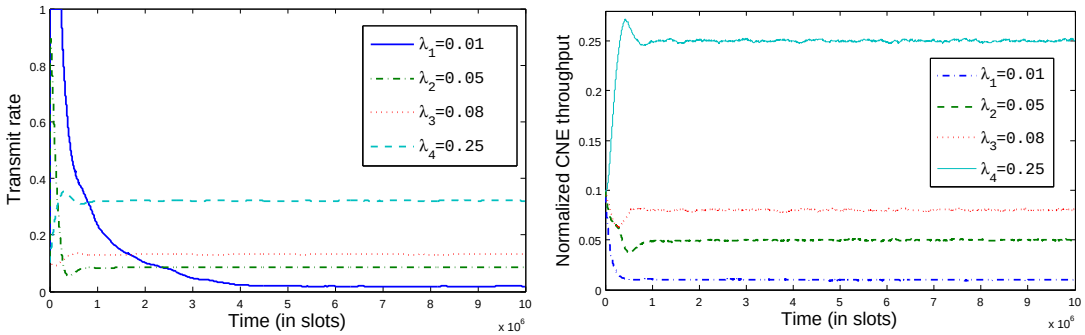
Figure 3.6: *Activity probability versus transmit rate for 3 symmetric users whose demand is $\lambda = 0.1$. In the case of small learning step ($\epsilon = 10^{-5}$), our stochastic algorithm performs as well as the BRA, either in term of accuracy and speed of convergence.*

3.6.2 Convergence of FDTPA and BRA algorithms

We now turn to check and compare the efficiency of our two algorithms. We simulate in Fig 3.6-b the collision channel where users use respectively BRA (dashed line)and FDTPA (solid line). We consider a learning step $\epsilon = 10^{-5}$. It is clear that information less stochastic scheme (FDTPA) tracks the desired Nash equilibrium as well as the version with partial information (BRA). We note similar trends when considering a variable learning step such as $\frac{1}{t+1}$ in Fig 3.7.

Fig 3.9-a and Fig 3.9-b show the crucial importance played by the initial conditions, in particular the initial transmit rate. Indeed if q_i^0 is not initialized judiciously the system can be absorbed by some non efficient equilibrium point. To avoid this problem, noting as said above that relation $\lambda_i \leq q_i$ always holds, initializing the transmit rate vector by the demand vector seems to provide a good start point. Another important factor that controls the speed of convergence, is the learning step, that should be chosen appropriately.
Next we simulate the behavior of the information less algorithm FDTPA in cases where no Nash equilibrium could exist. We depict in figure 3.10-a and figure 3.10-b the trans-

mit rate and the average throughput respectively. The well known result of decentralized slotted aloha with selfish users ([12], [9], [44]) is then obtained: Users transmit w.p. 1 which explain the throughput collapse and the congested situation of the whole system. A remark that might be interesting is that users are more aggressive as their demand is higher.



(a) (b)

Figure 3.7: *Convergence to the throughput demands λ_i under a variable update step $\epsilon = 1/(t+1)$.*

3.6.3 Re-convergence of FDTPA

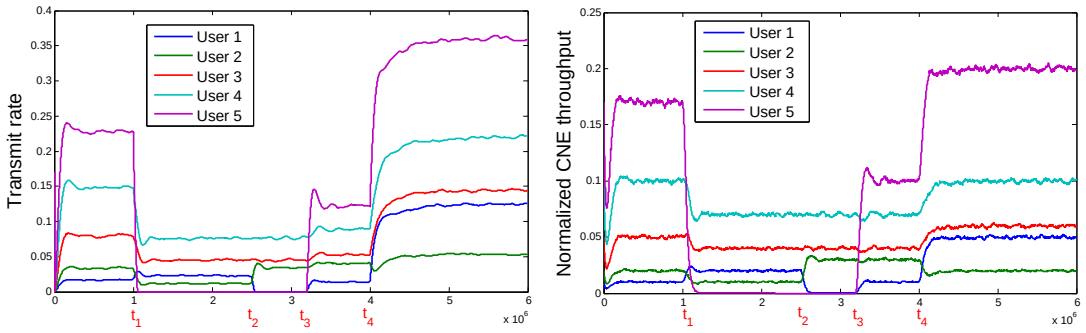
Instant of time	Perturbing event
t_1	User 5 leaves the system. $\underline{\lambda} = (0.02, 0.01, 0.04, 0.07, 0)$
t_2	User 1 leaves the system. $\underline{\lambda} = (0, 0.03, 0.04, 0.07, 0)$
t_3	Users 1 and 5 joint the system. $\underline{\lambda} = (0.01, 0.03, 0.04, 0.07, 0.1)$
t_4	Users change their demand. $\underline{\lambda} = (0.05, 0.02, 0.06, 0.1, 0.2)$

Table 3.1: *Re-convergence of FDTPA versus the users behavior.*

In order to track the behavior of our stochastic algorithm, we consider a system with some troubling events. We consider five users with initial throughput demands vector $\underline{\lambda} = (0.01, 0.02, 0.05, 0.1, 0.17)$. The illustrative events are summarized in table 3.1. We restrict here to the cases where an EEE indeed exists before and after the considered events. We plot in figure 3.8 the corresponding behavior of FDTPA. The users running FDTPA detect the perturbing events (i.e., a user joints/leaves the system or changes its throughput demand) and adapt their transmit rate to re-converge to the new EEE point.

3.6.4 Discussion

The design of new protocols has face to several challenges and hard issues. For instance environment, bandwidth under-utilization, energy constraint, evolution of hardware, resources allocation etc. Here we addressed an interesting problem that will drive progress in MAC layer protocols design. For saturated aloha system, authors in [103]



(a) (b)

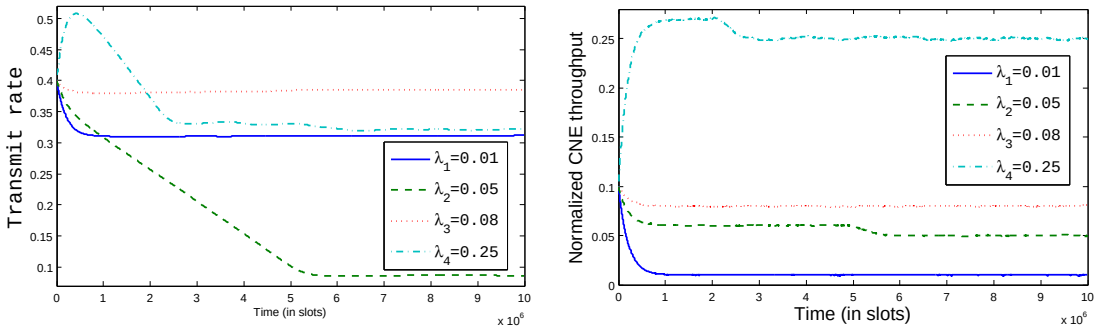
Figure 3.8: *Re-convergence of FDTPA with perturbing events ($\epsilon = 1/(t+1)$).*

presented a nice study for a saturated slotted aloha system. The main result is that when the throughput demands are within the demand feasible region, there exist exactly *two* Nash equilibria, with one strictly better (in terms of energy consumption) than the other for all users. However this seems not to be true for the same system with non saturated users. This can be simply explained as following: consider a NE $\underline{q}^1$, a tagged user i transmits at q_i^1 and is active w.p. $\pi(q_i^1)$, hence the transmit rate of user i perceived by other users is $q_i^1\pi(q_i^1)$. Operating in the stability region, even if tagged user changes its transmit rate, its activity probability varies so as to keep the perceived transmit rate almost constant. Hence there exist many infinite NE. The real issue we should be careful with is the initialization point of transmit rates vector to converge to the EEE. Yet, the fluctuations might bring the system to some non efficient equilibrium. Setting the vector $\underline{q}^0$ to $\underline{\lambda}$ is a judicious starting point that resolves the problem to a good extent.

Considering a fixed rate for every user can be of help to optimize the bandwidth utilization. On one hand, the minimum demand can be seen as the minimum QoS needed to keep the service reliability. On the other hand, this scheme can be seen as an alternative call admission control where the base station may be ensured that the system capacity is never exceeded.

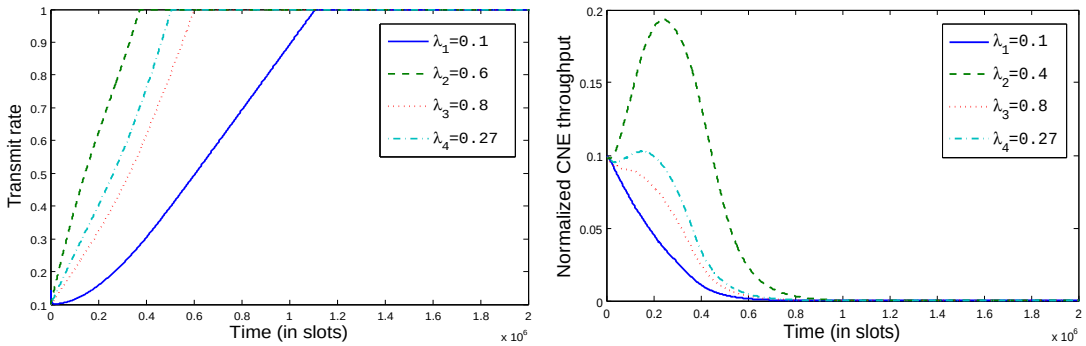
3.7 Concluding remarks

We studied the throughput of collision channels, without saturation condition (usually assumed in the literature), where users have some strict QoS to fulfill. We noted that the achievable throughput is not affected by the users asymmetry and the region of equilibrium is larger than the saturated case. Indeed, we showed existence of infinite number of Nash equilibria. In addition to providing an energy efficiency analysis, characterizing the efficient CNE (EEE), we adapted the algorithm proposed in [81] and [103] taking into account the instantaneous saturation level of users in order to converge to EEE. However this algorithm suffers from many problems and implementation difficulties such as bandwidth consumption, doubtful estimation of previous transmit rate



(a) (b)

Figure 3.9: *Importance of initializing the transmit rate is illustrated in Fig a and b where users converge to a non efficient CNE.*



(a) (b)

Figure 3.10: *When a CNE does not exist users assist to a typical Prisoner's Dilemma phenomenon; Mobiles become very aggressive and transmit w.p. 1 (a) which explains the throughput collapse observed in (b).*

vectors and processing time. This motivated us to propose an information less stochastic distributed algorithm. We showed theoretically and through extensive simulations, the accuracy of the new algorithm as well as its speed of convergence.

We turn in the next Part of this manuscript to analyze the end-to-end performance of heterogeneous wireless networks. We indeed addressed the case of a UMTS/ad hoc integrated network in [43] and then provide a more rigorous performance evaluation framework for WiMAX/ad hoc integrated networks in [45]. Later on, we derive the average end-to-end delay as well as the distribution of delay for more accuracy in a heterogeneous multihop system.

Part II

An end-to-end QoS Framework for multihop Wireless Heterogeneous Networks

Chapter 4

Performance Evaluation of WiMAX and Ad hoc Integrated Networks

Contents

4.1 Introduction

Over the past few years, wireless communication systems have been known a rapid growth and a notable success. Various wireless networks evolve into the next-generation networks (i.e., 4G) to provide better services and thus it is important to study how to integrate different technologies to synergies their advantages while allowing their combined strengths to make up for their individual limitations. In order to switch to 4G networks and to meet the increasing demand in data and user mobility, it is envisioned that the 4G networks will consist of many cooperating wireless technologies, and provide universal connectivity and opportunity for best suited services to users at any time from any where. Users are expected to have several radio interfaces providing the possibility to communicate simultaneously through the different interfaces and choose the “best" interface according to several parameters such as the application characteristics, the user preferences, the networks abilities, the operator policies, tariff constraints, etc. Basically, each interface has different access range, cost (Energy, Economical issue, etc.) and performance which may be a limitation of the always best connected project [61]. Next-generation wireless network (NGWN) will be heterogeneous radio access

and combines of UMTS, WiMAX, LTE, WLAN, etc. [43, 45, 54, 113, 165]. However, NGWN will also be a joint radio resource management among multiple operators [29]. We believe heterogeneous network architecture based on WiMAX networks and assisting ad hoc networks is one likely solution for distribution of high data rate services. This network insight allows simultaneous or alternative connections according to coverage and service constraints. An ad hoc wireless network is a collection of wireless nodes that communicate with each other without any established infrastructure or centralized control. Due to the limited transmission range of wireless network interfaces, multiple network "hops" may be needed for a given node to exchange data with another across the network. In such a network, the packets may have to be forwarded by several intermediate nodes before they reach their destinations, and therefore each node operates not only as a host but also as a router. Thus each node may be a source, destination and relay (intermediate). Many factors interact with each other to make the communication possible like routing protocol and channel access method.

Stub networks convey traffic only to or from their local hosts and never carry traffic for which they are neither the source nor the destination. Multihoming has been traditionally used by stub networks for improving network availability. Indeed, [142] has shown that "Multihoming Route Control" can improve significantly the Internet communication performance. A comparison study of overlay source routing and multihoming is provided in [137]. It also has shown how much benefit does overlay routing provides over Border gateway protocol (BGP), when multihoming and route control are considered. Analytical work on the stability of a system using multiple routes is presented in [74, 92]. In wireless networks, interest of connecting to Internet using different access technologies alternatively or even simultaneously has been described in various works [149]. The non-cooperative multihoming over IEEE 802.11 WLANs, where users can choose several access points and even split their traffic has been studied in [14].

In this chapter, we are interested in the interconnection between a multi-hop wireless ad hoc network and a WiMAX system. A multi-hop wireless ad hoc network is a collection of nodes those communicate with each other without any established infrastructure or centralized control. Many factors interact with each other to make the communication possible like routing protocol and channel access method. Recently, wireless ad hoc networks have been envisioned of commercial applications such as providing Internet connectivity for nodes those are not in transmission range of a wireless access point. Hence, WiMAX and ad hoc networks should be considered as complementary systems to each other. It is possible to use ad hoc as an extension of the cellular networks. Thus a multihomed node can be able to use different access technologies at any time from any where to assure a permanent connectivity. As a result, potential benefits can be envisaged for WiMAX system as: extend cellular coverage locally, reduce the power transmission which implies the reduction of intra-cell and extra-cell interferences. We assume that each node may have two kinds of packets to be transmitted: (i) Packets generated by the device itself. This can be sensed data if we are considering a sensor network. (ii) Packets from other neighboring devices that need to be *forwarded* till

achieving the final destination. For accuracy, we consider two separated queues scheduled using a Weighted Fair Queueing (WFQ) scheme [43]. The main contribution we present here is to provide approximate expressions for stability. This concerns mainly the stability of the forwarding queues, in particular, the nodes (gateways) which are connected to both the ad hoc and WiMAX subsystems. The achieved end-to-end (e2e) throughput is independent of the choice of the WFQ weight. Providing e2e delay has been an important issue in quality of service enabled networks as web phones and videoconferences typically require QoS guarantees in terms of e2e delays. In order to provide support for delay sensitive traffic in such network, an accurate evaluation of the delay is a necessary first step. As expected, the delay is minimized for full altruistic scenario, i.e., when the WFQ weight is set to 1.

Since WiMAX networks can cover a relatively large area, it is natural to imagine that many group communications, such as videoconferences, will be important applications in WiMAX networks. One of the features of the MAC layer of WiMAX is that it is designed to differentiate service among traffic categories with different multimedia requirements. For many applications, it is desirable that the network layer can provide a sufficient quality of service guarantee, usually in terms of bandwidth, data rate, delay, and delay jitter. But the characteristics of ad hoc network impose new constraints due to the variation of limited available resources which derive to a high delay and jitter. This motivates us to study the e2e delay. We then use a $G/G/1$ queueing to compute the waiting time at intermediate nodes between a source and a destination. This study may allow us to study some tradeoff between throughput and delay according to the initialized service.

4.2 Problem formulation

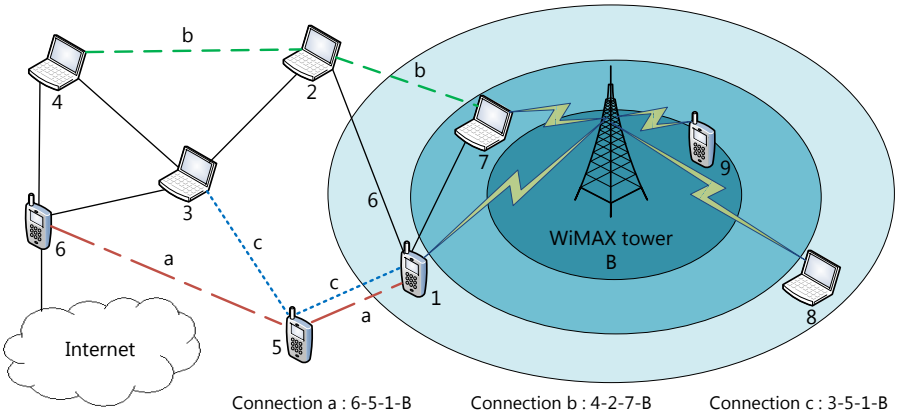


Figure 4.1: *WiMAX and ad hoc integrated network.*

Consider a wide geographical area served by WiMAX technology and assisting an

ad hoc network. We divide it according to available technologies into three geographical classes: (i) Class $\mathcal{C}_a$ that contains N_a terminals which communicate with each other using an ad hoc network, some of them could be covered by WiMAX tower B. (ii) Class $\mathcal{C}_w$ contains N_w terminals directly covered by WiMAX tower B. (iii) Class $\mathcal{C}$ is the overlap of these two classes, i.e., $\mathcal{C} = \mathcal{C}_a \cap \mathcal{C}_w$. It is composed of terminals able to connect to both ad hoc and WiMAX subsystems.

Each mobile is equipped by two separated network cards (IEEE 802.11 and IEEE 802.16). This allows simultaneous or alternative connections to two technologies according to the coverage criterion and offered services. On one hand, an ad hoc node (e.g., node 4 in figure 4.1) which is far from WiMAX tower will be able to benefit services from WiMAX in spite of coverage constraint using forwarding capability of its neighbors. On the other hand, a node (e.g., node 1 in figure 4.1) in the range of B with bad radio conditions cannot use internet at a high speed rate because it cannot use higher modulation level than Binary Phase Shift Keying (BPSK) scheme. An alternative solution is to connect to Internet over ad hoc subsystem (assume that ad hoc is connected to Internet), and hence it will get higher transmit rate and lower cost.

4.2.1 Cross layer architecture

Although we consider a heterogeneous system, the network layer is the same for both subsystems. In OSI model, layers are clearly separated. Here we propose a cross-layer architecture where both network and MAC layers parameters (see figure 4.2) are jointly considered. This allows communication and information sharing between different layers and henceforth is more powerful, flexible and in particular allows global optimization. The network layer handles two queues F_i and Q_i having infinite capacity and managed by the weighted fair queueing scheme. The first one (forwarding queue) carries packets originated from other nodes to some given destination, and the second one is the own queue which contains packets generated by node i itself (e.g., sensed data if we are considering a sensor network). Node i decides to transmit from F_i with probability f_i and henceforth transmits from Q_i with probability $1 - f_i$. This configuration allows nodes to have flexibility for managing at each node forwarded packets and its own packets differently. The system is assumed to be saturated, i.e., each node has always packets to transmit from Q_i whereas queue F_i could be empty. Instantaneous updates are made for each routing table using a proactive routing algorithm such as Optimized Link State Routing (OLSR) which corresponds well with our model. It is clear that each mobile handles two IP addresses and necessarily two multiple link-layer interfaces. We denote by $R_{s,d}$ the set of nodes between a source s and a destination d (s and d are not included in this notation). Indeed, attempting the channel starts by choosing the queue from which a packet would be selected. And then, this packet is moved from the network layer to the MAC layer where it will be transmitted and retransmitted (if needed) until success or definitive drop.

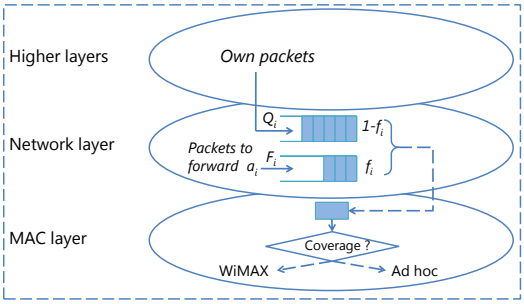


Figure 4.2: *Cross-layer architecture integrating WiMAX and ad hoc subsystems.*

4.2.2 WiMAX MAC layer

IEEE 802.16x MAC layer is designed to support multimedia services (streaming, MPEG video, etc.) and very high peak bit rates. It also fully manages the bandwidth utilization for both uplink and downlink schemes by the means of the polling process[1]. Moreover, it supports scalable QoS depending on the ongoing service by handling some parameters such as the tolerable rate, scheduling type, delay, etc. In this work, tower B uses an OFDMA scheme to communicate with N_w covered users. Now, it allocates dynamically and exclusively to each node i, a set of subchannels denoted by $\mathcal{L}_i = \{l_1, \ldots, L_i\}$ such that $\bigcup_{i=1}^{N_w} \mathcal{L}_i \subseteq [0, \beta]$ and $\mathcal{L}_i \cap \mathcal{L}_j = \emptyset$(non-overlapping channels), $i \neq j$. To improve the bandwidth utilization (results in performance enhancement), the tower B may employ a Dynamic Subcarrier Assignment scheme for subcarriers allocation. Indeed, B sends in each period the new resource allocation table. Therefore, each user scales its coding and modulation levels (Adaptive Coding and Modulation, ACM) for its assigned channels according to the instantaneous SINR. When channel conditions become bad, the ACM-enabled radio system automatically changes modulation and/or coding orders allowing applications to continue to run uninterrupted. Varying the modulation and/or coding orders will also vary the amount of bits that are transferred per signal, thereby enabling higher throughput and better spectral efficiency. Let M_w and $\tau^m_{i,B}$ be the number of bits in a WiMAX packet and the WiMAX packet transmission time (in seconds), respectively. We have $\tau^m_{i,B} = M_w/\rho^m_{i,B}$, where $\rho^m_{i,B} = \sum_{l \in \mathcal{L}_i} r^m_{i,B,l} \Delta f$ is the aggregate transmission rate (bps) allocated for user i when it uses a m-QAM modulation level; $r^m_{i,B,l}$ is the transmit rate (in bits per subcarrier) over subcarrier l and Δf is the bandwidth of one single subcarrier. The physical layer of WiMAX subsystem integrates a hybrid automatic repeat request mechanism (H-ARQ) with a Forward Error Correction (FEC) protection system. This provides improved link performance over traditional ARQ at the cost of increased implementation complexity. Blocks of data with a Cyclic Redundancy Check code are encoded using the FEC coder before transmission. Retransmission is requested if the decoder is unable to correctly decode the

[1]Polling is the periodic allocation mechanism of shared bandwidth, it allows to request bandwidth according to required QoS (rate, delay, etc.) of each user.

received block. We assume that ad hoc/WiMAX gateways retransmit a WiMAX packet to B (if needed) until success or definitive drop. Let K be the maximum number of transmissions allowed by a gateway i per packet for all paths.

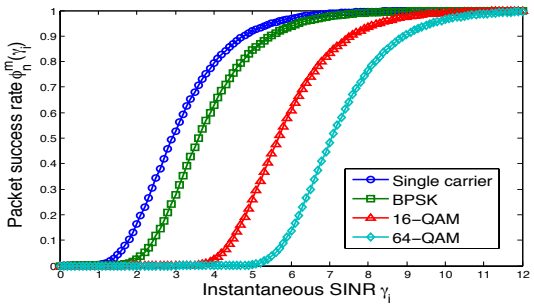


Figure 4.3: *Efficiency function (PSR) for multicarrier scenario, see [109] for single carrier case.*

We define $\phi_n^m(\gamma_{i,l})$ as the success probability for user i over subcarrier l, using an m-QAM modulation scheme, for the n^{th} transmission. It can be any continuous S-shaped function fulfilling $\phi_n^m(0) = 0$ and $\phi_n^m(+\infty) = 1$. The SINR w.r.t. subcarrier l (see [143] for more details) is given by

$$\gamma_{i,l} = \frac{\bar{p}_{i,l}|H_{i,l}|^2}{\bar{N}_{i,l}}, \tag{4.1}$$

where $\bar{p}_{i,l}$ is the transmit power density of user i over the sub-carrier l, $\bar{N}_{i,l}$ is the instantaneous noise power density, and $H_{i,l}$ is the impulse response of the fading channel. We consider here the function $\phi(\overrightarrow{\gamma}_i) = (1 - e^{-\gamma_i})^{M_w}$ to approximate the efficiency function on one single carrier, each WiMAX packet contains $M_w = 1$ Kbits of data and no overhead. Selective retransmission is not supported, i.e., if the signal of at most one subcarrier is corrupted, then the whole of block should be retransmitted. Hence the probability that signals of all subcarriers are received correctly is

$$\phi_n^m(\overrightarrow{\gamma}_i) = \prod_{l \in \mathcal{L}_i} \phi_n^m(\gamma_{i,l}), \tag{4.2}$$

where $\overrightarrow{\gamma}_i$ is the vector whose the j^{th} entry is $\gamma_{i,j}$ given by (4.1). When a retransmitted coded packet is received, it is combined with the previously detected coded packet and fed to the input of the FEC decoder. Combining the two received versions of the code block improves the chances of correctly decoding. This type of H-ARQ is called type-I chase combining. Hence $\phi_1^m(\overrightarrow{\gamma}_i) \leq \phi_2^m(\overrightarrow{\gamma}_i) \leq \cdots \leq \phi_K^m(\overrightarrow{\gamma}_i)$.

4.2.3 Ad hoc MAC layer

Over ad hoc subsystem, nodes use a single channel for transmission with an omnidirectional antenna. It follows that an ad hoc node i successfully transmits a packet to node j, only when there is no interference at node j from its ad hoc neighbors. Let τ^a denote the duration in seconds of one ad hoc slot and each packet is of length M_a-bits. We also

denote by s, d and i, respectively, a source (which generates the packet), the destination and an intermediate node on path $R_{s,d}$. Let π_i denote the probability that the queue F_i has at least one packet to be forwarded in the beginning of each cycle [2]. Also let $\pi_{i,s,d}$ be the probability that the queue F_i has a packet at the first position ready to be forwarded for the path $R_{s,d}$ in the beginning of each cycle. Thus, we have $\pi_i = \sum_{s,d} \pi_{i,s,d}$.

The medium access can be any Aloha-like protocol, CSMA/CA, etc. For instance, the attempt rate (for node i) in the IEEE 802.11 Distributed Coordination Function based multihop wireless network [166] is given by

$$P_i = \frac{2(1-2P_c)}{(1-2P_c)(CW_{min}+1) + P_c CW_{min}(1-(2P_c)^m)}, \tag{4.3}$$

where P_c is the conditional collision probability given that a transmission attempt is made, CW is the contention window and $m = log_2\,(CW_{max}/CW_{min})$ is the maximum of backoff stage. For example, in IEEE 802.11 standard, P_i depends on the number of neighbors, on the backoff mechanism and the probability of collision, see Yang et al. [166] for the ad hoc extension of Bianchi results [26]. Problems of hidden terminals or exposed terminals known with the IEEE 802.11 are included implicitly in the formula of P_i. It is clear that the scheduler of transmission overall the network depends on P_i. We assume that each node is notified about success or failure of its transmitted packets. To keep a reasonable reliability, we assume that a packet is retransmitted (if needed) until success or drop, $K_{i,s,d}$ is the maximum number of transmissions allowed by a mobile i per packet on path $R_{s,d}$. We also denote by $L_{i,s,d}$ the expected number of attempts till success or definitive drop from node i on the path from s to d. Let $\bar{L}_i$ be the average of all $L_{i,s,d}$ over all sources s and destinations d, and let $P_{i,j}$ be the probability that node i generates and transmits a packet to node j. Each mobile transmitting a packet accesses the channel with probability P_i [26]. The attempt rate P_i is only valid for IEEE 802.11 systems and does not hold in our heterogeneous network. Let x_i be the total proportion of WiMAX cycles for a given node i and τ_i^w be the average needed number of slots to send a WiMAX packet. The WiMAX proportion traffic and the transmit probability over the ad hoc subsystem are, respectively, given by the next proposition.

Proposition 4.2.3.1. *Consider a heterogeneous network composed of a WiMAX and an ad hoc subsystems, then we have*

1) The proportion of WiMAX traffic in a gateway is

$$x_i = P_{i,B}(1-\pi_i f_i) + f_i \sum_s \pi_{i,s,B}. \tag{4.4}$$

2) The attempt rate for any given node i in the system is

$$\bar{P}_i = \frac{\bar{L}_i(1-x_i)}{\bar{L}_i(1-x_i) + \tau_i^w x_i P_i} P_i. \tag{4.5}$$

[2] A cycle is the total number of slots needed to transmit a single packet until success or drop.

Proof. Here we develop a cycle-based method to prove the proposition 4.2.3.1. Let us observe the system (in particular a gateway node $i \in \mathcal{C}$) for $C_{i,t}$ cycles. Assume that $C^a_{i,t}$ is the number of ad hoc cycles until the t^{th} slot, and $C^w_{i,t}$ is the corresponding number of cycles for the WiMAX traffic. We denote by T^a_t the total number of transmission slots used by ad hoc connections till the t^{th} slot. We can write the proportion of WiMAX cycles as

$$x_i = \frac{C^w_{i,t}}{C_{i,t}} = \frac{C^{u,Q}_{i,t} + C^{u,F}_{i,t}}{C_{i,t}}.$$

On one hand, the term $\frac{C^{u,Q}_{i,t}}{C_{i,t}}$ is the probability to choose a WiMAX packet from Q_i. We have
$\frac{C^{u,Q}_{i,t}}{C_{i,t}} = \frac{C^{u,Q}_{i,t}}{C^Q_{i,t}} \cdot \frac{C^Q_{i,t}}{C_{i,t}} = P_{i,B}(1 - \pi_i f_i)$,
and on the other hand, we have
$\frac{C^{u,F}_{i,t}}{C_{i,t}} = \frac{C^F_{i,t}}{T^a_{i,t}} \cdot \frac{T^a_{i,t}}{C_{i,t}} \cdot \frac{C^{u,F}_{i,t}}{C^F_{i,t}} = f_i \pi_i \sum_s \frac{\pi_{i,s,B}}{\pi_i} = f_i \sum_s \pi_{i,s,B}$,
where $T^a_{i,t}$ denotes the number of cycles having ad hoc packets till slot t.
The proportion of WiMAX cycles becomes $x_i = P_{i,B}(1 - \pi_i f_i) + f_i \sum_s \pi_{i,s,B}$ and completes the proof of result (4.4).

The long term attempt rate is $\bar{P}_i = \lim_{t\to\infty} \frac{T^a_t}{t} = \lim_{t\to\infty} \frac{T^a_t}{C^a_{i,t}} \cdot \frac{C^a_{i,t}}{C_{i,t}} \cdot \frac{C_{i,t}}{t}$,
where $\lim_{t\to\infty} \frac{T^a_t}{C^a_{i,t}} = \bar{L}_i$ is exactly the average number of slots per cycle (Ad hoc or WiMAX), $\lim_{t\to\infty} \frac{C^a_{i,t}}{C_{i,t}} = 1 - x_i$ is exactly the proportion of ad hoc cycles among $C_{i,t}$, and $\lim_{t\to\infty} \frac{t}{C_{i,t}} = \frac{C^a_{i,t} \cdot \frac{\bar{L}_i}{P_i} + C^w_{i,t} \cdot \tau^w_i}{C_{i,t}} = \frac{\bar{L}_i}{P_i}(1 - x_i) + \tau^w_i x_i$ is the average number of slots per cycle (Ad hoc or WiMAX). The result (4.5) follows by substituting each term by its established expression. □

Remark 4.2.3.2. *It is important to note that a mobile may transmit on one interface and receive on the other, and vice-versa. Moreover, it could get simultaneously one ad hoc and one WiMAX receptions. Whereas we do not allow simultaneous transmissions over the two network interfaces. This is due to the transmission scheduler that we considered (selection of one queue at a time) and the omnidirectional antennas. If we allow parallel selections we could make multihoming possible, i.e., two simultaneous transmissions can be met.*

4.3 Throughput & Stability Region

4.3.1 Rate balace equation

In order to analyze the stability of forwarding queues, we derive the expression of their arrival and departure rates. Let $j_{i,s,d}$ denote the entry in the set $R_{s,d}$ just after the node $i \in \mathcal{C}_a \cup \mathcal{C}_w$, and define the set of neighbors of node i by $\mathcal{N}(i)$. Then, the probability that a transmission from node i on route from node s to node d over ad hoc network is successful is

$$P_{i,s,d} = \prod_{j \in j_{i,s,d} \cup \mathcal{N}(j_{i,s,d}) \setminus i} (1 - \bar{P}_j).$$

The expected number of attempts, per packet, till success or drop from i on route $R_{s,d}$ is

$$L_{i,s,d} = \frac{1 - (1 - P_{i,s,d})^{K_{i,s,d}}}{P_{i,s,d}}.$$

In WiMAX subsystem, the probability that a transmission is successful from a gateway $i \in \mathcal{C}$ (when using a modulation order m-QAM) to the tower B is

$$\varphi_i^m(\overrightarrow{\gamma}_i) = 1 - \prod_{n=1}^{K} \left[1 - \phi_n^m\left(\overrightarrow{\gamma}_i\right)\right]. \tag{4.6}$$

It follows that the expected number of attempts till success or drop from i on route $R_{s,B}$ is

$$L_{i,s,B} = \sum_{k=1}^{K} k \phi_k^m(\overrightarrow{\gamma}_i) \prod_{n=1}^{k-1} \left[1 - \phi_n^m\left(\overrightarrow{\gamma}_i\right)\right] + K \prod_{n=1}^{K} \left[1 - \phi_n^m\left(\overrightarrow{\gamma}_i\right)\right].$$

Let T_i be the average time to serve a packet at node i. Then

$$T_i = \sum_{s,d: i \in R_{s,d}} \pi_{i,s,d} f_i T_{i,s,d} + \sum_{d} (1 - \pi_i f_i) P_{i,d} T_{i,i,d},$$

where $T_{i,s,d}$ is the average time to serve a packet at node i on route $R_{s,d}$, it is given by

$$T_{i,s,d} = \begin{cases} \left\lceil \dfrac{\tau_{i,B}^m}{\tau^a} \cdot \dfrac{M_a}{M_w} \right\rceil L_{i,s,d} & \text{if } i \in \mathcal{C},\ d = B \\ \dfrac{L_{i,s,d}}{P_i} & \text{otherwise.} \end{cases} \tag{4.7}$$

It follows that the long term departure rate from the forwarding queue F_i of node i, for tagged path $R_{s,d}$ can be written as

$$d_{i,s,d} = \frac{\pi_{i,s,d} f_i}{T_i}. \tag{4.8}$$

When the WiMAX tower B transmits to a node d, let us denote by g (for gateway) the first hop on the route $R_{B,d}$. The long term arrival rate at the forwarding queue F_i of node i for the connection $R_{s,d}$ is expressed by

$$a_{i,s,d} = \alpha_{s,d} \prod_{k \in R_{s,i}} \left[1 - (1 - P_{k,s,d})^{K_{k,s,d}} \right], \tag{4.9}$$

where $\alpha_{s,d}$ indicates the successful departure rate of source s own packets. It is given by

$$\alpha_{s,d} = \begin{cases} \frac{P_{s,d}(1-\pi_s f_s)}{T_s} \left[1 - (1 - P_{s,s,d})^{K_{s,s,d}} \right] & \text{if } s \neq B \\ \rho^m_{B,g} . P^m_{g,B,d} \varphi^m_B(\overrightarrow{\gamma}_B) & \text{if } s = B \\ \frac{P_{s,d}(1-\pi_s f_s)}{T_s} \varphi^m_s(\overrightarrow{\gamma}_s) & \text{if } s \in \mathcal{C} \text{ and } d = B. \end{cases}$$

Note that $a_{s,s,d} = 0 \; \forall s$ and d. The end-to-end throughput between a couple of nodes s and d is exactly the arrival rate to the destination d. Namely $thp_{s,d} = a_{d,s,d}$. One can note that it does not depend on forwarding probability of intermediate nodes. The forwarding queue F_i is stable if the departure rate from it is at least equal to the arrival rate into it. We consider the extreme case of equality. This is a simple definition of stability that can be written with a *rate balance equation* (RBE). In the steady-state if all the queues are stable, then for each i, s and d such that $i \in R_{s,d}$ we get $d_{i,s,d} = a_{i,s,d}$. It is written

$$\pi_{i,s,d} f_i = \alpha_{s,d} T_i \prod_{k \in R_{s,i}} \left[1 - (1 - P_{k,s,d})^{K_{k,s,d}} \right]. \tag{4.10}$$

We sum over all sources s and destinations d (include WiMAX tower B) in (4.10) for all connections and get a unique global rate balance which is useful to study general case. It is given by

$$\sum_{s,d} \pi_{i,s,d} f_i = \sum_{s,d:i \in R_{s,d}} \alpha_{s,d} T_i \prod_{k \in R_{s,i}} \left[1 - (1 - P_{k,s,d})^{K_{k,s,d}} \right]. \tag{4.11}$$

Let $z_{i,s,d} = \pi_{i,s,d} f_i$, for all i, s and d, be the unknown of the rate balance system which is a non linear system because $\bar{P}_i$ and $\bar{L}_i$ usually depend on $z_{i,s,d}$. We remark that the solution is independent on the forwarding probability f_i, and consequently the e2e throughput is also not impacted by the scheduling mechanism. This result holds only when the forwarding queues are stable. As a consequence, the forwarding capabilities of gateway nodes do not affect their energy consumption. Hence, the node can fine-tune its forwarding probability f_i to improve the expected e2e delay without affecting the energy consumption or slowing down the throughput.

Another important remark is the interaction between the ad hoc and the WiMAX subsystems. Practically, the bit rate per node in the ad hoc network is locally larger than the WiMAX's, but the size of this former network and its characteristics (as channel access and routing) influence drastically its capacity. To connect two heterogeneous networks one normally needs a scaling and rate adaptation; if not, stability of the gateway nodes becomes a hard issue.

4.3.2 Special case : Uplink connections

Now, we analyze with more details the special case where all connections involve the WiMAX tower. We assume that all mobiles, either covered ones or those outside B range, are using some 802.16x services. This case will help us to develop the stability behavior of the system under asymmetric services. This could be in particular the case of services that do not require request-response mechanism. The RBE of ad hoc to WiMAX connection $R_{s,B}$ is then reduced to the unique following equation

$$\begin{aligned}\pi_{i,s,B} f_i &= \alpha_{s,B} T_i \prod_{k \in R_{s,i}} \left[1 - (1 - P_{k,s,B})^{K_{k,s,B}}\right] \\ &= \frac{T_i}{T_s} y_s \prod_{k \in R_{s,i} \cup s} \left[1 - (1 - P_{k,s,B})^{K_{k,s,B}}\right].\end{aligned}$$

For the uplink case, all nodes have the same destination and hence $P_{s,B} = 1$. This remark is important since it leads to a non-linear RBS. Now we have

$$\begin{aligned}\bar{L}_s &= \sum_{s'} \pi_{s,s',d} f_s L_{s,s',d} + (1 - \pi_s f_s) P_{s,d} L_{s,s,d} \\ &= \pi_s f_s L_{s,s,d} + (1 - \pi_s f_s) L_{s,s,d} = L_{s,s,d}.\end{aligned}$$

It is clear that $\bar{L}_s$ does not depend on π_i. Let us define $\omega_{s,i}$ as follows

$$\omega_{s,i} = \frac{T_i}{T_s} \prod_{k \in R_{s,i} \cup s} \left[1 - (1 - P_{k,s,B})^{K_{k,s,B}}\right].$$

This leads to the following linear system

$$y_i = 1 - \sum_{s \in \mathcal{C}_w \cup \mathcal{C}_a} y_s \omega_{s,i}. \tag{4.12}$$

Then equation (4.12) can be written in a matrix form and then can be resolved easily.

$$\underline{Y}(I + \bar{W}) = \underline{1}. \tag{4.13}$$

If we denote by $|\mathcal{C}_a \cup \mathcal{C}_w| = N$ the whole number of nodes present in the system, then $\bar{W}$ is the $N \times N$ matrix whose $(s,i)^{th}$ entry is $\omega_{s,i}$, $\underline{Y}$ is the unknown $N-$dimensional row vector which contains the stability values for each node, and $\underline{1}$ is a column vector of ones with the appropriate dimension.
We have $y_i = 1 - \pi_i f_i$. The system is as stable as π_i decreases, this is equivalent to have a high value of y_i. As the arrival rate increases, the stability of intermediate nodes becomes a hard issue. We derive some conditions to insure stability and system operability. Classically, a queue is stable when its arrival rate is less (or equal) than its departure rate, namely

$$\pi_{i,s,B} f_i \geq \frac{T_i}{T_s} y_s \prod_{k \in R_{s,i} \cup s} \left[1 - (1 - P_{k,s,B})^{K_{k,s,B}}\right], \quad \forall s.$$

Without loss of generality, let us consider a symmetric mesh network where each node has the same number of active ad hoc neighbors n, same forwarding probability $f_i \equiv f$ and $P_i \equiv P$. For more simplicity assume $K_{i,s,B} = 1$ (delay is minimized) or $K_{i,s,B} = \infty$ (throughput is maximized). As $\pi_{i,s,B} \leq 1$ and if we denote by $|(s, B)|$ the number of intermediate nodes between s and B then we have

$$f \geq \frac{T_i}{T_s} y_s (1 - P)^{n(|(i,B)|+1)}. \tag{4.14}$$

Considering the minimum arrival rate and the maximum departure rate of a node i, after some algebraic manipulations, we have the following proposition.

Proposition 4.3.2.1 (necessary condition)**.** *In a WiMAX and an ad hoc integrated network, the forwarding queues of all nodes are stable if and only if*

$$f \geq \frac{\omega'_{i,s,d}}{\omega'_{i,s,d} + 1}, \tag{4.15}$$

where $\omega'_{i,s,d} = \sum_s \frac{T_i}{T_s} (1 - P)^{n(|(i,B)|+1)}$.

An important interpretation of this proposition is that to insure the stability of the whole system, each node should forward with a minimum forwarding probability. In other words, *the system will not be stable if some nodes do not forward packets of their neighbors*. This problem can be efficiently addressed using the game theoretical tools and considering the forwarding probability f as strategy. We can also rewrite this condition of stability in term of service time as follows

$$T_i \leq \frac{f T_s}{(1 - \pi_s f)(1 - P)^{n(|(i,B)|+1)}}. \tag{4.16}$$

This later result is quite intuitive since the stability of node i is guaranteed whenever its average service time is less than some value depending on individual load and conditions.

4.4 Expected End-to-End Delay

4.4.1 Decomposition method

Here, our purpose is to obtain an expression of the delay of an arrival packet in the forwarding queue F_i at node i with the presence of a saturated source queue Q_i. We are interested in finding the delay function of several parameters belonging to different layers, so it will be easy to study their impact in a hybrid multi-hop wireless network.

Now, we focus our study on the forwarding queue F_i of an intermediate node i. We aim behind this to determine the sojourn time in the buffer. We derived above the arrival rate into F_i (i.e., packets received successfully by node i) and showed that packets

arrive according to a general process with average $a_i = \sum_{s,d:i\in R_{s,d}} a_{i,s,d}$. We note that a_i is exactly the aggregate arrival of packets from different paths and different kind of connections to the forwarding buffer F_i of node i. When a packet leaves the network layer it stays in the MAC layer (server) some arbitrary number of slots. One conclude that the forwarding queue F_i constitutes a $G/G/1$ queue that has some special characteristics due to the presence of Q_i.
Let $\bar{W}_i^w$ (respectively $\bar{W}_i^a$) be the mean waiting time in the forwarding queue F_i of a WiMAX arrival (respectively ad hoc) packet at node i. Also let $\bar{R}_i^w$ (respectively $\bar{R}_i^a$) be the mean residual service time of a WiMAX (respectively ad hoc) packet in MAC layer seen by an arrival. Then we have

$$\bar{W}_i^w = \bar{B}_i + \bar{R}_i^w \quad \text{and} \quad \bar{W}_i^a = \bar{B}_i + \bar{R}_i^a, \tag{4.17}$$

where $\bar{B}_i$ corresponds to the mean time to serve all packets arrived before it (in the buffer).

4.4.2 Mean residual service time

An arrival packet to F_i can find a packet in service corresponding to connection $R_{s,d}$. Let $\bar{R}_{i,s,d}^w$ (resp. $\bar{R}_{i,s,d}^a$) be the mean residual service time of a WiMAX (respectively ad hoc) packet in service.

Lemma 4.4.2.1. *The mean residual time of a packet in service for the connection (s,d) is given by*

$$\begin{aligned}
\bar{R}_i^w &= \sum_{s,d} \pi_{i,s,d} f_i \bar{R}_{i,s,d}^w + \sum_d P_{i,d}(1-\pi_i f_i)\bar{R}_{i,i,d}^w \\
\bar{R}_i^a &= \sum_{s,d} \pi_{i,s,d} f_i \bar{R}_{i,s,d}^a + \sum_d P_{i,d}(1-\pi_i f_i)\bar{R}_{i,i,d}^a,
\end{aligned}$$

where

$$\bar{R}_{i,s,d}^w = \begin{cases} \dfrac{T_{i,s,B}^{(2)}}{2T_{i,s,B}} - \dfrac{1}{2}, & \text{if } i \in \mathcal{C} \text{ and } d = B \\ \dfrac{T_{i,s,d}^{(2)}}{2T_{i,s,d}} + \dfrac{1}{2}, & \text{otherwise} \end{cases} \tag{4.18}$$

$$\bar{R}_{i,s,d}^a = \begin{cases} \dfrac{T_{i,s,B}^{(2)}}{2T_{i,s,B}} - \dfrac{1}{2}, & \text{if } i \in \mathcal{C} \text{ and } s = B \\ \dfrac{T_{i,s,d}^{(2)}}{2T_{i,s,d}} + \dfrac{1}{2}, & \text{otherwise} \end{cases} \tag{4.19}$$

with second moment of service time $T^{(2)}_{i,s,d}$ is given by

$$T^{(2)}_{i,s,d} = \begin{cases} \left\lceil \frac{\tau^m_{i,B}}{\tau^a} \right\rceil^2 L^{(2)}_{i,s,d}, & \text{if } i \in \mathcal{C} \text{ and } d = B \\ \frac{L^{(2)}_{i,s,d} + L_{i,s,d}(1-P_i)}{P_i^2}, & \text{otherwise.} \end{cases} \tag{4.20}$$

Proof. Based on [49] we derive the expression of the residual service time which has two forms: with minus term in (4.18) and (4.19) in presence of two different technologies, and possibility of simultaneous transmission and reception. And plus term in (4.18) and (4.19) for the same technology case. As $\tau^m_{i,B}$ is constant, the second moment of the service time over WiMAX is given by $\left\lceil \frac{\tau^m_{i,B}}{\tau^a} \right\rceil^2 L^{(2)}_{i,s,d}$ and for the ad hoc part it turns to be the same as [49]. □

4.4.3 Waiting time in the buffer

Now, we derive $\bar{B}_i$ as follows: the average service time of the forwarded connections and the node own connections are, respectively, given by

$$\tau^F_i = \sum_{s,d} \frac{\pi_{i,s,d}}{\pi_i} T_{i,s,d} \quad \text{and} \quad \tau^Q_i = \sum_d P_{i,d} T_{i,i,d}. \tag{4.21}$$

If we denote by $\bar{N}^F_i$, the mean number of packets in the queue F_i (without the MAC packet), then we can write

$$\bar{B}_i = \bar{N}^F_i \tau^F_i + (\bar{N}^F_i + 1)\bar{n}^Q_i \tau^Q_i, \tag{4.22}$$

where $\bar{n}^Q_i$ is the mean number of Q_i packets that are served before a packet in the head of queue F_i. After the departure of a forwarding packet, a head of queue F_i packet, if it exists, has to wait V (random variable) number of cycles needed to serve packets from Q_i before it can accesses the MAC layer. The probability to wait k cycles is: $P\{V = k\} = (1-f_i)^k f_i$. This is a geometric distribution with mean f_i. For each node i, the average value of the random variable V is $E[V] \equiv \bar{n}^Q_i \approx \frac{1-f_i}{f_i}$. This is an approximation of n^Q_i since V cannot take very large value in practice. From Little's formula, $\bar{N}^F_i = a_i \bar{W}_i$. Then, by replacing it in equation (4.22), and using formulas (4.17) we have

$$\bar{W}^w_i = \frac{\bar{R}^w_i + \tau^Q_i \frac{1-f_i}{f_i}}{1 - a_i(\tau^F_i + \tau^Q_i \frac{1-f_i}{f_i})}, \tag{4.23}$$

$$\bar{W}^a_i = \frac{\bar{R}^a_i + \tau^Q_i \frac{1-f_i}{f_i}}{1 - a_i(\tau^F_i + \tau^Q_i \frac{1-f_i}{f_i})}. \tag{4.24}$$

4.4.4 End-to-end delay

We add the service time τ_i^F to $\bar{W}_i$. For an additional accuracy, a packet belonging to the path $R_{s,d}$ waits the waiting time in queue F_i plus the service time of its corresponding path, thus $D_{i,s,d} = \bar{W}_i + \tau_{i,s,d}$. As we can see and guess from the expressions in (4.23) and (4.24), the delay is a decreasing function of f_i. In our scheduling mechanism of the WFQ, we can deduce that setting $f = 1$ (full altruist nodes) seems to be the best possible configuration. In fact, with $f = 1$ the delay of forwarding queue is minimized whereas the throughput and the energy consumption remains unchanged. The average e2e delay $D_{s,d}$ of a packet on a path $R_{s,d}$ is the average time from the instant the packet reaches the MAC layer of the source to the instant it is received by the destination. We have derived the average waiting time spent by a given forwarding packet at node i without considering if this packet will be successfully transmitted or dropped at the end of the service in the MAC layer. This delay time is for both successful and dropped packets. However, in the e2e delay formula, the dropped packets due to the finite number of transmissions must not be included in the calculation. Then we have

$$D_{s,d} = \frac{L_{s,s,d}^{succ}}{P_i} + \sum_{i=1}^{|R_{s,d}|} (\bar{W}_i + \tau_{i,s,d}^{succ}), \tag{4.25}$$

where $\tau_{i,s,d}^{succ}$ is the average service time of successfully transmitted packets in this same path $R_{s,d}$. $\tau_{i,s,d}^{succ}$ has the same form as $\tau_{i,s,d}$ and can be written: $\tau_{i,s,d}^{succ} = \frac{L_{i,s,d}^{succ}}{P_i}$, where $L_{i,s,d}^{succ} = \sum_{k=1}^{K_{i,s,d}} k(1 - P_{i,s,d})^{k-1} P_{i,s,d}$ is the average number of attempts till success.

4.5 Numerical examples

We reconsider the same illustrative network in figure 4.1. It is a heterogeneous network composed of one WiMAX cell and an ad hoc network. Gateway nodes maintain the system reliability by forwarding packets towards the two different technologies. They also adapt packet formats for each subnetwork. We consider three connections a, b and c for which we evaluate the e2e throughput when varying some parameters belonging to the two subnetworks. We assume that node 1 can split its crossing WiMAX traffic, i.e., it forwards a fraction δ of the WiMAX traffic to gateway node 7 and transmits directly the remaining fraction $1 - \delta$ to the WiMAX tower. This allows to study the interaction between the two subsystems and illustrates the multihoming capability of gateway nodes. Nodes 3, 4 and 6 cannot reach directly the WiMAX tower, however they can reach there through their neighbors 2 and 5 using their forwarding capability. We assume the ACM with corresponding spectrum efficiencies (rate) in Table 4.1. Gateway 7 experiences good channel conditions and uses 16-QAM scalable modulation. Therefore, it can benefits from relatively high transmission rate. Some parameters of the simulation are summarized in Table 4.2.

Modulation order	Coding order	Target SINR (db)	Spectral efficiency (bits/symbol)
BPSK	1/2	6.4	0.5
QPSK	1/2	9.4	1
QPSK	3/4	11.2	1.5
16-QAM	1/2	16.4	2
16-QAM	3/4	18.2	3
64-QAM	2/3	22.3	4
64-QAM	3/4	24.4	4.5

Table 4.1: *IEEE802.16e Adaptive Coding and Modulation settings.*

Gateway	Subcarriers	Modulation	Coding level
1	n_1	QPSK	1/2
7	n_7	16-QAM	3/4
$\Delta f = 10$KHz, $\tau^a = \tau^w = 2$ms, $M_a = M_w = 1$Kb, $P_1 = 0.5$ $P_7 = P_8 = P_9 = 0$, $P_2 = P_5 = 0.7$, $P_3 = P_6 = 0.4$, $P_4 = 0.3$			

Table 4.2: *Numerical values used for simulation.*

4.5.1 Impact of transmission probability over ad hoc

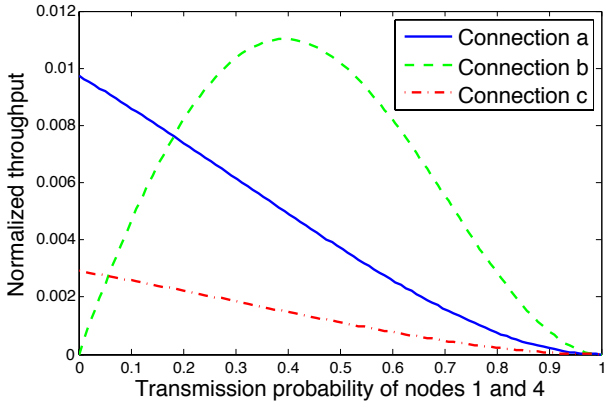


Figure 4.4: *End-to-end throughput versus transmission probability over ad hoc radio for $\delta = 0.5$. Here transmission probabilities $P_1 = P_4 = P$ are variable.*

Figure 4.4 depicts the e2e throughput of all connections as function of transmission probability of nodes 1 (gateway) and 4 (source), where only one single subcarrier is allocated to gateways 1 and 7, i.e., $n_1 = n_7 = 1$. Throughput of connections a and c is strictly decreasing. Indeed, at low transmission probability the collision probability of nodes 1 and 4 are minimized which explain the throughput behavior. Whereas throughput of connection b behaves as a typical slotted aloha scheme, i.e., maximum throughput is met at a moderate transmission probability and it vanishes for very low and very high transmission probability. Another factor is routing that impacts consid-

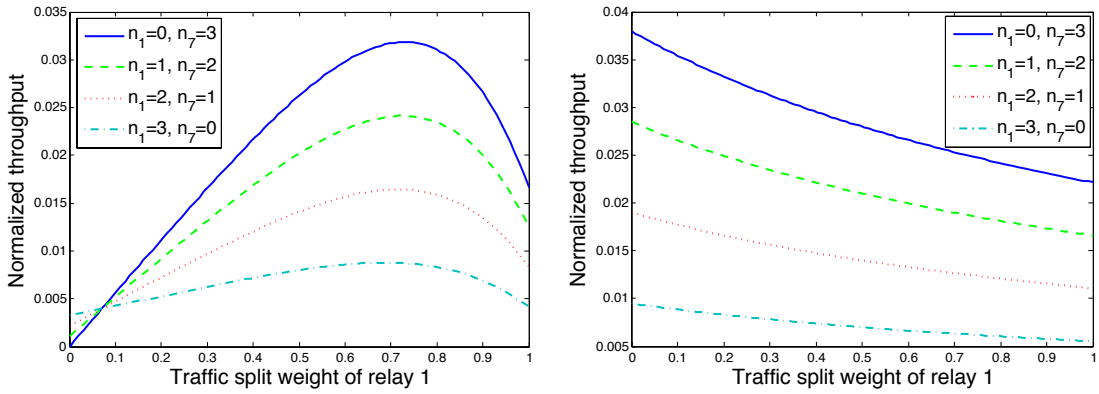
erably the e2e reliability. High performances could be met in a shortest path and low transmission aggressiveness. This can be expressed differently as a path with few number of hops and low transmission probabilities of neighbor nodes of crossing nodes.

4.5.2 Optimal traffic split & subcarriers assignment

Gateway 1 suffers from bad channel conditions, its geographical position allows it to either transmit directly to the WiMAX tower or to forward to another gateway with better channel conditions. It decides to forward a fraction δ of its crossing traffic to gateway 7 and transmits itself the remaining part. We show here the throughput effect when gateway 1 varies the amount of traffic forwarded to gateway 7. Figure 4.5 (a-c) depict the throughput of connections a, b and c under various subcarriers assignment schemes for gateways 1 and 7, other parameters are set as indicated in Table 4.2. For all subcarriers assignment, the throughput of connections a and c seems to be a unimodal function of the traffic amount forwarded to gateway 7. This supports our intuition of existence of an optimal fraction δ^* that achieves a maximum throughput when several routes may exist to reach the same destination. Throughput of connection b is strictly decreasing since other connections consume more and more allocated bandwidth of the link $(7, B)$. However, having better conditions is not a sufficient condition. In fact, existence of optimal split is also depending on the number of assigned resources as well as channel quality of ad hoc subsystem. Yet, we plot in figure 4.6 the throughput with $n_1 = n_7 = 1$ and other parameters are set as indicated in Table 4.2. Here, goodput of the whole connections keeps the same shape as figure 4.5. This says when a gateway suffers from bad channel conditions, WiMAX tower would do better to allocate dynamically available channels for gateways experiencing good channel conditions. This way, optimal subcarriers assignment turns to allocate zero subcarrier to gateway 1 and more bandwidth to gateway 7. This solution will then achieve a maximum goodput by judicious channel allocation as well as reducing the transmission power over both subsystems.

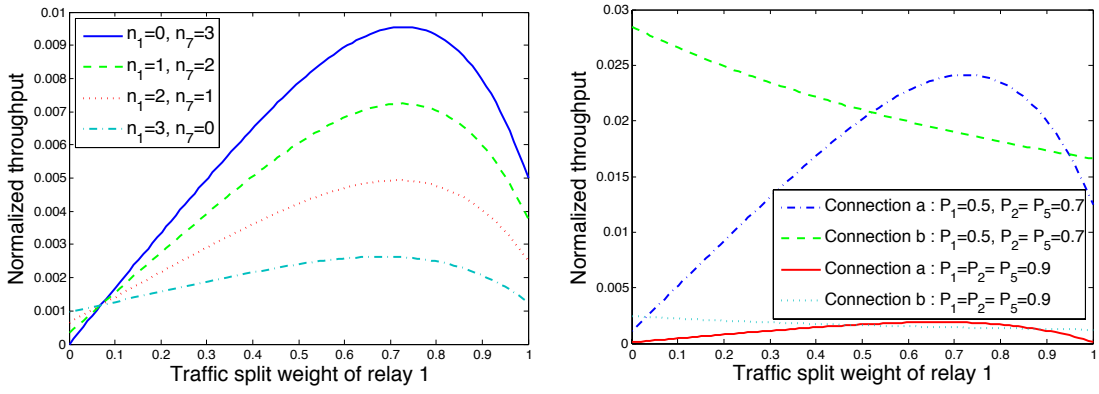
4.6 Concluding discussion

We studied a heterogeneous system composed of a WiMAX cell and an ad hoc network. In fact, hybridizing low cost and high rate technologies is a key towards 4G systems. Our main result is the characterization of the performance in terms of stability, throughput and delay. We also noted that WiMAX parameters do not impact the performance in term of throughput of pure ad hoc nodes and vice-versa. Whereas it may influence drastically the reliability of delay sensitive services. We also studied the impact of splitting WiMAX traffic even in covered area to achieve better connectivity and goodput. This work may have some interest for the future planning of wireless networks in presence of multiple inter-operating wireless technologies.



(Connection a) (Connection b)

Figure 4.5: *End-to-end throughput versus the fraction of traffic δ for all connections under diverse subcarriers assignment schemes. The total number of shared subcarriers is 4.*



(Connection c) (variable transmit probabilities)

Figure 4.6: *End-to-end throughput versus δ for different transmission probabilities of nodes 1, 2, and 5.*

We focused in this chapter on the average end-to-end throughput and delay over the WiMAX/ad hoc integrated network. In realistic systems, the topology is highly asymmetric (nodes density, environment, $\cdots$) and it is continuously changing, then the average delay may not be plausible and decisions based on it may not be optimal. For more accuracy and more rigorous delay derivation, we compute in the next chapter the delay distribution and use it to estimate the expected delay in intermediate relays but also the e2e delay.

Chapter 5

Asymptotic Delay in Wireless Ad hoc Networks with Asymmetric Users

Contents

5.1 Introduction

With the emergence of real-time applications in wireless networks, delay guarantees are increasingly required. In order to provide support for delay sensitive traffic in such network, an accurate evaluation of the distribution of delay is a necessary first step. Knowing the nature of the multihop ad hoc networks, many factors are crucial for the study of the end-to-end (e2e) delay. We cannot study separately the delay generated by a given layer without considering the others. Hence we adopt a cross-layer architecture with its potential synergy of information exchange between different layers, instead of the standard OSI non-communicating layers. Many studies of packet delay and loss in various network environments have been reported in the literature. A large number of studies on multi-hop wireless networks have been devoted to system stability and throughput. The delay performance has been studied under particular topologies (as linear networks or grid networks) or under uniform traffic distribution. It should be pointed out that due to the lack of analytic solutions, many studies of packet delay and

loss behavior have been conducted with simulation and experimental approaches. We provide in this chapter a framework for cross-layer of delay distribution in the context of wireless ad hoc networks. The analysis takes into account the queueing delays at source and intermediate nodes. The delay of a path (we will also refer to it as connection) of this network depends on the number of nodes, the source traffic characteristics, the number of retransmissions at nodes, the forwarding cooperation level and the behavior of the MAC protocol. We assume that time is slotted into fixed length time frames. At any time slot, a node having a packet to be transmitted to one of its neighboring nodes decides with some fixed probability in favor of a transmission attempt. If there is no other transmissions by other nodes those may interfere with the node under consideration, the transmission is successful. As examples of this mechanism, we find Aloha and Carrier Sense Multiple Access (CSMA) type protocols. We consider a parameter that measures the aptitude of a node to forward packets coming from its neighbors. At any instant of time, a node may have two kinds of packets to be transmitted: (1) packets generated by the node itself: data or control packets, and (2) packets from other neighboring nodes those need to be forwarded. To carry these two types of packets, we consider two separate queues handled with a weighted fair queueing (WFQ) discipline. We focus on the asymptotic properties of the delay due to buffering of packets at network layer and random access protocol on MAC layer. The analysis is done using the probability generating function approach which allows us to estimate the distribution of delay at intermediate nodes for all ongoing connections.

Supporting real-time flows with delay and throughput constraints is potentially important for wireless multi-hop networks. From the network designer's point of view, one major concern is how long it will take a packet to reach the destination. In particular, we investigate in this chapter an important issue for real-time multimedia applications in which a large delay and jitter will be unacceptable. Such application requires receiver playback buffers to smooth network delay variation and reconstruct the periodic nature of the transmitted packets. Packets arriving after their scheduled deadline are considered late and are not played out. This requires that the network should be able to offer quality of service (QoS) appropriate for the delay bounds of the real-time application constrains. Our analysis results allow us to study the impact of bounded delay on throughput. We compute the rate of packets arriving before their scheduled playout time (delay constraint). A major focus of the chapter is understanding the impact of multi-hops, the source traffic characteristics, the number of retransmissions at nodes, the forwarding cooperation level and the behavior of the MAC protocol. Based on the analysis, we provide a way to obtain a tradeoff between e2e delay and throughput. Furthermore, we propose a cross layer admission control including the network and the MAC layer to support real-time traffic. This scheme is useful to reduce the loss rate by decreasing the packets arrived after their scheduled deadline.

5.1.1 Main contributions

The main contributions addressed in this chapter are:

- The distributed cross-layer scheme proposed here, besides of its novelty and efficiency, is characterized by its high simplicity. It does not need any external information, but a local decision can be taken with the help of routing information from the network layer as well as the MAC layer.
- We derive a mathematical framework based on probability generating function (PGF) approach to estimate the distribution of delay.
- Unlike [26] and [166], we relax the symmetry assumption. Indeed, our analysis takes into consideration the users and the topology asymmetry. Therefore, each user may have different Network/MAC layer intrinsic parameters (such as attempt rate and cooperation level) and may experience different extrinsic factors (such as the collision probability) due to asymmetric topology and nodes local density.
- In contrast to [45, 49] where the average e2e goodput is calculated based on approximation, we derive here closed form of the e2e goodput and then conclude the exact value of packet admission rate (the probability that the end-to-end delay of a connection does not exceed the timeout delay).
- Getting the distribution of delay for each source/destination connection, we investigate an important issue for real-time and interactive data services (e.g., conversational and streaming flows) over multihop ad hoc networks. A fundamental feature of the streaming service is that the content is played back at the receiver during the delivery. Instead of satisfying a low delay bound as conversational services, streaming services need to maintain a continuous steady flow for smooth playback. In other words, conversational flow has a hard constraint on delay, whereas streaming flow has to solve the jitter problem in addition to the delay relatively soft constraint. In order to play the receiver stream, an application buffers the packets and plays them out after a certain deadline to get again a periodic stream at the application level. Packets arriving after their corresponding delay timeout are lost and then not played out. We further define an admission control to bound the e2e delay and then an acceptable service quality may be guaranteed.

5.1.2 Prior work

In [26], Bianchi model dealt with the behavior of the binary backoff counter at one tagged node as a discrete Markov chain with two-dimensional state. At steady state and based on the remark that each transmission “sees" the system in the same state, computation of the transmission and the collision probabilities becomes possible. Then, he analyzes the saturation throughput under the assumption that in each transmission attempt, regardless of the number of retransmissions, each packet collides with constant and independent probability. Kumar et al. [85] present a fixed point analysis of Bianchi’s model, and give closed form for the collision probability, the aggregate attempt rate, and the aggregate throughput in the asymptotic regime of a large number of nodes. All these studies focus on single-cell WLANs. Yang et al. [166] extend and

characterize the channel activities in IEEE 802.11 DCF-operated multi-hop wireless networks from the perspective of an individual sender and under saturation condition. This is natural since a consistent view for the entire network cannot be symmetric in a multi-hop topology : A node may detect the channel to be busy while another node senses the channel to be idle. Later, they study the impact of the transmit power and the carrier sense threshold on the channel efficiency.

Many papers in the literature have studied the problem of cooperation in ad hoc networks, see [145, 156]. In [49] and [82], authors worked with the above mentioned system model, and studied the impact of routing, channel access rates and weights of the weighted fair queueing on throughput, stability and fairness properties of the network. Important insights were revealed into various tradeoffs that can be achieved by fine-tuning certain network parameters. The throughput maximization of the multi-hop wireless networks has been extensively studied in [66] and [84]. However, it is shown that the high throughput in an ad hoc network is achieved at the cost of a high amount of delay. In [60], the authors characterized the delay-throughput tradeoffs in wireless networks with stationary and mobile nodes. These problems have drawn our attention to the relation between the delay characteristic and the throughput. However, most of the related studies do not consider the problem of forwarding. The authors in [121] contributed to quantifying the impact of hidden nodes on the performance of linear wireless networks based on the IEEE 802.11 protocol and taking into consideration the effects of queueing and retransmissions at each node. In [153], authors provided closed form expressions for the queue length in the presence of arbitrary arrival patterns, packet size distributions and finite network load. In [163], using the decomposition approach authors analyzed the e2e delay of wireless multihop networks for two MAC schemes, m-phase TDMA and slotted aloha, and related references therein. They considered the arrival processes to every node are only relayed versions of the original traffic flow generated at the source node.

5.2 Wireless model

We consider a collection of autonomous nodes able to communicate with other nodes in their respective direct range. Each one can reach nodes those are outside of its direct range by communicating indirectly through intermediate nodes those forward packets towards the required destination. We assume that nodes use the same channel for transmitting with an omnidirectional antenna. A node j receives successfully a packet incoming from a node i if and only if there is no interference at the node j due to another simultaneous transmission. It also follows that a node cannot receive and transmit at the same time slot because of the use of a single channel. Each node i handles two separated buffers: buffer Q_i carries the own packets of i and buffer F_i carries packets originated from a given source, to be forwarded to neighbors till achieving the final destination. Figure 5.1 depicts a typical example the associated double buffering net-

work that can be used to study the multihop ad hoc network. These two queues are

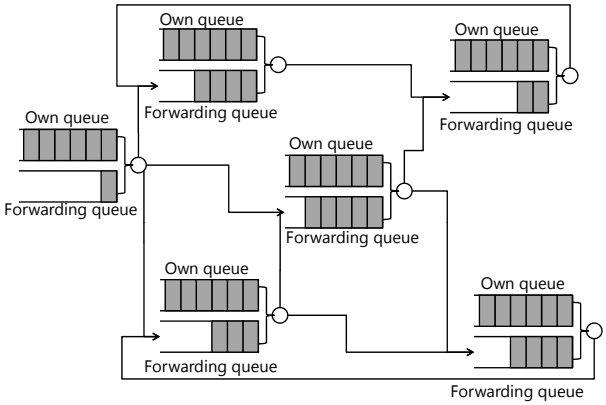


Figure 5.1: *Queuing network model for multihop wireless ad hoc network with double buffering.*

considered to have infinite storage capacity, packets inside are served with a First-In-First-Out fashion and are managed using Weighted Fair Queueing scheduling. The buffer F_i is selected for transmission with probability f_i. Since we assume that each node has always packets to send from queue Q_i, then it follows that queue Q_i is selected with probability $1 - \pi_i f_i$, where π_i is the probability that queue F_i has at least one packet. The forwarding capability permits to each node to behave as a router and this allows to relay packets originated from a source s to a destination d. Routing tables that ensure the network reachability and define which neighbors to use to reach any given destination are periodically updated using a proactive routing protocol as OLSR (Optimized Link State Routing). We use throughout this chapter the notation $R_{s,d}$ to denote the set of intermediate nodes in a path between a source s and a destination d (s and d not included).

MAC layer protocols play doubtlessly the most important role in the communication chain. Consequently, studying medium access methods have got a particular attention by the researchers community from earlier years. This way, many access and resources reservation methods have been elaborated to ensure performance guarantee. There exists two major families of dynamic access methods : Deterministic access such as token ring and token bus, and random access such as aloha and CSMA with all their variants and improved versions. It has been shown that an IEEE 802.11-operated multihop ad hoc network is reasonably equivalent to a multihop ad hoc network operating with slotted aloha protocol. Indeed, this result becomes intuitive by considering the definition of virtual slot, i.e., the mean time (in slots) that the system stays in a given state (idle, busy, success and collision). This way, we only need to be careful about the individual transmission probabilities that should be the solution of Bianchi [26], Kumar et al. [85] and Yang et al [166] fixed point problems. Henceforth, it is plausible to consider a channel access mechanism only based on a probability to access the network, i.e., when a node i has a packet to transmit, it accesses the channel with a probability $\bar{P}_i$. In IEEE 802.11 Distributed Coordination Function based ad hoc system, the attempt

rate is given by formula (4.3). We assume that each node is notified about the success or failure of its transmitted packets. A transmission only fails when there is an interference on the intended receiver, in other terms, when a collision occurs on the receiver. Henceforth, the only source of packet loss is due to collisions.

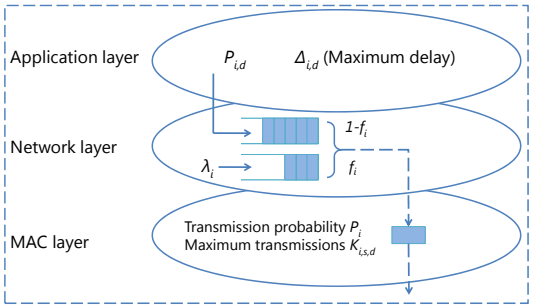


Figure 5.2: *Proposed cross-layer architecture: Accessing the channel starts by choosing the queue from which a packet would be selected. Then, this packet is moved from the network layer to the MAC layer, where it will be transmitted and retransmitted, if needed, until success or definitive drop. This way, it is clear that the e2e QoS (mainly, throughput and delay) depends on several layers as well as the cooperation level f_i of intermediate nodes, i.e., those that play relays role.*

For a reliable communication, we fix a limit number of successive transmissions of a single backlogged packet, after that it will be dropped definitively. We denote $K_{i,s,d}$ the maximum number of successive collisions allowed for a single packet sent from the node i on the path $R_{s,d}$. Unlike the OSI model where layers are clearly separated, we jointly consider network and MAC layer parameters, see figure 5.2. This allows communication and information exchange between different layers and henceforth is more powerful, flexible, allows global optimization and in particular permits manipulating the cooperation level. For ease of reading, we summarize the main notations in Table 5.1.

5.3 Delay distribution analysis

For soft real-time applications, which are delay sensitive but loss tolerant, delay distribution is an important quality of service (QOS) measure of interest. In order to effectively support delay-sensitive applications such as video streaming and interactive gaming in an ad hoc wireless network, it is crucial and challenging to develop feasible methodologies and techniques for accurately analyzing, predicting and guaranteeing the e2e delay performance over multi-hop wireless networks. Our analysis takes into account the queueing delays at source and intermediate nodes of random access multihop wireless ad hoc networks. However the delay is defined as the time taken by a packet to reach the destination after it has left the source. We aim behind this to determine the distributions of the number of forwarding packets and their sojourn times in the system. We focus our study on the forwarding queue F_i of a given node i. In

f_i	Forwarding probability, i.e., the probability that node i picks a packet from queue F_i.
$\bar{P}_i$	Probability that node i attempts a transmission.
$\pi_{i,s,d}$	Probability that the queue F_i has a packet at the first position ready to be forwarded for the path $R_{s,d}$ in the beginning of each transmission cycle.
π_i	Probability that the queue F_i has at least one packet to be forwarded. We have $\pi_i = \sum_{s,d} \pi_{i,s,d}$.
$K_{i,s,d}$	Maximum number of transmissions allowed by a node i per packet of the path $R_{s,d}$.
$P_{i,d}$	Probability that node i transmits its own packet to node d.
$P_{i,s,d}$	Probability that a transmission from node i on the path from s to d is successful.
$\Delta_{s,d}$	Maximum delay that can be tolerated for packets originating from node s to node d.
$L_{i,s,d}$	Expected number of attempts till successful or a drop from node i on the path from s to d.
$\overline{L}_i$	Expected number of attempts till success or definitive drop from node i.
δ	Time slot duration in seconds.

Table 5.1: *Summary of the main notations used in the chapter.*

earlier work [49], authors have shown that packets arrive to F_i (i.e. packets received successfully by node i) according to a random arrival process with average λ_i, it can be written as:

$$\lambda_i = \sum_{s,d:i\in R_{s,d}} \lambda_{i,s,d} = \sum_{s,d:i\in R_{s,d}} \frac{(1-\pi_s f_s)\, P_{s,d}\bar{P}_s}{\overline{L}_s} \prod_{k\in R_{s,i}\cup s} \left[1-(1-P_{k,s,d})^{K_{k,s,d}}\right], \qquad (5.1)$$

where π_s is the probability that queue F_s has at least one packet to be forwarded in the beginning of each cycle[1], $\lambda_{i,s,d}$ is the arrival flow into relay i concerning path $R_{s,d}$ and $P_{s,d}$ is the probability that the node s generates and sends a packet to node d. We distinguish two types of cycles : The forwarding cycles related to the packets of F_i and the source cycles related to the packets coming from Q_i. Moreover, due to asymmetric topology each transmission cycle has a different size for each path. Indeed, the beginning of each cycle represents both the choice of the queue from which we choose a packet and the choice of the transport layer connection where to send it, i.e., the tagged couple of source s and destination d. Whereas, the slots that constitute the cycle represents the attempts of the packet itself to the channel, including its retransmissions in case of collision. An illustrative example of the cycle approach is depicted in figure 5.3. The mean number of attempts till successful or drop from node i on the path $R_{s,d}$ is given by $L_{i,s,d} = \dfrac{1-(1-P_{i,s,d})^{K_{i,s,d}}}{P_{i,s,d}}$, please refer to [49] for detailed derivation.

[1] A cycle is the total number of slots needed to transmit a single packet until success or definitive drop.

Let us denote the mean number of attempts till success or definitive drop by $\overline{L}_s$, it can be computed by averaging over all possible connections

$$\overline{L}_i = \sum_{s,d:i\in R_{s,d}} \pi_{i,s,d} f_i L_{i,s,d} + \sum_d (1 - \pi_i f_i) P_{i,d} L_{i,i,d}, \tag{5.2}$$

The quantity $P_{k,s,d} = \prod_{j\in j_{k,s,d}\cup\mathcal{N}(j_{k,s,d})\setminus k} (1 - \bar{P}_j)$ is the success probability of a transmission initialized by node k for connection $R_{s,d}$, where $j_{k,s,d}$ is the entry in the set $R_{s,d}$ just after k and $\mathcal{N}(j_{k,s,d})$ is the set of neighbor nodes of node $j_{k,s,d}$. In other words, $P_{k,s,d}$ is the product of idle probabilities of all neighbors of the next hop after k on the path from s to d. We note that λ_i is exactly the aggregate arrival rate of packets from different paths and different kind of connections to the forwarding buffer F_i of node i.

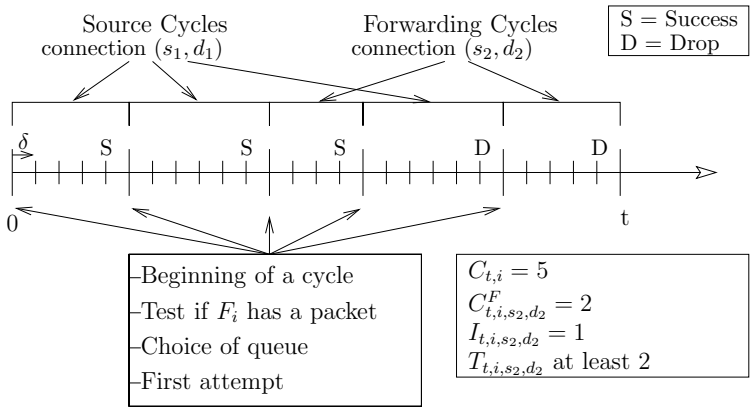


Figure 5.3: *Illustrative example of node i with cycles approach.*

When a packet leaves the network layer, it stays in the MAC layer (server) for some arbitrary number of slots depending on the attempt rate and the collision probability. One conclude that the forwarding queue F_i constitutes a $G/G/1$ queue that has some special characteristics due to the presence of saturated queue Q_i. Furthermore, we will derive the desired distributions using the PGF approach. For sake of simplicity, in the following, we omit the index i that identify the node i itself to facilitate the notations and the reading, e.g., $F_i \equiv F$. Also, the notations indicating connections identities will be omitted until contraindicate. Each connection $R_{s,d}$ has its own service time which depends on the topology (set of neighbors), the transmission probability of nodes and the limit number of transmissions per packet.

- Let r denote the number of arrival packets to the buffer F during the residual service time of a packet which is picked from buffer Q and seen by an arrival packet to buffer F.
- Let a^F denote the number of arrival packets to the buffer F during a service time of a packet picked from forwarding buffer F.

- Let a_j^Q denote the number of arrival packets to the buffer F during the jth packet service time of packets picked from the own buffer Q.
- Let n_i be the instantaneous number of packets in the buffer F.

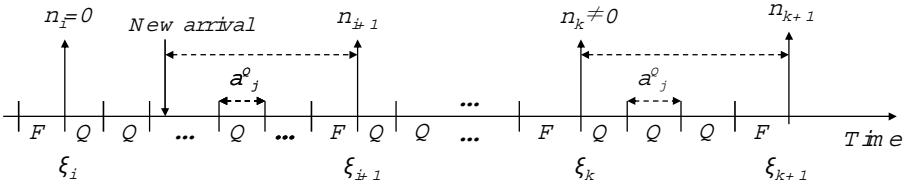


Figure 5.4: *Departure instances from network layer to MAC layer.*

Figure 5.4 shows the evolution of MAC service in terms of cycles (Q and F transmission cycles), it shows also the departure instants ξ_i of forwarding packets in the two cases of the system (for $n_i = 0$ and $n_i \neq 0$). This leads to the following balance equation of the number of packets in F at departure instants:

$$n_{i+1} = \begin{cases} r + \sum_{j=1}^{m} a_j^Q + a^F, & \text{for} \quad n_i = 0 \\ n_i + \sum_{j=1}^{m} a_j^Q + a^F - 1, & \text{for} \quad n_i \neq 0, \end{cases} \tag{5.3}$$

where m represents the number of consecutive packets which are from buffer Q, taking service before the next packet from buffer F. For the case $n_i = 0$, the second term $\sum_{j=1}^{m} a_j^Q = 0$ if there is no packet in the buffer Q (except the residual service packet) which will take for service before the next packet from buffer F. Putting the two cases of n_{i+1} together in one equation, we get

$$n_{i+1} = n_i + rI(n_i) + \sum_{j=1}^{m} a_j^Q + a^F - 1 + I(n_i), \quad \forall i, \tag{5.4}$$

where $I(n_i)$ is an indicator defined by

$$I(n_i) = \begin{cases} 1, & \text{if } n_i = 0; \\ 0, & \text{else.} \end{cases} \tag{5.5}$$

We now focus on the solution of the difference equation (5.4) in the domain of the generating functions to derive all distributions of interest. The following proposition gives the distribution of the length of buffer F_i.

Proposition 5.3.1. *The PGF of the number of packet* $P(z) = \sum_{n=0}^{\infty} P_n z^n$ *in the forwarding queue seen by a departure is given by*

$$P(z) = \frac{P_0(zR(z) - 1)fA^F(z)}{z - (1 - f)zA^Q(z) - fA^F(z)}. \tag{5.6}$$

Proof. See Appendix B. □

For detailed derivation of distributions of new arrivals in a service time denoted by $A^Q(z)$ and $A^F(z)$ and the distribution of new arrivals in residual time $R(z)$, see Appendix A. We now turn to the calculation of how long a packet spends in an intermediate node. From queueing theory, there is a relation between the PGF of the number of packets in the buffer and the PGF of waiting time. Considering a First-In-First-Out fashion, it is clear that the packets left behind are precisely those arrived during its stay in the buffer.

Remark 5.3.2. *Due to the presence of the saturated queue Q_i, the minimum waiting time in the forwarding queue F_i is one time slot. Similarly, the service time is at least one time slot. It follows that the average delay (waiting+service) in the relay node i cannot be less than two time slots.*

Thus, we have

$$\begin{aligned} P(z|t) &= \sum_{n=0}^{t-2} z^n \binom{t-2}{n} \lambda^n (1-\lambda)^{t-2-n} \\ &= (1-\lambda+\lambda z)^{t-2}. \end{aligned}$$

Denote the total time spent in the system for this customer by the random variable D with distribution

$$\begin{aligned} P(z) &= \sum_{t=2}^{\infty} P(z|t)P(D=t) \\ &= \sum_{t=2}^{\infty} (1-\lambda+\lambda z)^{t-2} P(D=t) \\ &= \frac{D\,(1-\lambda+\lambda z)}{(1-\lambda+\lambda z)^2}, \end{aligned}$$

and the next result follows.

Proposition 5.3.3. *The PGF of the waiting time in the system $D(z) = \sum_{n=2}^{\infty} d_n z^n$ is given by*

$$D(z) = z^2 P\left(\frac{z-1+\lambda}{\lambda}\right). \tag{5.7}$$

and then we can come out easily the expression of end-to-end delay PGF using $D_{s,d}(z) = \prod_{i \in R_{s,d}} D_i(z)$.

Lemma 5.3.4. *The expected waiting time and the variance of waiting time at node i are, respectively, given by*

$$D'(1) = 2 + \frac{P'(1)}{\lambda} \tag{5.8}$$

and

$$Var[D] = D''(1) + D'(1) - \left[D'(1)\right]^2, \tag{5.9}$$

where $D''(1) = 2 + \frac{4}{\lambda}P'(1) + \frac{1}{\lambda^2}P''(1)$, where $\phi'(1)$ and $\phi''(1)$, respectively, represent the first and second order derivatives of any PGF $\phi(z)$ at $z = 1$. See Appendix C for detailed derivation of $P_0, P'(1) and P''(1)$. Since we are interested to derive the end-to-end delay of some given connection, we should not consider time elapsed due to dropped packets. Then, we have to deduct it from the total delay (waiting and service times) at node i. The average number of successful transmissions and the average time spent by dropped packet at source s are, respectively, given by

$$L_{s,s,d}^{succ} = \sum_{k=1}^{\widetilde{K}} k \left(1 - P_{s,s,d}\right)^{k-1} P_{s,s,d},$$

and

$$\tau_{i,s,d}^{drop} = \frac{\widetilde{K}\left(1 - P_{i,s,d}\right)^{\widetilde{K}}}{\bar{P}_i},$$

where $\widetilde{K} = K_{i,s,d}$. Finally, the expression of the end-to-end delay of the path $R_{s,d}$ can be derived by

$$\widehat{D}_{s,d} = \frac{L_{s,s,d}^{succ}}{\bar{P}_s} + \sum_{i \in R_{s,d}} \left(D_i'(1) - \tau_{i,s,d}^{drop}\right). \tag{5.10}$$

The first term is the average service time at the source s, whereas the term inside symbol sum is exactly the average waiting and service time in intermediate nodes.

5.4 Application : Playout delay control

Due to its vast potential for providing ubiquitous communication, ad hoc and the emerging mesh networking have received overwhelm interest over the last years. This way, several research works have been done to claim the ability of supporting multimedia applications over ad hoc networks [30] and [154]. In this section, we deal with multimedia applications over ad hoc networks. Our analysis is applicable for both conversational (e.g., VoIP, Gaming, ...) and streaming services (e.g., VoD, TV, ...). Supporting these classes of services over wireless medium is very challenging due to many

factors that cause high error rate. In such interactive multimedia applications, packets loss and connection reliability deterioration are generally caused by delay, jitter (unexpected phase variation) and decoding errors. With a self-managing nodes such as ad hoc networks, the problem becomes more complicated due to the absence of a central entity that monitors the instantaneous changes in the network. In order to enable supporting real-time services, some QoS demand should be stochastically fulfilled (e.g., the average goodput should be strictly guaranteed or a maximum delay should not be exceeded).

The newly designed H.264 video coding standard has been developed such as to support wireless medium [123]. Here, we assume that the ongoing application buffers received packets and plays them out after a given deadline. Henceforth, a packet arriving after the deadline will not be played out. Real-time Transport Protocol (RTP) defines a standardized packet format for delivering audio and video over the Internet. It is used extensively in communication and entertainment systems that involve streaming media, such as telephony, videoconference applications and web-based push to talk features. Each multimedia packet is labeled by a sequence number that is incremented by one for each RTP data packet sent and is to be used by the receiver to detect packet loss and to restore packet sequence. The RTP does not take any action when it sees a packet loss, but it is left to the application to take the desired action. For example, video applications may play the last known frame in place of the missing frame.

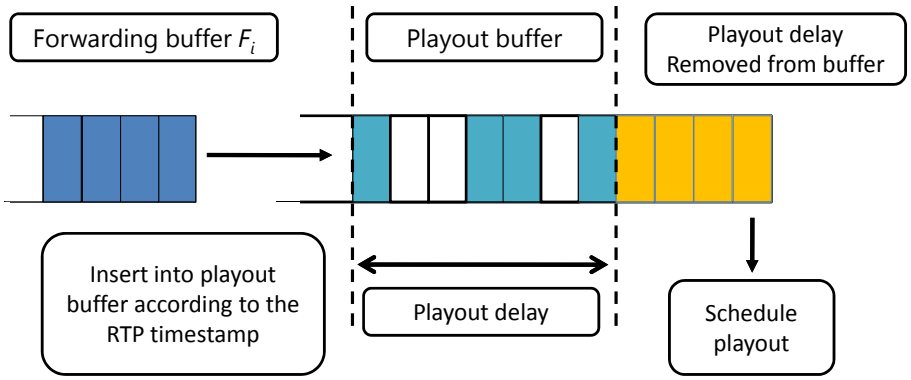


Figure 5.5: *Illustration of the playout process.*

Data packets are extracted from the forwarding queue F_i and inserted into a specific playout buffer sorted by their RTP timestamps[2], see figure 5.5. Frames are held in the playout buffer for a period of time to smooth timing variations generated while crossing intermediate nodes in the network. Holding the data in a playout buffer also allows the pieces of fragmented frames to be received and grouped, and it allows any error correction data to arrive. Potential remaining errors of the frames are then concealed, and

[2]The timestamp is used to place the incoming audio/video packets in the correct timing order (playout delay compensation)

the media is rendered for the user.

5.4.1 End-to-end goodput

If we denote the maximum tolerable delay for connection $R_{s,d}$ by $\Delta_{s,d}$, then the end-to-end goodput (effective throughput) can be written as

$$goodput_{s,d} = thp_{s,d} \cdot P\left(\widehat{D}_{s,d} \leq \Delta_{s,d}\right), \tag{5.11}$$

where $thp_{s,d}$ is the end-to-end throughput without the delay control which is given by $thp_{s,d} = \lambda_{d,s,d}$, see equation (5.1), and $P\left(\widehat{D}_{s,d} \leq \Delta_{s,d}\right)$ is the probability that the accumulative delay does not exceed the application layer threshold $\Delta_{s,d}$, we call it *the end-to-end packets admission rate*. It is given by next lemma.

Lemma 5.4.1.1. *Let $l = |R_{s,d}|$ be the number of intermediate nodes (s and d not included) in route $R_{s,d}$. The probability that the end-to-end delay is exactly j slots is*

$$P\left(D_{s,d} = j\right) = \sum_{j_1=2}^{j-2l+2} \sum_{j_2=2}^{j-2l+2-j_1} \cdots \sum_{j_l=2}^{j-2l+2-j_1-j_2\cdots-j_{l-1}} \prod_{i=1}^{l} P\Big(D_{i,s,d} = j_i\Big), \quad j \geq 2l. \tag{5.12}$$

Proof. We present here a mathematically non-rigorous, but intuitive, derivation of the probability that the e2e delay is exactly j time slots. It is easy to see that $P(\widehat{D}_{s,d} \leq \Delta_{s,d}) = \zeta(\Delta_{s,d})$, where $\zeta(\cdot)$ is the Cumulative Distribution Function (CDF) of the e2e delay, i.e., the probability that the random variable $\widehat{D}_{s,d}$ takes on a value less than or equal to $\Delta_{s,d}$. Since a packet cannot stay less than two slots on an intermediate relay, then the minimum delay on route $R_{s,d}$ after leaving the source s is $2l$, it follows that

$$P\left(\widehat{D}_{s,d} \leq \Delta_{s,d}\right) = \sum_{j=2l}^{\Delta_{s,d}} P\left(D_{s,d} = j\right), \tag{5.13}$$

where $P(D_{s,d} = j)$ is the probability that the e2e delay is exactly j slots. It can be computed considering the set of all partitions of the integer j and taking into account the possible permutations. It follows that

$$P\left(D_{s,d} = j\right) = \sum_{j_1=2}^{j-2l+2} \sum_{j_2=2}^{j-2l+2-j_1} \cdots \sum_{j_l=2}^{j-2l+2-j_1-j_2\cdots-j_{l-1}} \prod_{i=1}^{l} P\left(D_{i,s,d} = j_i\right), \quad j \geq 2l,$$

given that $d_n = P(D_{i,s,d} = n)$ can be calculated by differentiating n times the polynomial $D(z) = \sum_{n=2}^{\infty} d_n z^n$ at $z = 0$, provided by result (5.7).

Example : Let us consider a linear topology with five nodes. Node 1 is the source, nodes 2, 3 and 4 are relays and node 5 is the final destination. Consider the simple scenario

where source 1 forwards to relay 2, relay 2 forwards to relay 3, relay 3 forwards to relay 4 and this latter forwards to the destination 5. The probability that the e2e delay is exactly 8 slots is $P(D_{1,5} = 8) =$

$$P(D_{2,1,5} = 2)P(D_{3,1,5} = 2)P(D_{4,1,5} = 4) + P(D_{2,1,5} = 2)P(D_{3,1,5} = 3)P(D_{4,1,5} = 3)$$
$$+P(D_{2,1,5} = 2)P(D_{3,1,5} = 4)P(D_{4,1,5} = 2) + P(D_{2,1,5} = 3)P(D_{3,1,5} = 3)P(D_{4,1,5} = 2)$$
$$+P(D_{2,1,5} = 3)P(D_{3,1,5} = 2)P(D_{4,1,5} = 3) + P(D_{2,1,5} = 4)P(D_{3,1,5} = 2)P(D_{4,1,5} = 2).$$

□

5.4.2 Distributed Dynamic Retransmissions Algorithm (DDRA)

When a packet is traveling over a multihop ad hoc network, it experiences different transmission success probabilities due to the asymmetric topology. Intuitively, to enhance the multihop reliability and then to improve the success probability on the intermediate relays, each relay should fine-tune its intrinsic network/MAC parameters according to its instantaneous environment perception. This kind of auto-configuration scheme is quite simple but needs a relatively high amount of external information. Here, we describe and extend a simple and fully distributed algorithm that was first described in [48]. We remark that giving more chance to packets that arrive near final destination or the ones whom accumulative delay is less than the threshold, could improve the end-to-end goodput. This can be seen as a cross-layer congestion control that may enhance the e2e admission rate and allows to support multimedia streams. The key idea is to set a different (re)transmissions limit per packet according to the hop sequence number of the relay in the path $R_{s,d}$ while keeping a fixed average transmissions limit $K_{s,d}$ per path, i.e.,

$$K_{s,d} = \frac{\sum_{i \in R_{s,d} \cup s} K_{i,s,d}}{|R_{s,d}| + 1}.$$

Thus, each relay node transmits a forwarded packet with larger number of transmissions limit compared to the previous relay. For instance, the value of $K_{s,d}$ can be defined as the default number of transmissions allowed per packet. Before transmitting to a given path, each source s estimates the total number of hopes to reach the destination (this information is provided by the routing algorithm) and then computes its transmissions limit $K_{s,s,d}$ as well as a fixed step $\eta_{s,d}$. Later, the source node s introduces on each packet of $R_{s,d}$ the values of $K_{s,s,d}$, $\eta_{s,d}$ and the integer h that represents the hop sequence of the next relay ($h = 0$ for s and $h = |R_{s,d}| + 1$ for the final destination). Then, each crossed relay set its transmissions limit as $K_{i,s,d} = \min(K_{s,s,d} + h \cdot \eta_{s,d}, K_{max})$ and update the hop sequence $h = h + 1$, where K_{max} is the maximum transmission limit per packet allowed by the network. Algorithm 4 summarizes how each node should set its retransmission limit for each crossing connection. After some number of iterations, the average queue size (or equivalently the load) of F_i may increase drastically which will induce a huge waiting time (a threshold is to be defined by higher layers). To reduce the appeared huge delay, we propose to integrate the dynamic retransmission scheme

jointly with a *reset* technique, i.e., the $K_{i,s,d}$ of the connection that suffers from huge delay is set to its respective $K_{s,d}$. When congestion is below some threshold, the node can decide to restart the dynamic retransmission scheme again.

Algorithm 4 : Distributed Dynamic Retransmission Algorithm (DDRA)

1: **for** each relay node $i \in R_{s,d}$ **do**
2: Update the retransmissions limit : $K_{i,s,d} = \min\left(K_{s,s,d} + h \cdot \eta_{s,d}, K_{max}\right)$;
3: Update the hop sequence : h=h+1;
4: **end for**

5.5 Numerical examples

We now turn to study a typical example of ad hoc networks. We consider an asymmetric static network formed by 11 nodes as shown in figure 5.6. We established five connections (or streams) a, b, c, d and e. Two nodes are neighbors if they are connected with a dashed or solid line. For illustrative purpose, we consider that the time slot duration is $\delta = 100\mu s$[3] and all nodes are supposed to have the same transmissions limit K per packet. In order to get stability for all nodes, let $\bar{P}_2 = \bar{P}_3 = \bar{P}_7 = \bar{P}_8 = 0.3$, $\bar{P}_4 = \bar{P}_{10} = 0.4$ and $\bar{P}_5 = 0.5$ over all the realized simulations. We present extensive numerical and simulation results to show the accuracy of our method. For that aim, a discrete time simulator that implements the model of Section 5.2 is used to simulate the former network. In order to smooth out the simulation plots, we performed at least 10 runs per simulation and then took the average values.

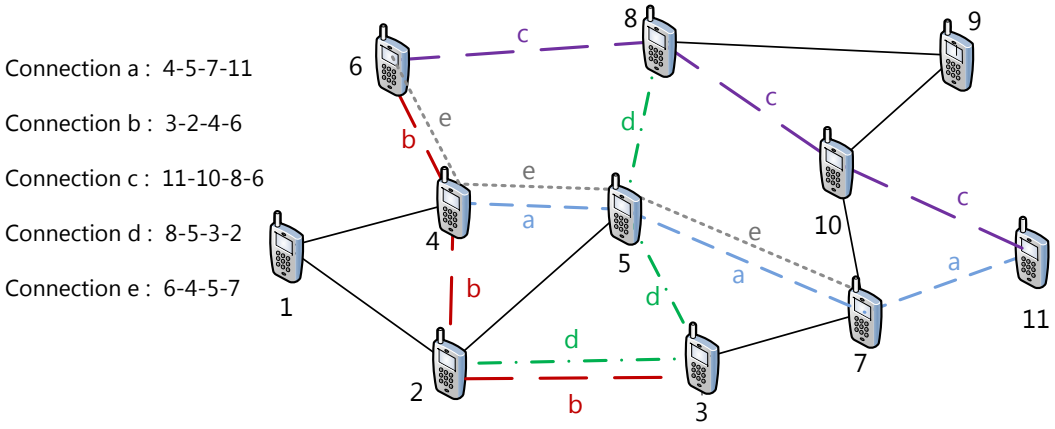


Figure 5.6: *The ad hoc network used for simulation and numerical examples.*

[3] In IEEE 802.11b/g and IEEE 802.11a, the slot duration is $20\mu s$ and $8\mu s$, respectively.

5.5.1 Model validation

All involved nodes are considered to be cooperative and their forwarding probabilities (cooperation levels) are set to $f_i \equiv f \equiv 0.8$. While $\bar{P}_i \equiv \bar{P}$ is varying for all nodes 1, 6, 9 and 11. Figures 5.7 and 5.8, respectively, show (from analytical model as well as simulation results) the average delay in intermediate relays and the average end-to-end delay of considered connections. The results based on our analytical approach are close to the simulation results. This is also true in the case where nodes forward to different neighbors on different paths. However, one can see a sharp gap which is perhaps due to the approximation of the number of consecutive Q_i cycles.

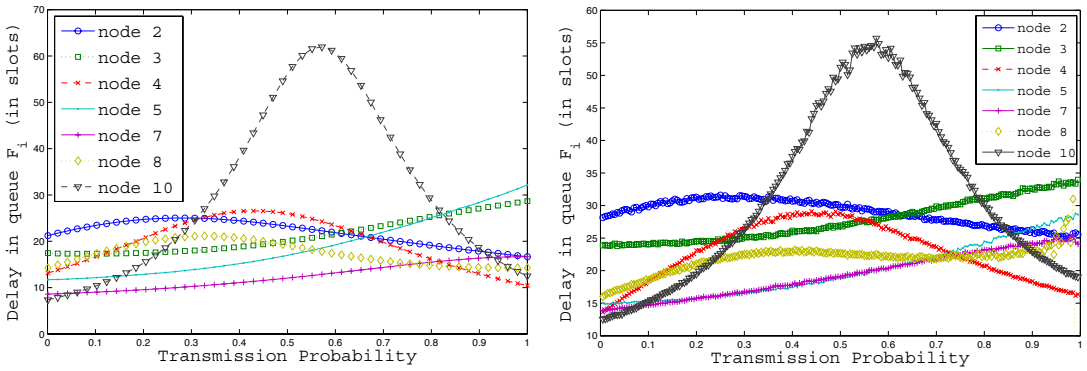


Figure 5.7: *Delay in F_i from analytical model (a) compared to simulation (b) as function of transmission probability for $K = 4$ and $f = 0.8$.*

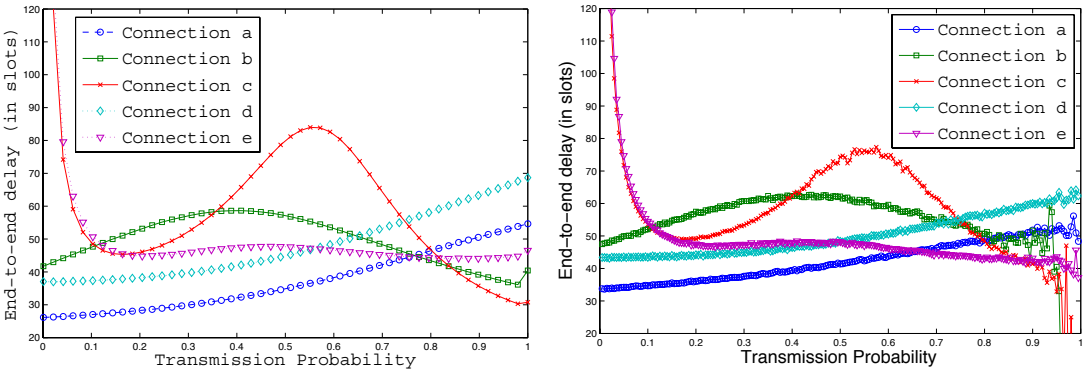
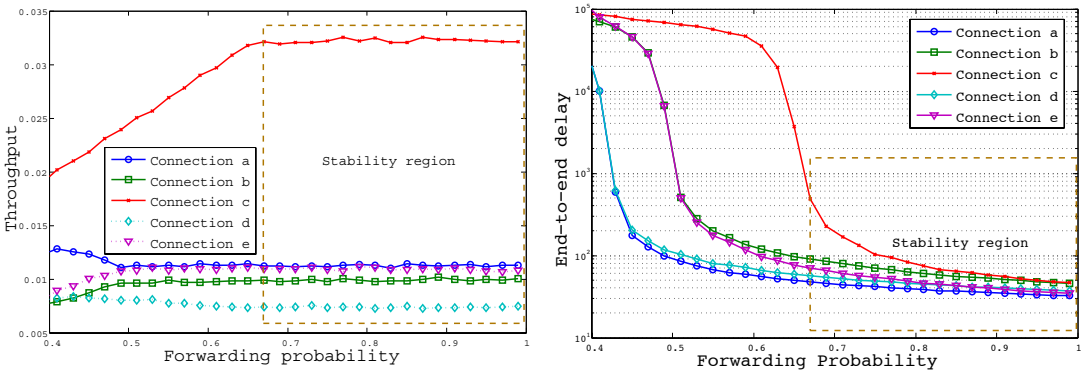


Figure 5.8: *Analytical (a) and simulation (b) end-to-end delay versus transmission probability for $K = 4$ and $f = 0.8$.*

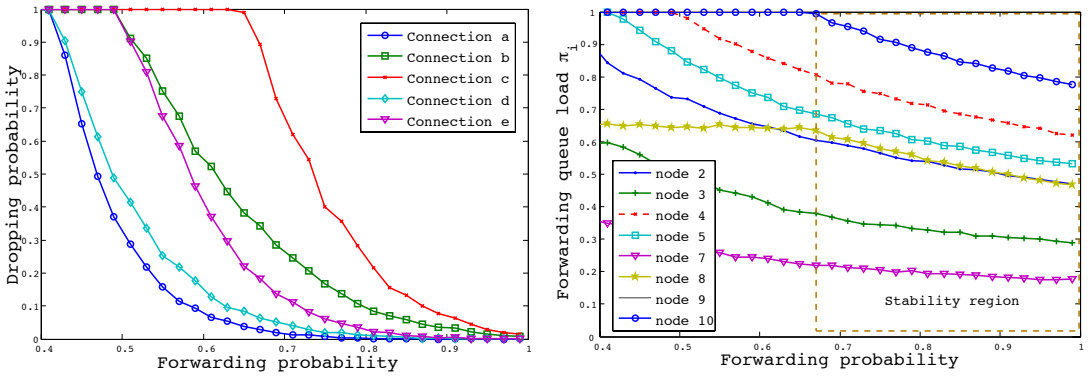


(a) (b)

Figure 5.9: *Average e2e throughput (a) and average e2e delay (b) versus cooperation level for static retransmissions limit* $K = 4$, $\bar{P}_1 = \bar{P}_6 = \bar{P}_{11} = 0.4$ *and* $P_9 = 0.3$.

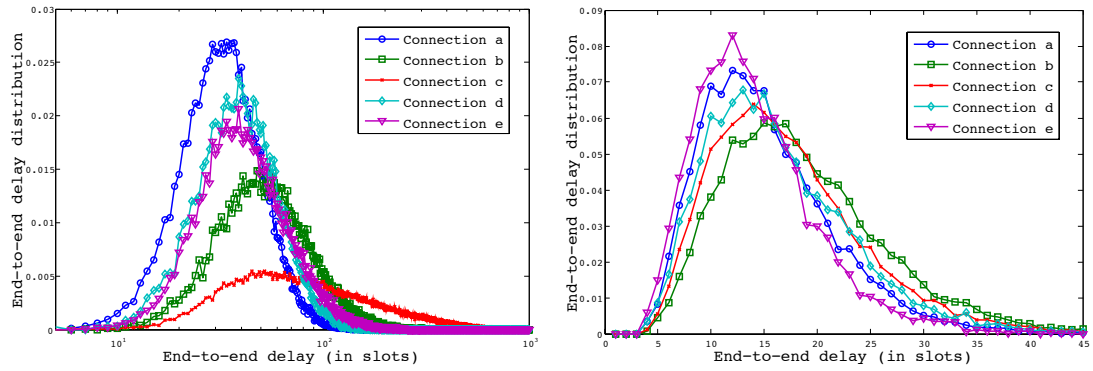
5.5.2 Impact of cooperation level f_i

Here, we address a crucial parameter that impacts significantly the e2e reliability. We fix here the transmission probabilities as follows $\bar{P}_1 = \bar{P}_6 = \bar{P}_{11} = 0.4$, $\bar{P}_9 = 0.3$. For simplicity and without any loss of generality, we consider same cooperation level for all nodes, i.e., $f_i \equiv f$. If nodes were selfish ($f = 0$) then the e2e delay may goes to infinity and the throughput is minimized. When nodes are altruistic ($f = 1$), the e2e delay is minimized and a maximum throughput can be achieved. But when $0 < f < 1$, we note that above some threshold (depending on transmissions probability vector) all forwarding buffers become stable. We refer to this region as the system stability region; therein throughput becomes insensitive to the cooperation level, see figure 5.9 (a). At any fixed transmission probability vector, we come out that the e2e delay is strictly decreasing with f, see figure 5.9 (b). We note here that connection c outperforms other connections in term of throughput. Analyzing the topology, it is clear that connection c has no common segments with other connections. However, it seems that this connection suffers from relatively high delay compared to other connections. This can be explained by high load of forwarding queues in path c, in particular forwarding queue of relay 10, see figure 5.10. Next, we depict the distribution of the average delay in figure 5.11 (a), when cooperation level is set to 0.7. A similar behavior is observed at $f = 0.99$, but the curve is shifted to the left where probability to have small or average delay becomes greater. The delay distribution is as narrow as the cooperation level increases. This imply that the average e2e delay decreases with f whereas the e2e throughput may remain constant. Furthermore, we see clearly that the e2e dropping probability is strictly decreasing with the forwarding probability of intermediate relays, see figure 5.10(a).



(a) (b)

Figure 5.10: *The figure depicts here the e2e dropping probability and the load of forwarding queues versus the cooperation level for static transmissions limit $K = 4$, $\bar{P}_1 = \bar{P}_6 = \bar{P}_{11} = 0.4$ and $P_9 = 0.3$. Giving more weight to F_i decreases the load π_i of node i and henceforth enhances the corresponding stability. The dashed rectangle shows the region where all forwarding queues are stable. In this region, based on QoS requirement, each node can fix its own cooperation level so as to achieve a good throughput/delay tradeoff. The stability region can also be defined as the region where all e2e dropping probabilities are strictly less than 1.*

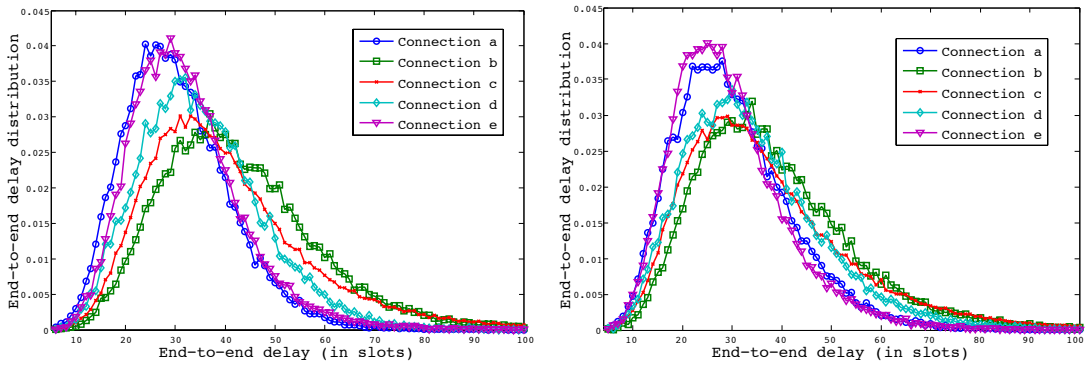


(a) (b)

Figure 5.11: *The cooperation level is set to $f = 0.7$, $\bar{P}_1 = \bar{P}_6 = \bar{P}_{11} = 0.4$ and $P_9 = 0.3$. Subfigures (a) and (b) show the delay distribution for $K = 4$ and $K = 1$ (no retransmissions), repectively.*

5.5.3 Impact of transmissions limit K

Another important factor that impacts e2e performances is the maximum number of transmissions per packet. It is clear that the waiting time does not depend on K whereas the service time depends strongly on it. With $K = \infty$, the throughput is maximized $thp_{s,d} = \dfrac{(1-\pi_s f_s)P_{s,d}\bar{P}_s}{\overline{L}_s}$, see eq. (5.1), corresponding to a huge delay (may goes to infinity due to long service time caused by successive collisions, in particular when neighbor nodes are very aggressive). When $K = 1$, a minimum average throughput is obtained $thp_{s,d} = \dfrac{(1-\pi_s f_s)P_{s,d}\bar{P}_s}{\overline{L}_s} \prod_{k \in R_{s,d} \cup s} P_{k,s,d}$, whereas the delay is minimized. Figure 5.11 (b) and figure 5.12 (a) show the distribution of delay with $f = 0.7$ for $K = 1$ and $K = 4$, respectively. One note that the distribution becomes larger when increasing K.

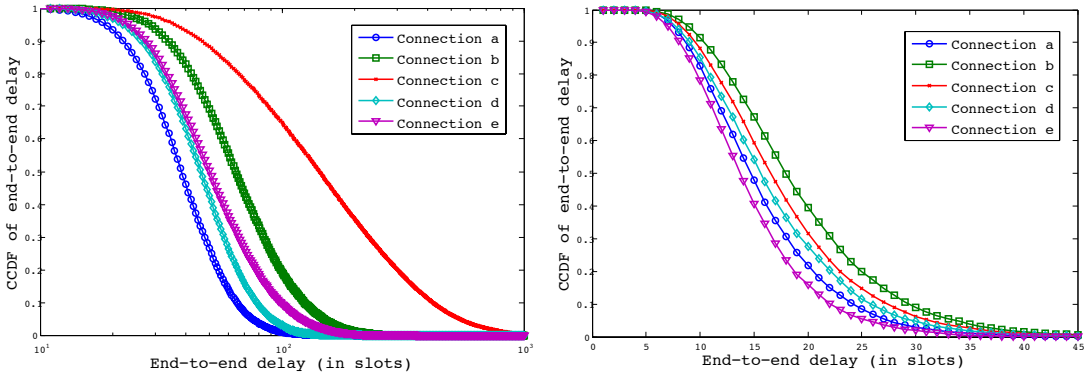


(a) (b)

Figure 5.12: *The cooperation level is set to $f = 0.99$ (altruistic nodes), $\bar{P}_1 = \bar{P}_6 = \bar{P}_{11} = 0.4$ and $P_9 = 0.3$. This figure shows the distribution of e2e delay for static transmissions limit $K = 4$ (a) and dynamic K (with step 2, mean $\bar{K} = 4$ and reset mechanism) (b).*

5.5.4 Static transmission Vs. Dynamic transmissions

Consider a step $K' = 2$, i.e., $K_{i,s,d} = K_{j,s,d} + 2$, where node j is just before node i in route $R_{s,d}$. We note that performances are improved since this new scheme gives more chances to packets arrived near the final destination, see figure 5.12 (b). Indeed, the delay distribution of dynamic case is more narrow than static retransmissions limit under same value of parameters. In contrast, a huge delay may be observed at intermediate nodes and the use of reset mechanism becomes crucial. Using this new routing, we achieve a better average delay (resp. throughput) for each connection without changing the average throughput (resp. delay). In extreme cases, a reset technique is introduced to reduce congestion and optimize the e2e performances. This scheme seems to be very interesting for delay-sensitive traffic.



(a) (b)

Figure 5.13: *To insure stability of forwarding queues, the cooperation level is set to* $f = 0.7$, $\bar{P}_1 = \bar{P}_6 = \bar{P}_{11} = 0.4$ *and* $P_9 = 0.3$. *Subfigures (a) and (b) show the delay complementary cumulative distribution function for* $K = 4$ *and* $K = 1$ *(no retransmissions), respectively.*

5.5.5 A Throughput−Delay tradeoff

From proposed cross-layer point of view, see figure 5.2, the delay at an intermediate node can be written (with some abuse of notation and neglecting the transient effect of other parameters) as delay $(f, k, \bar{P})$ = waiting-time(f) + service-time$(K, \bar{P})$. Using a dynamic transmission scheme and based on figure 5.9 (a and b), one can find an appropriate tradeoff between the throughput and the delay, so that the average delay will be less than some threshold while keeping the average throughput almost constant. This way, making the system running in such a region improves considerably the e2e reliability and makes the system able to support several classes of services with different QoS requirements, in particular real-time traffic. Another way to get an appropriate tradeoff between throughput and delay is to fully use the information of our cross-layer model. For instance, by exploiting the instantaneous length of the forwarding queue, a node may efficiently adjust its cooperation level as well as its maximum number of transmissions per packet.

We depict in figure 5.13 (a) and (b), the complementary cumulative distribution function of e2e delay for different values of K. As stated before, the e2e delay is minimized for small values of K. This is not efficient since it results in low throughput and a loss probability close to 1 when the path length becomes large. Further, we compare the static retransmission and the dynamic retransmission schemes in terms of delay CCDF, see figure 5.13 and figure 5.14. On one hand, we note that the e2e delay of dynamic retransmission scheme is slightly higher compared to the static case. On the other hand, the e2e goodput is improved which confirms the interest of defining a throughput-delay tradeoff .

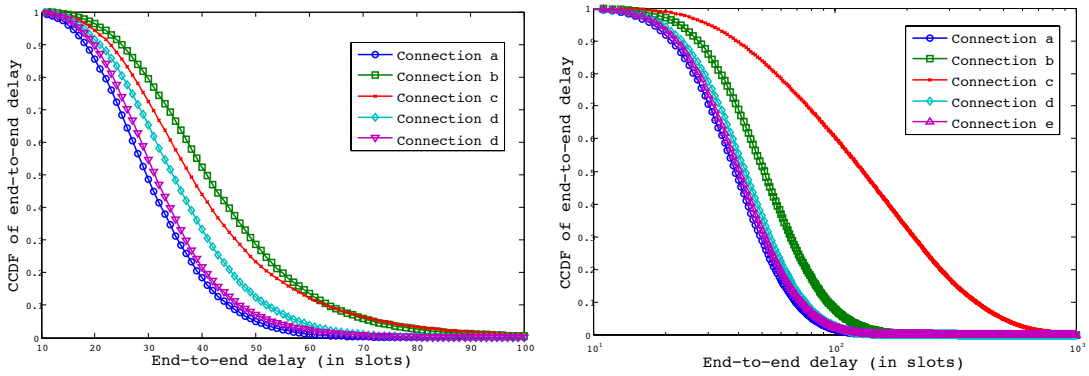


Figure 5.14: *Transmission pobabilities are set to* $\bar{P}_1 = \bar{P}_6 = \bar{P}_{11} = 0.4$ *and* $P_9 = 0.3$. *Subfigure (a) shows the complementary cumulative distribution function of e2e delay for static transmissions limit* $K = 4$ *(with altruistic nodes* $f = 0.99$*), and dynamic* K *(with step 2, mean* $\bar{K} = 4$*,* $f = 0.7$ *and reset mechanism), see subfigure (b).*

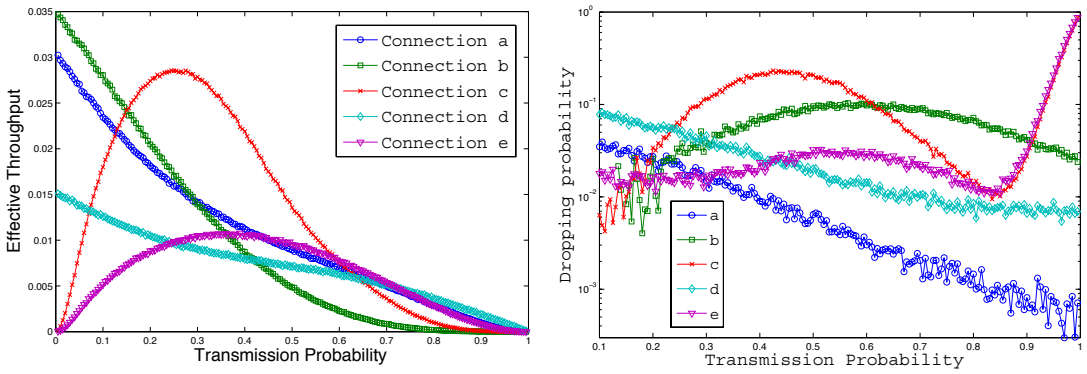
5.5.6 Delay control for real-time media streaming

We depict the variation of average goodput with respect to transmission probability for all established connections. We consider a service requiring a delay threshold value $\Delta_{s,d} \equiv \Delta = 10$ms (100 time slots). This means that a packet arrives after 100 slots is dropped and not played out. The goodput turns to decrease and vanishes when nodes become very aggressive (transmit at probability close to 1). This situation is similar to the well-known prisoners dilemma in game theory where cooperation between players is crucial. This control mechanism causes packets drop and therefore the goodput is deteriorated (figure 5.15 a).

Next, we depict in figure 5.15 (b) the dropping probability for connections a, b, c, d and e as a function of transmission probability. It represents the amount of packets lost due to delay time-out. When we fix the forwarding probability at $f_i \equiv f = 0.8$ and vary the transmission probability, we note a clear correlation with the corresponding e2e delay. Indeed, when average delay is huge, the dropping probability tends to increase and vice-versa. One can note that the dropping probability may increase when the transmission probability goes to 1. This is not due to a huge delay but because of retransmissions expiration. Whereas for fixed transmission probability vector and variable forwarding probability, we note that the admission probability increases with f.

5.6 Concluding discussion

We have presented a framework to derive the end-to-end delay in ad hoc networks taking into account the parameters related to several layers (cross-layer architecture).



(a) (b)

Figure 5.15: *End-to-end goodput (a) and packets dropping probability (b) when cooperation level is set to* $f = 0.8$, $K = 4$ *and* $\bar{P}_1 = \bar{P}_6 = \bar{P}_9 = \bar{P}_{11} \equiv P$ *is varying.*

We have obtained the distribution of the forwarding queue size and then the average delay based on the probability generating function approach. As an application of our results, we have considered the case of real-time traffic which requires delay/jitter constraints. By using the delay analysis we calculate the admission rate/loss rate of this traffic. We also adapted the dynamic retransmission scheme proposed in [48] to improve end-to-end QoS. Preliminary investigations show good match with our experimental illustrations. A part of future guidelines is to address the choice of cooperation level (forwarding probability and retransmission limit) in a game theoretical perspective and analyze the behavior of the selfish nodes. We are also interested in extending our results for wide Mesh networks as well as heterogeneous systems.

Throughout the last two chapters, we considered a slotted aloha based wireless multihop network. This model is quite simple and give good insights to how to fine-tune network parameters for better performances. However unfortunately, it may suffer from some imprecisions mainly due to idle periods in IEEE 802 based operated networks. In the next two chapters, we extend this model to IEEE DCF/EDCF function. Our work extend also the famous models of Bianchi [26], Kumar et al. [85] and Yan et al. [166] to multihop ad hoc networks with asymmetric topology and service differentiation. The modeling and performance evaluation are done using a cross-layer architecture that includes APPLICATION, NETWORK, MAC and PHY layers.

Part III

On Modeling and Understanding IEEE 802.11-Operated Multi-hop Wireless Ad hoc Networks

Chapter 6

A Cross-layered Modeling of IEEE 802.11-Operated Ad hoc Networks

Contents

6.1 Introduction

In next-generation wireless networks, it is likely that the IEEE 802.11 wireless LAN (WLAN) will play an important role and affect the style of people's daily life. People want voice, audio, and broadband video services through WLAN connections. Unlike traditional best effort data applications, multimedia applications require quality of service (QoS) support such as guaranteed bandwidth and bounded delay/jitter. There was a lot of interest in modeling the behavior of the IEEE 802.11 DCF and studying its performances in both the WLAN networks and the multi-hop context. Medium access control protocol has a large impact on the achievable network throughput and stability for wireless ad hoc networks. So far, the ad hoc mode of the IEEE 802.11 standard has been used as the MAC protocols for MANETs. This protocol is based on the CSMA/CA mechanism in DCF (Distributed Coordination Function). There have been a number of studies on the performance of IEEE 802.11 in ad hoc network. All these studies focus on MAC layer without take into account the routing and the cooperation level of nodes in ad hoc networks, see e.g. [6, 7, 22, 23, 102, 157, 158] and [166]. In multi-hop ad hoc

networks, the majority of efforts was concentrated to extend Bianchi's model in saturated or unsaturated network. Now, the problem of hidden terminals and the channel asymmetry are a real issue. A non rare assumption is to consider implicitly symmetric traffic distribution or nodes randomly distributed on a plane following a Poisson point process. Hence, the collision probability and attempt rate are the same for all users. Yang et al. [166] propose an extension of Bianchi [26] and Kumar et al. [85] model and characterize the channel activities from the perspective of an individual sender. They studied the impact of carrier sensing range and the transmit power on the sender throughput. The PHY/MAC impact was clearly considered. Basel et al. [6] were also interested in tuning the transmit power relatively to the carrier sense threshold. They offer a detailed comparison performance between the two-way and the four-way handshake. A three dimensional Markov chain was proposed in [73] to derive the saturation throughput of the IEEE 802.11 DCF. The collision probability is now function of the distance between the sender and its receiver. The unsaturated node state was introduced in the Markov chain in [7] and the channel state too. A performance analysis was performed for a single-hop and a multi-hop case considering that a node can carry different traffic load. Medepalli et al. [102] propose an interesting framework model for analyzing the throughput, the delay and fairness characteristics of IEEE 802.11 DCF multi-hop networks. The applicability of the model in terms of network design is also presented.

In this chapter, we extend Yang's model [166] to an asymmetric ad hoc network in which the nodes have not the same channel vision and then the attempt rate may not always describe the real channel access. Moreover, this model is extended to the IEEE 802.11e which provides differentiated channel access to packets by allowing different back-off parameters. Several traffic classes are supported. We also allow that each traffic may have different limited number of packet transmissions after which it is dropped. From analyzing the model, we find that the performance measures of MAC layer are affected by routing and the intensity of traffic of a connection that an intermediate node forwards. More precisely, the attempt rate and collision probability are now dependent on the traffic flows, topology and routing.

6.2 Problem formulation

6.2.1 Overview on IEEE 802.11 DCF/EDCF

The distributed coordination function (DCF) of the IEEE 802.11 is based on the CSMA/CA protocol in which a node starts by sensing the channel before attempting any packet. Then, if the channel is idle it waits for an interval of time, called the Distributed Inter-Frame Space (DIFS), before transmitting. But, if the channel is sensed busy the node defers its transmission and waits for an idle channel. In addition, to reduce collisions of simultaneous transmissions, the IEEE 802.11 employs a slotted binary exponential back-off where each packet in a given node has to wait for a random number of time

slot, called the back-off time, before attempting the channel. The back-off time is uniformly chosen from the interval $[0, W-1]$, where W is the contention window that mainly depends on the number of collisions experienced by the packet. The contention window W is dynamic and given by $W_i = 2^i W_0$, where i represents the stage number (usually, it is considered as the number of collisions or retransmissions) of the packet, and W_0 is the initial contention window. The back-off time is decremented by one slot each time when the channel is sensed idle, while it freezes if it is sensed busy. Finally, when the data is transmitted, the sender has to wait for an acknowledgement (ACK) that would arrive after an interval of time, called the Short Inter-Frame Space (SIFS). If the ACK is not received, the packet is considered lost and a retransmission has to be scheduled. When the number of retransmissions expires, the packet is definitively dropped.

To consider multimedia applications, the IEEE 802.11e uses an enhanced mode of the DCF called the Enhanced DCF (EDCF) which provides differentiated channel access for different flow priorities. In this manner, the delay to access the channel is reduced for those who have a delay constraint. The main idea of EDCF is based on differentiating the back-off parameters of different flows. So, priorities can be distinguished due to different initial contention window, different back-off multiplier or different inter-frame space. An Arbitration IFS (AIFS) is used instead of DIFS. The AIFS can take at least a value of DIFS, then, a high priority flow needs to wait only for DIFS before transmitting to the channel. Whereas a low priority flow waits an AIFS greater than DIFS. In the next paragraph, we used a generalized model of the back-off mechanism.

Wireless communications using radio signals are subject to many degradation and attenuation specially in ad hoc dynamic environment. The spatial reuse was one of the important issues that impacts the utilization of the bandwidth optimally. It is directly related to the transmit power of nodes, and then to the carrier sensing threshold. This determines the contention degree on the shared channel, and then it is responsible of long back-off freeze. This is the well known problem of exposed terminals. In addition, overlapping of different transmissions at a given receiver may cause packets loss. Collision is mainly due to hidden terminals from the sender, it can occur also due to accumulative transmissions. In sum, simultaneous transmissions in the interference zone of a receiver cause collisions. Assuming that the same frequency for transmitting with an omni-directional antenna is used by all nodes transmit with a fixed power. This assumption can be extended to consider power control each time a node wants to transmit to a different path. On the other hand, the topology of the network is asymmetric, thus now the distribution of nodes is not uniform and the channel attempt is different from a node to another. This also means that collision probabilities are different for each transmission in a given path.

6.2.2 Problem modeling and cross-layer architecture

The network layer of the tagged node i contains two queues are associated. The first one is the forwarding queue F_i which carries packets originated from some source nodes and destined to some destinations. The second one is Q_i which carries own packets of

node i. We assume that each node has an infinite capacity of storage for the two queues. Packets are served with a first in first served fashion. When F_i has a packet to be sent, the node chooses to send it from F_i with a probability f_i, and it chooses to send from Q_i with probability $1 - f_i$. When one of these queues is empty then we choose to send a packet from the non empty queue with probability 1. When node i decides to transmit from the queue Q_i, it sends a packet destined for node d, $d \neq i$, with probability $p_{i,d}$. This latter parameter characterizes somehow the QoS (Quality of Service) required by the initiated service from upper layers. We consider that each node has always packets to be sent from queue Q_i, whereas F_i can be empty. Consequently, the network is considered saturated and mainly depends on the channel access mechanism.

In ad hoc networks, each node may behave as a router. At each time, it has a packet to be sent to a tagged destination and starts by finding the next hop neighbor where to transmit the packet. Clearly, each node must carry routing information before sending the packet. Proactive routing protocols as OLSR (Optimized Link State Routing) construct and maintain a routing table that carries routes to all nodes of the network. To do so, it has to send periodically some control packets. These kind of protocols correspond well with our model, specially that we have considered non-empty Q_i. Nodes form a static network where routes between any source s and destination d are invariant. To consider routing in our model, we denote by $R_{s,d}$ the set of nodes between a source s and destination d (s and d not included). Each node in our model can handle many connections on different paths. The traffic flow leaving a tagged node i is determined by the channel allocation using IEEE 802.11 DCF. However, differentiating the flow leaving F_i and the flow leaving Q_i, allows us to determine the load of and the intensity of traffic crossing F_i. We denote here the probability that the forwarding queue F_i is non-empty by π_i. Similarly, we denote the probability that a packet of the path (s,d) is chosen in the beginning of a transmission cycle by $\pi_{i,s,d}$. This latter quantity is exactly the fraction of traffic related to the path (s, d) crossing F_i, thus $\pi_i = \sum_{s,d:i\neq s} \pi_{i,s,d}$.

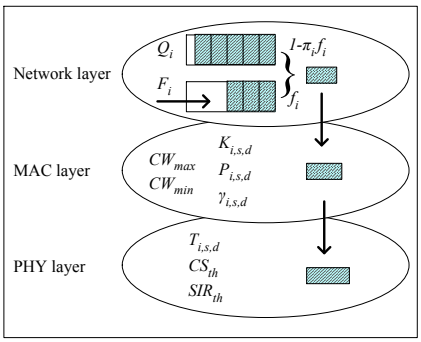


Figure 6.1: *The interaction between NETWORK, MAC and PHY layers is now clear. The end-to-end performance should be studied using a suitable cross-layer model instead of the standard non-communicating OSI model.*

Each node i can be represented with the figure 6.1, with two storing buffers in the network layer and a MAC layer. Attempting the channel begins by choosing the queue from which a packet must be selected. And then, this packet is moved from the corre-

sponding queue from the network layer to the MAC layer where it will be transmitted according to the IEEE 802.11 DCF protocol. In this manner, when a packet is in the MAC layer, it is itself attempted until it is removed from the node. We analyze in the following each layer separately, show how coupled they are and derive the metrics of interest.

6.2.3 PHY layer : Propagation and capture model

Accumulative Interference and virtual node : Let us now define the propagation and interference model in the asymmetric environment. Here we extend the formulas presented in [166]. During a communication between a sender node i and a receiver node j in a given path from s to d (where the source node of a connection is s and the destination node is d), the node i transmits to j with a power $T_{i,s,d}$. The received power on j can be related to the transmitted one by the propagation relation $T_{i,s,d} \cdot h_{i,j}$, where $h_{i,j}$ is the channel gain experienced by j on the link (i,j). In order to decode the received signal correctly, $T_{i,s,d} \cdot h_{i,j}$ should exceeds the receiver sensitivity denoted RX_{th}, i.e., $T_{i,s,d} \cdot h_{i,j} \geq RX_{th}$. Under symmetry assumption and no accumulative effect of concurrent transmissions, the carrier sense range forms a perfect circle with radius r_1. Even when considering accumulative interference, the carrier sense can be reasonably approached by a circle with radius $r_2 \geq r_1$.

Definition 6.2.3.1. We call the group $\mathcal{Z}$, composed of nodes that cannot be heard individually by a tagged sender i but their accumulative signal may jam the signal of interest, a **virtual node**. This way, the effect of a virtual node $\mathcal{Z}$ is equivalent to the behavior of a **fictive node** being in the carrier sense range of tagged sender i.

We can then formulate the carrier sense set of a node i by the following expression

$$CS_i = \left\{ \mathcal{Z} : \forall s,d,k' \in \mathcal{Z}, \begin{array}{c} \sum_{k\in\mathcal{Z}} T_{k,s,d} \cdot h_{k,i} \geq CS_{th} \\ \\ \sum_{k\in\mathcal{Z}\setminus k'} T_{k,s,d} \cdot h_{k,i} < CS_{th} \end{array} \right\} \tag{6.1}$$

where CS_{th} is the carrier sense threshold. One can see CS_i as the set of virtual nodes that may be heard by the tagged sender i when it is sensing the channel in order to transmit on the path $R_{s,d}$. In other words, CS_i is the set of all real nodes (if they are neighbors of i) and virtual nodes (due to accumulative interferences) that may interfere with node i transmission. Now, we define $H_{i,s,d}$ as the set of nodes that may sense the channel busy when the node i is transmitting on the path (s,d). Then

$$H_{i,s,d} = \{k : T_{i,s,d} \cdot h_{i,k} \geq CS_{th}, \forall s,d\}. \tag{6.2}$$

For sake of clarity, we are restricted in our formulation to the case of single transmission power. However, our model can be straightforward used for studying power control

from nodes individual point of views. An interesting feature is that when the transmission power level is the same for all nodes and accumulative interferences are neglected, we have $CS_i = H_{i,s,d}$. *Later result says that under considered assumptions, the set of nodes tagged node i can hear is exactly the set of nodes that can hear node i when transmitting*. The receiver $j_{i,s,d}$ can correctly decode the signal from sender node i if the Signal to Interference Ratio (SIR) exceeds a certain threshold SIR_{th}. Let the thermal noise variance, experienced on the path (s,d), be denoted by $N_{i,s,d}$, then

$$SIR_{j_{i,s,d}} = \frac{T_{i,s,d} \cdot h_{i,j}}{\sum\limits_{k \neq i} T_{k,s',d'} \cdot h_{k,j} + N_{i,s,d}} \geq SIR_{th}, \ \forall s, d, s', d'. \tag{6.3}$$

We define now the interference set of a tagged receiver $j_{i,s,d}$ in a path (s,d), denoted by $\mathcal{T}_{j_{i,s,d}}$, as the collection of its virtual nodes, i.e., all combination of nodes whose the accumulative signal may cause collisions at $j_{i,s,d}$. For instance, the virtual node $\mathcal{Z}$ is in the interference set of node $j_{i,s,d}$ iff the received signal from node i is completely jammed when nodes in $\mathcal{Z}$ are transmitting all together. The interference set of node j is then written as

$$\mathcal{T}_{j_{i,s,d}} = \left\{ \begin{array}{ll} & \dfrac{T_{i,s,d} \cdot h_{i,j}}{\sum\limits_{z \in \mathcal{Z}} T_{z,s',d'} \cdot h_{z,j} + N_{i,s,d}} < SIR_{th}, \\ \mathcal{Z}: & \dfrac{T_{i,s,d} \cdot h_{i,j}}{\sum\limits_{z \in \mathcal{Z} \setminus z'} T_{z,s',d'} \cdot h_{z,j} + N_{i,s,d}} \geq SIR_{th}, \\ & \forall z', s', d', z' \neq i, s' \neq s, \ d' \neq d. \end{array} \right\} \tag{6.4}$$

Remark 6.2.3.2. *We note here that the interference set of a tagged receiver may be of any arbitrary shape. This is due to the heterogeneous node density and the transmit power that may be controlled by each node per path. Indeed, the higher transmit power as a sender uses, the larger the transmission range and then the lower the carrier sense it sets as well.*

Figure 6.2 shows explicitly two different areas that need to be considered when a couple of nodes are communicating. Here, we distinguish (i) the transmission area where two nodes can send and receive packets mutually, (ii) the set of nodes that may hear ongoing transmissions of node i, and (iii) implicitly the carrier sense area where two nodes may hear each other but cannot decode the transmitted data. In figure 6.2, we have situated the communication of i and j on the path (s,d), so we can integrate the impact of the routing in the model. Figure 7.1 illustrates the effect of accumulative interference on transmission cycles of node i. For illustrative purpose, we consider the following virtual nodes : $\{6\}$ and $\{5,7,8\}$. Node 6 is a neighbor of receiver j which causes collision whenever they both, i.e., nodes i and 6, are transmitting simultaneously. Whereas a failure may only occur when nodes of virtual node $\{5,7,8\}$ are all transmitting altogether with tagged sender i.

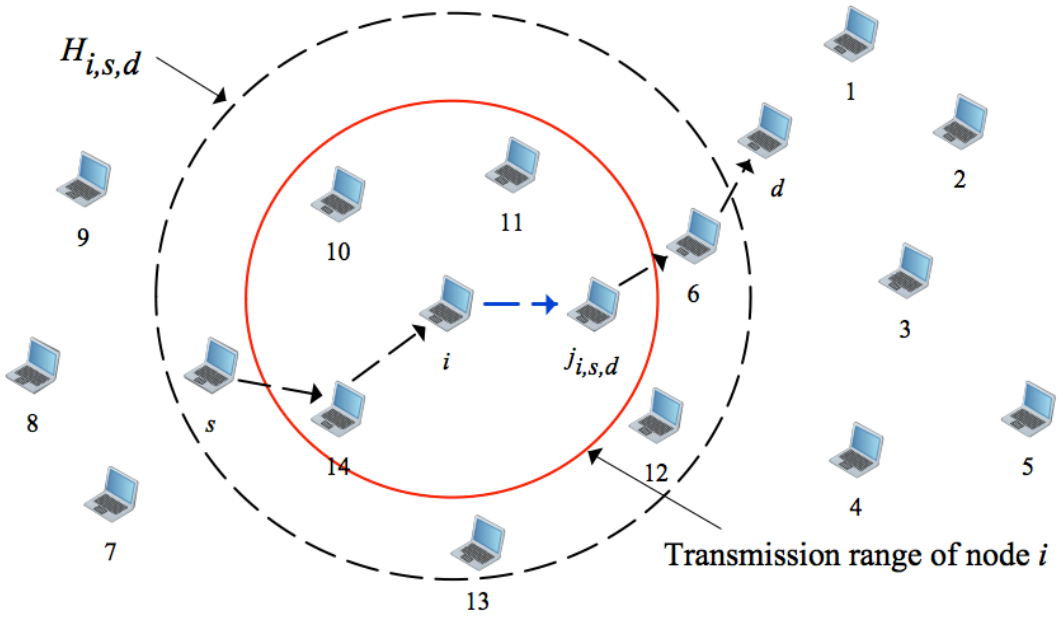


Figure 6.2: *This plot shows the transmission range of node i and the set of real nodes $H_{i,s,d}$ that can hear i when transmitting to node $j_{i,s,d}$. The carrier sense CS_i of node i and the interference set $\mathcal{T}_{j_{i,s,d}}$ are not plotted because they depend on transmit powers of all nodes in the network as well as the topology and scale of the network. For instance $H_{i,s,d} = \{\{s\}, \{j_{i,s,d}\}, \{6\}, \{10\}, \{11\}, \{12\}, \{13\}, \{14\}, \{7,8,9\}, \{d,4\}, \{1,5,7,8,\}, \cdots\}$.*

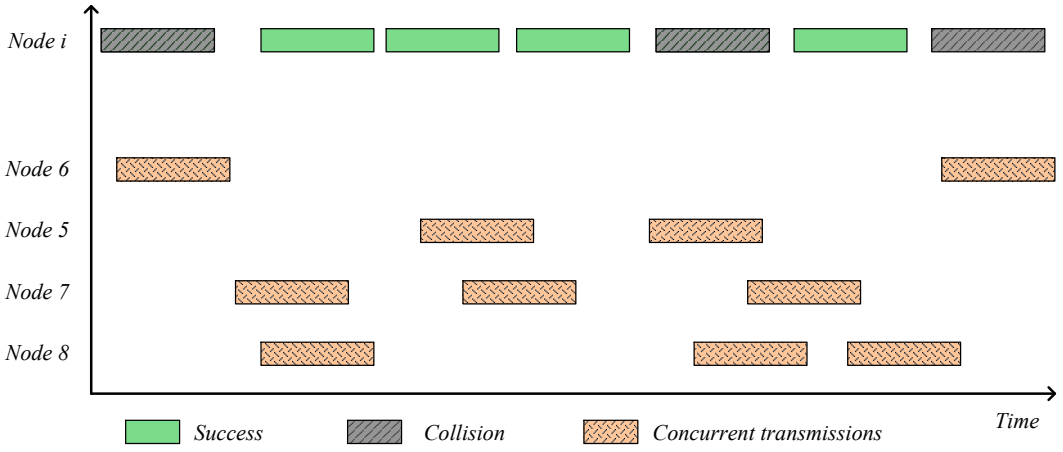


Figure 6.3: *Effect of accumulative interferences on transmission of node i to node $j_{i,s,d}$, see figure 6.2.*

6.2.4 MAC layer modeling

Each node uses the IEEE 802.11 DCF to access the channel and each one can use different back-off parameters. Let $K_{i,s,d}$ be the maximum number of transmissions allowed by a node i per packet on the path (s,d). Then after $K_{i,s,d}$ retries, the packet is automatically dropped. Also let p_i be the back-off multiplier of a given node i. The maximum stage number of node i is obtained from $W_{m,i} = p_i^{m_i} W_{0,i}$, where $W_{m,i}$ and $W_{0,i}$ are, respectively, the maximum and initial contention window for node i. If $K_{i,s,d} < m_i$ then m_i takes the value of $K_{i,s,d}$, otherwise $m_i = log_{p_i}\left(\frac{W_{m,i}}{W_{0,i}}\right)$. Using a contention window $W_{k,i}$ for stage k of node i, the average back-off time for this stage is $b_{k,i}$. Remark that back-off parameters of different nodes may be different. Then, the system of nodes are nonhomogeneous as defined by [119].

We consider the modeling problem of the IEEE 802.11 using the perspective of a sender which consists on the channel activity sensed by a sender, or on the state (success or collision) of its transmitted packet. This will facilitate the problem in the ad hoc environment where nodes have an asymmetric vision of the channel. We start by defining the notion of virtual time slot and channel activity, then we write the expression of the attempt probability for the asymmetric topology. Consider that time is slotted with a physical slot duration τ. Nodes transmit in the beginning of each slot and the transmission duration depends on the kind of the transmitted packet. A data packet has a fixed length and takes $Payload$ (integer) slots to be transmitted (it includes the header transmission time). While an acknowledgment packet spends ACK slots. In our model we consider the two-way handshaking scheme, but it is easily extended to the four-way handshaking scheme. On one hand, a sender node before transmitting would see the channel either **busy** or **idle**. On the other hand, its transmitted packet may encounter a **success** or a **collision**. These four states define all the possibilities that a sender may observe. Therefore, the average time spent in a given state (seen by this sender) will be referred as the **virtual slot** of this sender. A remarkable feature here is that this virtual time would depend on the receiver, i.e., on the path where the packet is transmitted. In fact, the success or the collision of the transmitted packet is itself function of the actual receiver interferences state. For that, we denote by $\Delta_{i,s,d}$ the virtual slot seen by node i on the path (s,d) that we will derive later on. Considering any asymmetric topology, we will always note the metrics functions of the path chosen for transmission. We recall that when we mention the node $j_{i,s,d}$, it will be clear that this latter is the receiver of node i on the path (s,d).

At steady state and such as [26], we use the key assumption which states that at each transmission attempts, and regardless of the number of retransmissions suffered, each packet collides with constant and independent probability. However, collisions may depend only on the receiver channel state. For that we denote by $\gamma_{i,s,d}$ the probability that a transmission of a packet of relay i on the path (s,d) fails due to either a corruption of the data or of its acknowledgment. Thus, $(1-\gamma_{i,s,d})$ is the probability of success in the path (s,d). Henceforth, the attempt probability seen by a sender also depends on the receiver, and the well known formula of [26] can be used in the ad hoc network as

confirmed in [166]. However, in the asymmetric network the attempt probability ($P_{i,s,d}$) (in a virtual slot) for a node i will be different for each path (s,d) and can be written as in [85]:

$$P_{i,s,d} = \frac{1 + \gamma_{i,s,d} + \gamma_{i,s,d}^2 + \cdots + \gamma_{i,s,d}^{K_{i,s,d}-1}}{b_{0,i} + \gamma_{i,s,d} b_{1,i} + \gamma_{i,s,d}^2 b_{2,i} + \cdots + \gamma_{i,s,d}^{K_{i,s,d}-1} b_{K_{i,s,d}-1,i}} \tag{6.5}$$

where $b_{k,i} = (p_i^k W_{0,i} - 1)/2$. In the average, a node i will attempt the channel (for any path (s,d)) with a probability P_i which mainly depends on the traffic and the routing table (here, it is maintained by OLSR protocol). Then

$$P_i = \sum_{s,d:i \in R_{s,d}} \pi_{i,s,d} f_i P_{i,s,d} + \sum_d (1 - \pi_i f_i) p_{i,d} P_{i,i,d}. \tag{6.6}$$

Similarly, the average virtual slot seen by node i is written

$$\Delta_i = \sum_{s,d:i \in R_{s,d}} \pi_{i,s,d} f_i \Delta_{i,s,d} + \sum_d (1 - \pi_i f_i) p_{i,d} \Delta_{i,i,d}. \tag{6.7}$$

Remark 6.2.4.1. *The attempt probability (or attempt rate) must be differentiated from the transmission probability : This latter refers to the probability that a node transmits at any slot. Therefore, the transmission probability, if found, can characterize the channel allocation per node. In WLAN, we usually don't need it and it is sufficient to analyze the back-off rate to determine the channel allocation rate.*

Note that $1 - \pi_i f_i$ is the probability to find a packet from Q_i in the MAC layer. It seems important to note that the attempt probability represents the back-off expiration rate. It is the transmission probability in an idle slot (only when the channel is sensed idle). For that, it is convenience to work with MAC protocols that are defined by only an attempt probability, this kind of definition may englobe both slotted Aloha and CSMA type protocols including IEEE 802.11. The problem in ad hoc is that nodes have not the same channel vision (or different back-off parameters) and then the attempt probability may not always describe the real channel access. In [119], the problem of short term unfairness was studied in the context of a WLAN. We expect that in ad hoc networks, the problem is more crucial. Other subtlety concerning the fairness after a busy period was studied by [77]. Our model will be easily extended to take into consideration this subtlety in accessing the channel.

Collision probability and virtual slot expressions. The collision probability of a packet occurs when either the data or the acknowledgment experiences a collision. If we note by $\gamma_{i,s,d}^D$ and $\gamma_{j_{i,s,d},s,d}^A$, respectively, the collision probability of a data packet and its acknowledgement, then we have

$$\gamma_{i,s,d} = 1 - \left(1 - \gamma_{i,s,d}^D\right)\left(1 - \gamma_{j_{i,s,d},s,d}^A\right), \tag{6.8}$$

The attempt probability of a virtual node $\mathcal{Z}$ is defined by $P_{\mathcal{Z}} = \prod_{z \in \mathcal{Z}} P_z$. Therefore, the virtual slot of a virtual node $\Delta_{\mathcal{Z}}$ can be reasonably estimated using the minimum virtual

slot among all nodes in $\mathcal{Z}$, i.e., $\Delta_{\mathcal{Z}} = \min_{j \in \mathcal{Z}} \Delta_j$. Thus the probability that transmitted data collides with other concurrent transmissions can be written as

$$\gamma^{D}_{i,s,d} = 1 - \prod_{k \in H_{i,s,d} \cap \mathcal{T}_{j_{i,s,d}}} (1 - P_k) \left(1 - \sum_{\mathcal{Z} \in \mathcal{T}_{j_{i,s,d}} \setminus H_{i,s,d}} P_{\mathcal{Z}}^{\frac{Payload}{\Delta_{\mathcal{Z}}}} \right). \tag{6.9}$$

Indeed, nodes in area $H_{i,s,d} \cap \mathcal{T}_{j_{i,s,d}}$ must be silent at the beginning of node i transmission. While nodes in $\mathcal{T}_{j_{i,s,d}} \setminus H_{i,s,d}$ are hidden to i (they constitute the virtual nodes of i) and needs to be silent during all the data transmission time which is a vulnerable time. $\frac{Payload}{\Delta_j}$ is the normalized vulnerable time. After the beginning of data transmission, nodes in $H_{i,s,d}$ will defer their transmission to $EIFS$ (Extended Inter-Frame Space) duration, which would insure the good reception of the acknowledgment. In practice, acknowledgement are small packets and less vulnerable to collision, for that it is plausible to consider $\gamma^{A}_{j_{i,s,d},s,d} \simeq 0$. Then, we can write $\gamma_{i,s,d} = \gamma^{D}_{i,s,d}$.

Considering the previously defined four states and from node i vision, the network stays in a single state a duration equal to $\Delta_{i,s,d}$. Now, the virtual slot, seen by an attempted packet in the path (s,d), Namely

$$\Delta_{i,s,d} = P^{succ}_{i,s,d}.T_{succ} + P^{col}_{i,s,d}.T_{col} + P^{idle}_{i}.T_{idle} + P^{busy}_{i}.T_{busy}, \tag{6.10}$$

$$\text{where} \quad \begin{cases} T_{succ} = Payload + ACK + SIFS + DIFS, \\ T_{col} = Payload + ACK + DIFS, \\ T_{idle} = \tau, \\ T_{busy} = Payload + DIFS, \\ P^{succ}_{i,s,d} = P_{i,s,d}(1 - \gamma_{i,s,d}), \\ P^{col}_{i,s,d} = P_{i,s,d}\gamma_{i,s,d}, \\ P^{idle}_{i} = \prod_{\mathcal{Z} \in CS_i \cup \{i\}} (1 - P_{\mathcal{Z}}), \\ P^{busy}_{i} = (1 - P_i) \sum_{\mathcal{Z} \in CS_i} P_{\mathcal{Z}}. \end{cases} \tag{6.11}$$

Finally, let us denote the equations (6.5), (6.6), (6.8) and (6.10) by *system I*. Normally, it is sufficient to solve the *system I* to derive the fixed points of each node. In fact, by introducing the traffic metric in equations (6.6) and (6.7), these latter equations cannot be solved without knowing the $\pi_{i,s,d}$ which is defined as the traffic intensity for each path (s,d) at each node i. Therefore, in Section 6.3, we proceed in writing the rate balance equations at each node, from which $\pi_{i,s,d}$ can be derived function of P_j and $\gamma_{j,s,d}$, for all j. These rate balance equations that give the traffic intensity will be referred as *system II*. The problem resides in the complexity of the systems and in the computational issue. For that, by considering that the traffic in each node is uniformly distributed, the problem will be simplified. In this case, the system of equations (6.5), (6.6), (6.8) and (6.10) can be solved independently of the rate balance equations.

6.3 End-to-end throughput and traffic intensity system

We are interested in this section to derive the end-to-end throughput per connection, function of different layer parameters, including the IEEE 802.11 parameters. It is clear that the average performances of the system is hardly related to the interaction PHY/MAC/NETWORK. We are now focusing on the traffic crossing the forwarding queues. In fact, it is an issue on the forwarding queues stability. We say that a queue F_i is stable if the departure rate of packets from F_i is equal to the arrival rate into it. This is a simple definition of stability that can be written with a *rate balance equation*. We are going to derive this equation for each node i and each connection (s,d) using the cycle approach. The system of these equations, for all i and (s,d), will form the traffic intensity system. In sum, we are writing a system that determines $\pi_{i,s,d}$ for all i and (s,d). For that, we start by deriving the average length of a transmission cycle, writing the departure rate from F_i and then the arrival rate into it.

6.3.1 Average length of a transmission cycle

A cycle is defined as the number of slots needed to transmit a single packet until its success or drop. It is formed by the four channel states seen by a sender and described previously: idle slots, busy slots, transmissions with collisions and/or a success. We distinguish two types of cycles: the *forwarding cycles* relative to the packets of F_i and the *source cycles* relative to the packets coming from Q_i. Also, each cycle is affected to a connection. The beginning of each cycle represents the choice of the queue from which we choose a packet and the choice of the connection where to send it. Whereas, the slots that constitute the cycle represents the attempts of the packet itself to the channel, including its retransmissions. For each path (s,d), we have different channel activities and nodes behaviors. Based on cycle approach, we handle the back-off evolution and the node transmissions without worried about busy periods. It is also very useful when treating with multi-connections. In fact, the differences of attempting the channel will be hidden. This means that it will not be needed to differentiate between the view of a given node on each path, because cycles do englobe the transmission evolution of the MAC layer. In other words, a transmission cycle may contain idle periods, busy periods, collision periods or/and at most one successful transmission period. Let the random variable l (resp. h, k and t) be the number of idle period (resp. the number of busy period, the number of collision period and the number successful period) in a cycle on the path $R_{s,d}$. Hence the average length in slots of this cycle is given by

$$\hat{C}_{i,s,d} = \sum_{l=0}^{\infty}\sum_{h=0}^{\infty}\left[\sum_{k=1}^{K_{i,s,d}-1}\binom{l+h+k}{l}\binom{h+k}{h}T(l,h,k,1)\cdot P(l,h,k,1) + \binom{l+h+K_{i,s,d}-1}{l}\binom{h+K_{i,s,d}-1}{h}T(l,h,K,0)\cdot P(l,h,K,0)\right], \quad (6.12)$$

where $P(l,h,k,t) = (P_i^{idle})^l (P_i^{busy})^h (P_{i,s,d}^{col})^k (P_{i,s,d}^{succ})^t$ and $T(l,h,k,t) = l\cdot T_{idle} + h\cdot T_{busy} + k\cdot T_{col} + t\cdot T_{succ}$. When a node transmits to several paths, we need to know the average

cycle length, it is almost surely given by

$$C_i = \sum_{s,d:i\in R_{s,d}} \pi_{i,s,d} f_i \hat{C}_{i,s,d} + \sum_d (1-\pi_i f_i) p_{i,d} \hat{C}_{i,i,d}. \tag{6.13}$$

In order to derive the traffic intensity system, we need to consider the following counters:

- $C_{t,i}$ is the number of cycle of the node i till the t^{th} slot, where t slots means t physical slots and it is equivalent to $t \cdot \tau$ seconds with $\tau = 20\mu s$ in the IEEE 802.11.
- $C^F_{t,i}$ (resp. $C^Q_{t,i}$) is the number of all *forwarding cycles* (resp. *source cycles*) of the node i till the t^{th} slot.
- $C^F_{t,i,s,d}$ (resp. $C^Q_{t,i,s,d}$) is the number of *forwarding cycles* (resp. *source cycles*) corresponding to the path $R_{s,d}$ of the node i till the t^{th} slot.
- $T_{t,i,s,d}$ is the number of times we found at the first slot of a cycle and at the first position in the queue F_i a packet for the path $R_{s,d}$ of the node i till the t^{th} slot.
- $I_{t,i,s,d}$ is the number of cycles corresponding to the path $R_{s,d}$ of the node i, where a cycle is ended by a success of the transmitted packet till the t^{th} slot.
- $A_{t,i,s,d}$ is the number of arrival packets to node i on the path $R_{s,d}$.

6.3.2 Departure rate

The departure rate from F_i is the probability that a packet is removed from node i forwarding queue by either a successful transmission or a drop after successive $K_{i,s,d}$ failures. The departure rate concerning only the packets sent on the path $R_{s,d}$ is denoted by $d_{i,s,d}$. Formally, for any node i, s and d such that $p_{s,d} > 0$ and $i \in R_{s,d}$, the long term departure rate of packets from node i on the route from s to d is given by the following proposition:

Theorem 6.3.2.1. The long term departure rate from node i related to path $R_{s,d}$ is given by

$$d_{i,s,d} = \frac{f_i \pi_{i,s,d}}{C_i}. \tag{6.14}$$

Proof. The long term departure rate of packets from node i on the route from s to d is, by definition, given by

$$d_{i,s,d} = \lim_{t\to\infty} \frac{C^F_{t,i,s,d}}{t} = \lim_{t\to\infty} \frac{T_{t,i,s,d}}{C_{t,i}} \cdot \frac{C^F_{t,i,s,d}}{T_{t,i,s,d}} \cdot \frac{C_{t,i}}{t}. \tag{6.15}$$

- $\lim_{t\to\infty} \frac{T_{t,i,s,d}}{C_{t,i}}$ is the probability that F_i carries a packet to the path $R_{s,d}$ at the beginning of each cycle. Therefore $\lim_{t\to\infty} \frac{T_{t,i,s,d}}{C_{t,i}} = \pi_{i,s,d}$.

- $\lim_{t\to\infty} \frac{C^F_{t,i,s,d}}{T_{t,i,s,d}}$ is exactly the probability that we have chosen a packet from F_i to be sent when F_i carried a packet to the path $R_{s,d}$ in the first position and in the beginning of a forwarding cycle. Therefore, $\lim_{t\to\infty} \frac{C^F_{t,i,s,d}}{T_{t,i,s,d}} = f_i$.

- $\lim_{t\to\infty} \frac{t}{C_{t,i}}$ is the average length in slots of a cycle of the node i.

□

Hence, it is easy to derive the total departure rate d_i on all paths:

$$d_i = \sum_{s',d':i\in R_{s',d'}} d_{i,s',d'} = \frac{\pi_i f_i}{C_i}.$$

6.3.3 Arrival rate and end-to-end throughput

The probability that a packet arrives to the queue F_i of the node i is also called the arrival rate, we denote it by a_i. When this rate concerns only packets sent on the path $R_{s,d}$, we denote it by $a_{i,s,d}$. Formally, for any nodes i, s and d such that $p_{s,d} > 0$ and $i \in R_{s,d}$, the long term arrival rate of packets into F_i for $R_{s,d}$ is provided by the following

Theorem 6.3.3.1. The long term arrival rate into node i forwarding queue, related to path $R_{s,d}$, is given by

$$a_{i,s,d} = (1 - \pi_s f_s) \cdot \frac{p_{s,d}}{C_s} \cdot \prod_{k\in R_{s,i}\cup s} \left(1 - \gamma_{k,s,d}^{K_{k,s,d}}\right). \tag{6.16}$$

Proof. Using the transmission cycles method, the long term arrival rate of packets into F_i for $R_{s,d}$ is

$$a_{i,s,d} = \lim_{t\to\infty} \frac{A_{t,i,s,d}}{t} = \lim_{t\to\infty} \frac{C^Q_{t,s}}{C_{t,s}} \cdot \frac{C^Q_{t,s,s,d}}{C^Q_{t,s}} \cdot \frac{C_{t,s}}{t} \cdot \frac{I_{t,s,s,d}}{C^Q_{t,s,s,d}} \cdot \frac{A_{t,i,s,d}}{I_{t,s,s,d}}, \tag{6.17}$$

- $\lim_{t\to\infty} \frac{C^Q_{t,s}}{C_{t,s}} = 1 - \frac{C^F_{t,s}}{C_{t,s}} = 1 - \pi_s f_s$ is exactly the probability to get a source cycle, i.e., to send a packet from the queue Q_s.

- $\lim_{t\to\infty} \frac{C^Q_{t,s,s,d}}{C^Q_{t,s}}$ is the probability to choose the path $R_{s,d}$ to send a packet from Q_s. Therefore, $\lim_{t\to\infty} \frac{C^Q_{t,s,s,d}}{C^Q_{t,s}} = p_{s,d}$.
- $\lim_{t\to\infty} \frac{C_{t,s}}{t} = \frac{1}{C_s}$.
- $\lim_{t\to\infty} \frac{I_{t,s,s,d}}{C^Q_{t,s,s,d}}$ is the probability that a source cycle on the path $R_{s,d}$ ends with a success, i.e., the packet sent from Q_s is received on the queue $F_{j_{s,s,d}}$. Therefore, $\lim_{t\to\infty} \frac{I_{t,s,s,d}}{C^Q_{t,s,s,d}} = 1 - \gamma^{K_{s,s,d}}_{s,s,d}$.
- $\lim_{t\to\infty} \frac{A_{t,i,s,d}}{I_{t,s,s,d}}$ is the probability that a packet received on the node $j_{s,s,d}$ is also received on the queue F_i of the node i. For that, this packet needs to be received by all the nodes in the set $R_{s,i} \cup s$. Therefore, $\lim_{t\to\infty} \frac{A_{t,i,s,d}}{I_{t,s,s,d}} = \prod_{k\in R_{s,i}\cup s} \left(1 - \gamma^{K_{k,s,d}}_{k,s,d}\right)$.

□

End-to-end throughput : The global arrival rate at F_i is $a_i = \sum_{s,d:i\in R_{s,d}} a_{i,s,d}$. Remark that when the node i is the final destination of a path $R_{s,d}$, then $a_{d,s,d}$ represents the end-to-end average throughput of a connection from s to d. Practically, $a_{d,s,d}$ is the number of delivered (to destination) packet per slot. Let ρ be the bit rate in bits/s of the wireless network. Therefore, the throughput in bits/s can be written as following:

$$thp_{s,d} = a_{d,s,d} \cdot Payload \cdot \rho. \tag{6.18}$$

6.3.4 Rate balance equations: traffic intensity system

Finally, in the steady state if all the queues in the network are stable, then for each i, s and d such that $i \in R_{s,d}$ we get $d_{i,s,d} = a_{i,s,d}$, which is the rate balance equation on the path $R_{s,d}$. For all i, s and d we get the traffic intensity system: *system II*. And when we sum the both sides of this last system, we get the global rate balance equation: $d_i = a_i$.

Let $y_i = 1 - \pi_i f_i$ and $z_{i,s,d} = \pi_{i,s,d} f_i$. Thus $y_i = 1 - \sum_{s,d:i\in R_{s,d}} z_{i,s,d}$. Then, the rate balance equation can be written in the following form:

$$\sum_{d:i\in R_{s,d}} z_{i,s,d} = \frac{y_s(\sum_{s',d'} z_{i,s',d'}\hat{C}_{i,s',d'} + \sum_{d''} y_i p_{i,d''}\hat{C}_{i,i,d''})w_{s,i}}{(\sum_{s',d'} z_{s,s',d'}\hat{C}_{s,s',d'} + \sum_{d''} y_s p_{s,d''}\hat{C}_{s,s,d''})}, \tag{6.19}$$

where $w_{s,i} = \sum_{d:i\in R_{s,d}} p_{s,d} \prod_{k\in R_{s,i}\cup s} \left(1-\gamma_{k,s,d}^{K_{k,s,d}}\right)$.

An interesting interpretation and application of equation (6.19) are the following : (i) $z_{i,s,d}$ and y_i (can be considered as the stability region of node i) are independent of the choice of f_i. (ii) For some values of f_i the forwarding queue of node i will be stable. Concerning P_i, we notice that it can be written as $P_i = \sum_{s,d:i\in R_{s,d}} z_{i,s,d}P_{i,s,d} + \sum_d y_i p_{i,d} P_{i,i,d}$. Then it depends on $z_{i,s,d}$ and y_i, but it is not affected by f_i. A similar deduction is also observed for the energy consumed when sensing the channel or transmitting data. It turns out to be independent of the choice of cooperation level f_i. Hence, the node can fine-tunes f_i to improve the expected delay without affecting the throughput or the energy consumption.

6.3.5 Resolving PHY/MAC/NETWORK coupled problems

As we showed previously, the MAC layer systems of fixed points and the Network layer rate balance systems (non linear systems) could not be resolved separately. Moreover, due to dependance on topology, routing and users behaviors, we cannot show analytically existence of a unique solution of the fixed point systems. However, for several scenarios and network topologies, *system I* and *system II* always provide the same solution as obtained from simulation. We give here a sketch of the algorithmic way we follow to solve mutually the above systems (including the correlation between layers).

Algorithm 5 : Fixed point and rate balance resolution

Require: $\pi_{i,s,d}^0 = \epsilon_{i,s,d}$, δ : convergence indicator

1: **for** each source s, relay i and destination d **do**
2: **while** $\left|\frac{\pi_{i,s,d}^{t+1} - \pi_{i,s,d}^t}{\pi_{i,s,d}^t}\right| \geq \delta$ **do**
3: Compute $P_{i,s,d}$ using fixed point such as [85]
4: Update $\gamma_{i,s,d}$ using equation (6.9)
5: Estimate cycles size using equation (6.13)
6: Update $\pi_{i,s,d}^{t+1}$ by solving the rate balance system (6.19) using for example the Gaussian elimination method
7: **end while**
8: **end for**

6.3.6 Energy consumption

Let $E_{i,s,d}$ be the expression of the energy consumed per cycle by each node on the path (s,d). Let also E_i^s be the energy consumed per (virtual) slot in sensing the channel, and $E_{i,s,d}^{tx}$ be the energy consumed per transmission of a single packet on the same path (s,d). Therefore, we can derive $E_{i,s,d}$ from the average cycle length of equation (6.12) as

following:

$$E_{i,s,d} = \sum_{l=0}^{\infty}\sum_{h=0}^{\infty}\left[\sum_{k=1}^{K_{i,s,d}-1}\binom{l+h+k}{l}\binom{h+k}{h}E(l,h,k,1)\cdot P(l,h,k,1) + \binom{l+h+K_{i,s,d}-1}{l}\binom{h+K_{i,s,d}-1}{h}E(l,h,K,0)\cdot P(l,h,K,0)\right], \quad (6.20)$$

where $P(l,h,k,t) = (P_i^{idle})^l (P_i^{busy})^h (P_{i,s,d}^{col})^k (P_{i,s,d}^{succ})^t$ and $E(l,h,k,t) = l \cdot T_{idle} \cdot E_i^s + h \cdot T_{busy} \cdot E_i^s + k \cdot T_{col} \cdot E_{i,s,d}^{tx} + t \cdot T_{succ} \cdot E_{i,s,d}^{tx}$. This quantity turns out to be independent of the choice of f_i. Hence, the node can use f_i to improve the expected delay without affecting the energy consumption. Note that the value $\pi_{i,s,d} f_i E_{i,s,d}$ represents the energy consumption used by node i to forward packets to path (s,d). It is clear that one can reduce the amount of forwarded packets at the expense of queue average load. The average energy spent per transmission cycle is almost surely given by

$$E_i = \sum_{s,d:i\in R_{s,d}} \pi_{i,s,d} f_i E_{i,s,d} + \sum_d (1-\pi_i f_i) p_{i,d} E_{i,i,d}. \quad (6.21)$$

6.4 Numerical discussion

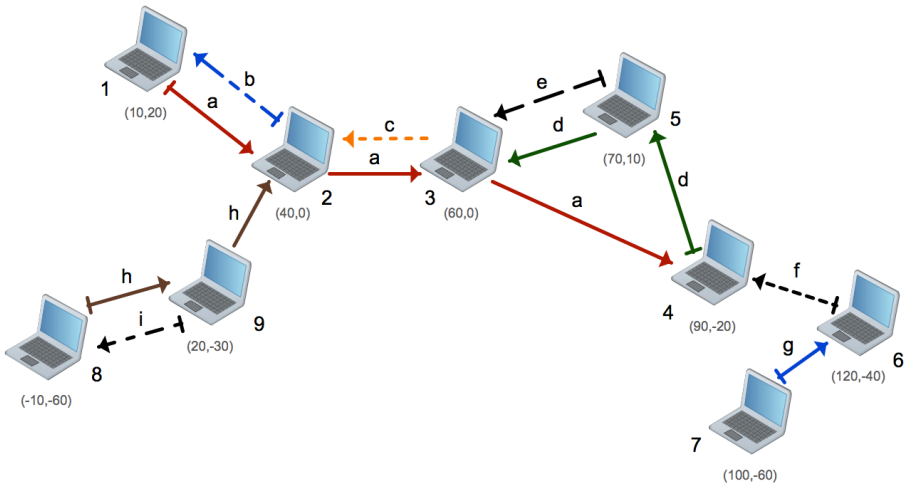
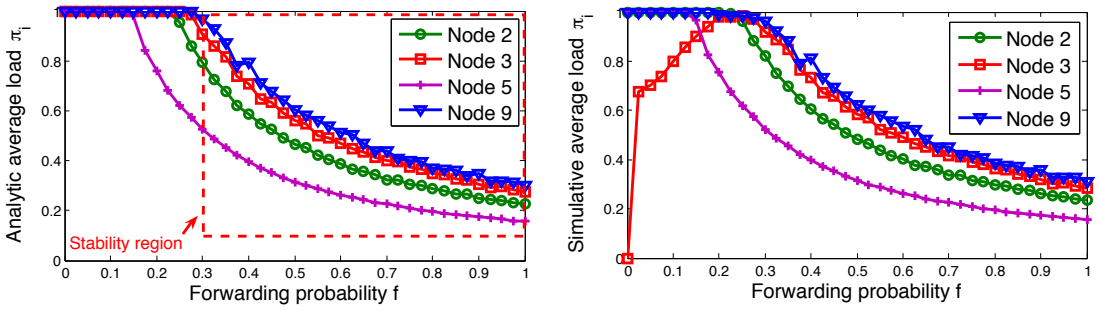


Figure 6.4: *The multi-hop wireless ad hoc network used for simulation and numerical examples.*

We turn in this section to study a typical example of multi-hop ad hoc networks. We consider an asymmetric network formed by 9 nodes, identified using an integer from 1 to 9 as shown in figure 6.4. We establish 9 connections (or paths) labeled by a letter from a to i. Each node is located by its plane Cartesian coordinates expressed in meters. Except contraindication, the main parameters are fixed to the following values : $CW_{min} = 32$, $CW_{max} = 1024$, $K_{i,s,d} \equiv K = 5$, $f_i \equiv f = 0.9$ (to insure operating in the

stability region of all forwarding queues), $T_{i,s,d} \equiv T = 0.1W$ $(\forall i, s, d)$, $CS_{th} = 0$ dBm, $RX_{th} = 0$ dBm, $SIR_{th} = 10$ dB (target SIR), $\rho = 2$ Mbps (bit rate), $\tau = 20\mu s$ (physical slot duration), $DISF = 3\tau$ and $SIFS = \tau$. For sake of illustration, we assume that the signal attenuation is only due to the path-loss phenomenon, i.e., a tagged receiver experience a signal power of $c \cdot T_{i,s,d}^{-\alpha}$, where $\alpha = 2$ (path-loss exponent factor) and $c = 6$ dBi (antenna isotropic gain).



(a) (b)

Figure 6.5: *Average forwarding queues load from model versus simulation as function of forwarding probability.*

Model validation : We now present extensive numerical and simulation results to show the accuracy of our model and then study the impact of joint PHY/MAC/NETWORK parameters. For that aim, a discrete time simulator that implements the IEEE 802.11 DCF, integrating the weighted fair queueing over the two buffers previously discussed, is used to simulate the former network. Each simulation is realized during 10^6 physical slots, repeated at least 20 times and then averaged to smooth out the fluctuations caused by random numbers generator of the discrete events simulator (back-off counters). We checked the validity of the model by extensively considering different network scenarios (different connections and nodes parameters), several topologies (linear, circular and arbitrary topologies) and different network population size. We depict in figure 6.5(a) and figure 6.5(b) the analytic and the simulative average load π_i of forwarding queues respectively. Numerical plots show that analytic model match well with simulation results, in particular under the stability region which is the main applicability region of our model. With some abuse we refer to the interval of forwarding probability that insure a load strictly less than 1 for all queues, as the stability region of the system. The main difference seen between individual loads is mainly due to the topology asymmetry. Based on figure 6.6 (a) and (b), we note that our analytic finding saying that under the stability condition, the end-to-end throughput does not depend on the choice of the WFQ weight, i.e., the cooperation level or also the forwarding probability f, is confirmed. Therefore, one can judiciously fine-tune the cooperation level value to decrease the delay while the average throughput remains almost constant. This mechanism may play a crucial role in delay sensitive traffic support over multi-hop networks. One can note that the system stability region is strongly impacted by the nodes density. Indeed, in regions with relatively high or high nodes density, it is crucial that relay nodes increase their cooperation level in order to insure stability. Otherwise, the waiting of packet in forwarding queues may grow drastically and the

network reliability becomes a real issue to face with.

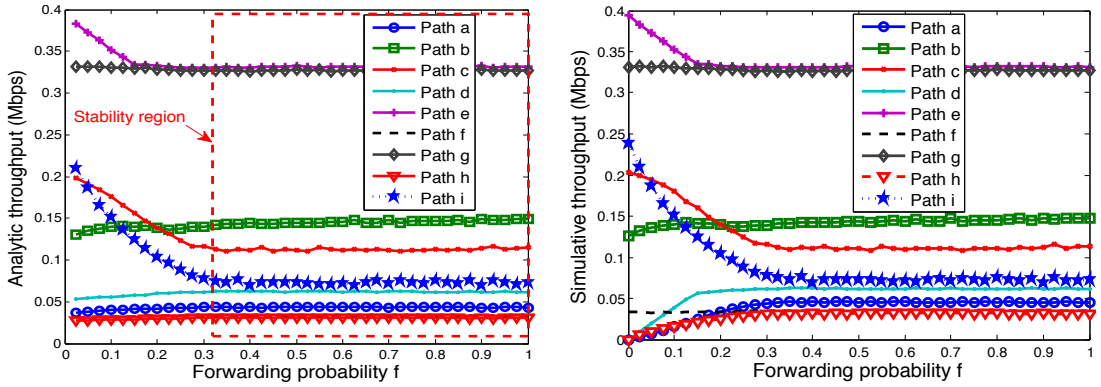


Figure 6.6: *Average end-to-end throughput from model versus simulation as function of forwarding probability.*

Interested reader is referred to [46] for more details, extensive simulations and complete performance evaluation in terms of the considered cross-layer architecture. For instance many results on how to set values of CW_{min}, Payload size and other nodes intrinsic parameters were discussed. We first stipulate that an optimal payload size may not exist, see figure 6.7(a). Indeed, we note that some specific payload size is providing good performances in term of average throughput over some paths, but may hurt drastically the throughput on other links and then the reachability becomes a real issue. *Setting the payload size to a fixed value over the whole network is, in general, unfair and is not suitable for multi-hop networks*. However fortunately, locally optimal payload size may exist. This way, it depends strongly on the topology and the local node densities, i.e., the number of neighbors per m^2, their respective distances with respect to a tagged node and how they are distributed in the network. In terms of the minimum contention

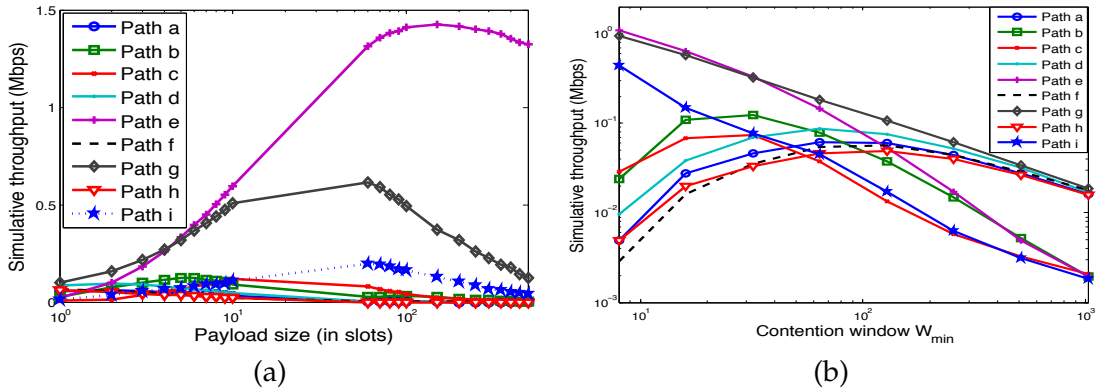


Figure 6.7: *Average end-to-end throughput from model versus simulation as function of payload and contention window CW_{min}.*

window CW_{min}, see figure 6.7(b), the throughput has two different behaviors accord-

ing to the topology of the multi-hop network. Indeed, when the nodes density is low, the throughput is maximized for short backlog duration. Here, nodes take advantage from local nodes density and tend to transmit more *aggressively*, having a relatively low collision probability due to low number of competitors. We also note that contention windows tends to increase as the nodes density becomes high. This latter statement is quite intuitive and due to the competitions that becomes colossal. In terms of queues load (equivalently delay), it is clear that when the contention windows increases it implies the increase of queues load, thus tagged node may suffer from huge delay.

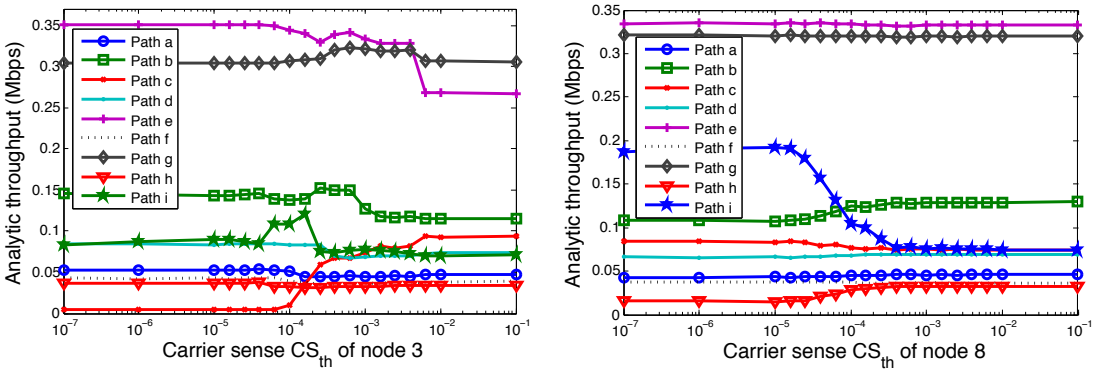


Figure 6.8: *End-to-end throughput from model versus simulation for variable carrier sense threshold (in Watt).*

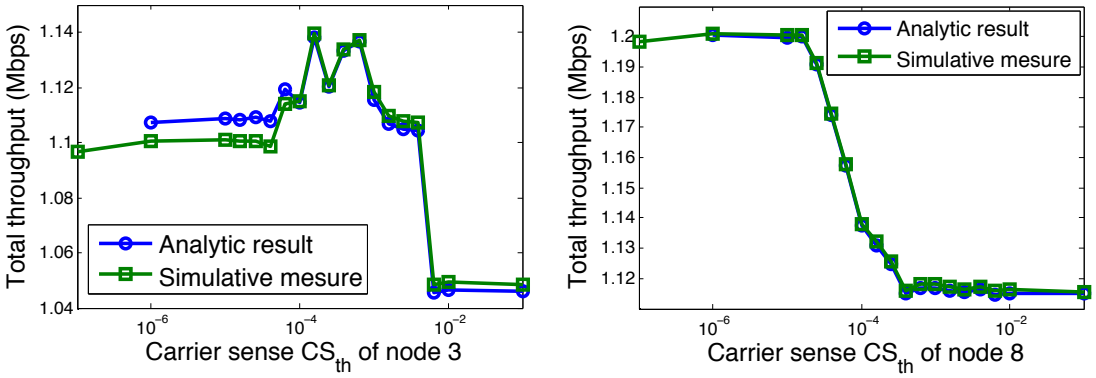


Figure 6.9: *The system capacity from model versus simulation for variable carrier sense threshold (in Watt).*

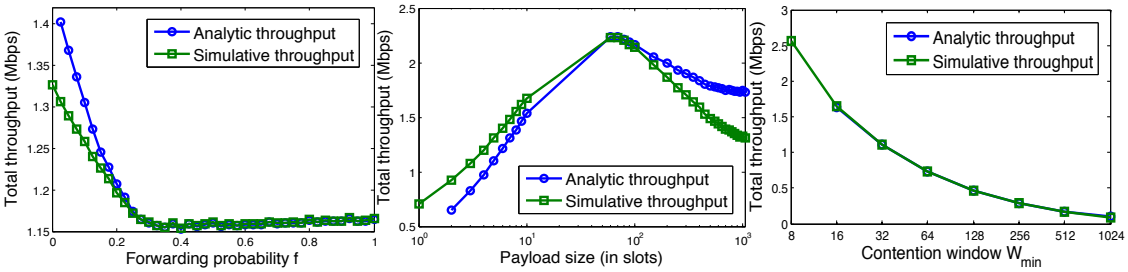
Per-path joint power and carrier sense control : Since the network topology is asymmetric , we can reconsider here the Spanning tree-based algorithm proposed in [90] to compute the optimal transmit power per-path. *Each node sets its transmit power to a level that allows reaching the farthest neighbor*, i.e. the received power is at least equal to the receiver sensitivity. The per-path power control may then improve the spatial reuse over ad hoc networks. In order to analyze the impact of carrier sense threshold on network performances, we vary CS_{th} for some tagged node and fix it to the

default value for others, i.e., $CS_{th} = 0$ dBm. We plot in figure 6.8 the average throughput on all paths when tuning the carrier sense threshold of node 3 which is located in a relatively dense zone. The throughput of all connections continues to decrease (in particular connections crossing node 3 or its immediate neighbors) with CS_{th} except connections originated from node 3. Now we analyze the interplay of node 8 (in a low dense zone) carrier sense on network throughput. We note that the only negatively impacted connection is the connection i originated from node 9 (immediate neighbor of node 8). When carrier sense of node 8 is increasing, it becomes more nose-tolerant which implies high transmission aggressiveness. Which explains the throughput decrease of connection i. Thus connections crossing neighbors of node 9 take benefit from the low attempt rate of node 9 to improve their throughput, for instance connections a, b and h.

In terms of total capacity, see figure 6.9, and depending on the local nodes density, the carrier sense control may increase the network capacity. Indeed, when a node in a dense zone fine-tunes its carrier sense threshold, we note existence of a region where the total capacity is maximized. This region correspond to a CS_{th} interval where tagged node benefits from relatively high throughput and other nodes don't suffer much. Whereas, it seems that tuning carrier sense by nodes in low dense parts of the network may cause a throughput decrease. To sum up, we can say that on one hand, a higher carrier sense threshold encourages more concurrent transmissions but at the cost of more collisions. On the other hand, a lower carrier sense threshold reduces the collision probability but it requires a larger spatial footprint and prevents simultaneous transmissions from occurring, which may result in limiting the system capacity. Analyzing figure 6.10 where the behavior of the total capacity is depicted as a function of nodes intrinsic parameters (f_i, $Payload_i$ and CW_{min}) we note the following : As expected from equation (6.19), the total capacity is insensitive for all cooperation level in the stability region. However, the cooperation is crucial to maintain the network connectivity. In terms of minimum contention window is seems that a as the CW_{min} increases as the total capacity decreases, and an optimal payload length that maximizes the total capacity may exist.

Discussion : In contrast to classical systems where all users communicate with an access point and have, in general, the same channel/environment view, in ad hoc networks, the main difference is the variable topology and the asymmetric environment. A judicious and punctual solution is to auto-configure parameters of the PHY/MAC/NETWORK by the node itself. However unfortunately, this may result in a performance collapse due to users selfishness (similar to prisoners dilemma in game theory). We also suggest to run a MAC/PHY cross-layer control where each node is increasing the transmit power whenever a retransmission is needed. Unfortunately, this power control seems to be unfair since the benefit is strongly depending on the topology. Due to asymmetry, many nodes take benefit from this policy but others may hardly suffer from it. To sum up, under topology asymmetry, the problem is not how to choose parameters such as the network may operate in optimal way; but the problem is how to define a cooperation level and a tradeoff between end-to-end throughput and delay. A challenging but

promising concept is then to enable an autonomous location and environment-aware feature. Then, each node may sense the channel, learn the channel state/network topology, decide the best setup, adapt its parameters and reconfigure them till desired QoS is met. Nodes can then share their respective information for better environment awareness and less signaling traffic.



Impact of forwarding probability. Impact of payload size. Impact of contention windows.

Figure 6.10: *The system capacity from model versus simulation as functions of different parameters.*

Remark 6.4.1. *It was shown in earlier works, see e.g., [83], that a maximum throughput is achieved in the shortest path. A high amount of traffic in the topology of figure 6.4 is issued from one hop paths, which explains the continuous decrease of the capacity with cooperation level f until saturation when forwarding queues become stable.*

6.5 Concluding discussion

Multi-hop ad hoc networks are based on the fact that nodes would cooperate with each other to accomplish safe communications. In addition, a stack of protocols would interact with each other to accomplish a successful packet transfer. In this context, we have developed the model by using the IEEE 802.11 DCF MAC protocol. We are interested to study the effect of forwarding on nodes using the IEEE 802.11 (and its extensions) based ad hoc networks. This have led us to consider different traffic flows in the network. We have discovered that the modeling of the IEEE 802.11 in this context is not yet mature in the literature. Furthermore, to the best of our knowledge, there is no study done that considers jointly the PHY/MAC/NETWORK cross layer interaction in a non-uniform traffic and an asymmetric network topology. Therefore, we have extended Yang's model [166] using the perspective of an individual sender. The attempt rate and collision probability are now depending on the traffic intensity, on topology and on routing decision. The fixed point *system I* is now related to the traffic intensity *system II*. For that, with the means the cycle of transmissions approach and by writing the rate balance equations, we have expressed this latter system. In fact, the two systems are complicated to solve and computational expensive. For that, special cases can be useful. The two main results concerning stability-throughput in the network are as follows: (1) stability of forwarding queues and attempt rate are independent of the forwarding probabilities, this latter can be judiciously used to design an admission control to limit the average end-to-end delay while the throughput remains constant, and

(2) the end-to-end throughput of a connection does not depend on the load of the intermediate forwarding queues between a source and a destination. Numerical results and simulations have shown the interest of our model by presenting the end-to-end throughput and stability function of several parameters. The problem addressed here opens many interesting directions to study in future, as power control problem and the effect of different parameters of the IEEE 802.11e and the physical layer on routing and performances of the network. In particular, the issue of cooperation in this kind of networks which can be efficiently studied in a game theoretical perspective.

To cooperate or not to cooperate ? In terms of energy criterion, the main battery consumption of a node is due to the channel sensing and transmission of data packets. The issue of cooperation in this kind of networks can be efficiently studied in a game theoretical perspective. Indeed, starting from users rationality and the punishment mechanism, it is plausible that each node will think the following : If a tagged user i cooperates and forwards others packets, he/she will be able to benefit from the network and then may sends its packets! else he/she will be punished and his packets will be discarded! This way, forwarding packets of i can be seen as its payoff after forwarding packets of other nodes. This scheme looks like a non-cooperative game where each player has to decide either to cooperate or not to cooperate. The key assumption that makes this statement relevant and justified is that own queues are assumed to be *saturated* and each user aims to transmit its packets as soon as possible. However setting the cooperation level to the same value for all nodes, may be unfair, in particular when relaxing the saturation assumption. Indeed, some sources may have more packets to transmit than others and then this may penalize nodes in towards destination by consuming rapidly their battery.

We notice that due to topology and nodes density asymmetry impacts strongly the success probability of a tagged node. This implies a non negligible fairness issue for IEEE 802.11e in ad hoc mode. This is why we propose, for nodes that suffer from high collision probability, to encode their packets before transmitting them. Therefore, we don't change the standard but we add and encoder/decoder at the MAC layer of each node. The encoding/decoding scheme we propose is the well known and recently invented Fountain codes. Using incremental combination, we expect that the retries limit per-packets is reduced and the Jain's fairness index will be improved.

Chapter 7

Fountain Code-based Fair IEEE 802.11-Operated Ad hoc Networks

Contents

7.1 Introduction

Error correction coding is incontestably one of the most important elements of a modern communication system. Here, a redundant information (parity bits) is added to the original message to make it more resilient against noise/errors induced during transmission. Nevertheless, the field of error correcting coding is not new, since it dates back to the famous paper published by Shannon [139]. Shannon predicted that reliable communication is achievable, if redundancy is added to the message across a memoryless channel, as long as the communication rate does not exceed the capacity of the channel. Shannon however did not propose any coding scheme.

In TCP/IP network, the file delivery works as if follows : The sender emits data packets one after the next, and when the receiver misses a packet, it informs the sender so the packet can be retransmitted. This back-and-forth dialogue between sender and receiver forms the basis of many modern applications, including email, web, and FTP. In Mobile Broadcast File Networks, things are very different. With a potentially infinite number of receivers, each receiver must take whatever packets it can get. Since

the broadcast channel is one way (half-duplex), there is no opportunity for a tagged receiver to request retransmission of a lost packet. Moreover, if receivers were actually able to request packet retransmissions, doing so would be utterly impractical [152]. Even with minimal packet loss, the aggregate requests from a large receiver population would completely swamp the network. This way, file delivery over a broadcast network inevitably results in lost packets on the individual receivers. The losses may appear in bursts of missed packets and/or as individual dropped packets. Erasures in a received file -even just a single byte- can render a file completely useless. However fortunately, there is a technique for delivering files reliably over a network despite the inevitable erasures. That technique is called Fountain Codes.

Fountain Codes are a class of codes designed for reliable transmission over a channel with unknown quality. The concept behind them is elegant and simple, see [97] and [152] for a very nice overview. An encoder generates a stream (with some abuse, a "fountain") of packets from an original file. Any receiver trying to download an encoded file need only receive enough packets from the fountain -any packets, in any order, continuous or with holes- and the receiver is then able to successfully reconstitute the file from the received packets. The question might be now : "How many packets must the receiver collect before it can reproduce the original file ?" The answer is : "Surprisingly few". The total number of received packets altogether is only slightly larger than the total number in the original file. In other words, a fountain encoder produces, for a given vector of M input bits, a potentially limitless stream of output bits. Each output bit is produced independently and randomly from the M input bits. The receiver collects output bits of the encoder from the channel, and with each bit, it records the reliability of the bit. The receiver collects bits until the sum of information of individual bits is $M(1+\epsilon)$, where ϵ is an appropriate constant, called the overhead, that allows the receiver to recover the correct input bits with high probability. The magic of Fountain Coding is fully described in the papers and patents describing the technique; See e.g., [31, 95, 97, 140]. Digital Fountain's technology has been deployed in the automotive, military, and digital radio industries. Their announced customers include XM, Sirius, Honda, Sumitomo Electric, NTT, Nokia, and Northrop Grumman. The technique has been published in the IEEE Transactions on Information Theory, and was honored with the prestigious IEEE 2002 Best Paper award. Nowadays, Raptor codes have been standardized in :

- 3GPP MBMS: 3rd Generation Partnership Project for use in mobile cellular wireless broadcast and multicast,
- DVB-H IPDC standards for IP datacast.

Authors in [33] studied the performance of broadcasting algorithms for underwater acoustic sensor networks (UWASNs). They considered a single source and a number of nodes randomly placed within a given geographical area. Later, a hybrid ARQ scheme is used. Fountain Codes are exploited to enhance the efficiency of the data dissemination process in the face of poor and possibly unknown channel conditions. The main contribution of this paper is a mathematical model to characterize the performance of Fountain codes as applied to broadcasting in underwater networks. Relevant tradeoffs

are highlighted and quantified; in particular the implications of transmission power on covered distance, rate and delay are discussed. Kumar et al. [86] presented a Fountain codes based Transport protocol as an alternate paradigm to that of the ubiquitous TCP. The new scheme abolishes the need for a reverse feedback mechanism usually essential to provide reliability in packet data transmission. Indeed, absence of a reverse feedback mechanism can substantially improve the performance of networks with half-duplex wireless channels (such as 802.11 WLANs), where collisions between forward and reverse MAC frame transmissions contribute significantly towards performance degradation. Based on a Markovian stochastic model, authors analyzed the performance of a simple Fountain codes based Transport protocol in a single cell IEEE 802.11 WLAN. Using renewal theory, Kumar et al. provided an explicit expression for the average downlink throughput. ns2 simulations are used to validate their model and the analytically obtained throughput metric.

We propose here a new scheme using Fountain encoder/decoder at MAC layer in multihop ad hoc context. Our work is original since, to the best of our knowledge, we are the first to use Fountain codes to design a HARQ-like mechanism at IEEE 802.11e EDCF MAC layer. We conclude by observing that Fountain coding and incremental redundancy are very effective in terms of error correction capability and needed transmission retries. However unfortunately, using Fountain coding in multihop ad hoc context does not always provide better performances. Indeed, they are only suitable for large networks where accumulative interference may hurt transmissions. Here, the joint Fountain coding and incremental redundancy introduce a better Fairness and outperform performances of IEEE 802.11e EDCF in terms of end-to-end throughput.

7.2 Fountain code-based IEEE 802.11e DCF/EDCF

7.2.1 Fountain code-based MAC layer

We consider the modeling problem of the IEEE 802.11 using the perspective of a sender which consists on the channel activity sensed by a sender, or on the state (success or collision) of its transmitted packet. This will facilitate the problem in the ad hoc environment where nodes have an asymmetric vision of the channel. We start, again, by defining the notion of virtual time slot and channel activity, then we write the expression of the attempt probability in the asymmetric network. Consider that time is slotted with a physical slot duration τ. Nodes transmit in the beginning of each slot and the transmission duration depends on the kind of the transmitted packet. A data packet has a fixed length and takes M slots to be transmitted without encoding. Whereas a data encoded packet takes L (integer) slots to be transmitted (it includes the header transmission time) such as $L = M(1+\epsilon)$, where $\epsilon \cdot M$ is a small number excess packets. While an acknowledgment packet spends ACK slots. The encoder of a Fountain code is a metaphorical fountain that produces an endless supply of water drops (encoded packets) [97], here we consider that after coding phase, the users schedules transmis-

sion of $\bar{L}$ encoded Fountain Codes Packets, say FCPs. Then, at the input of the receiver decoder, it suffices to correctly decode M FCPs to fully recover the original M-length packet.

Example : Imagine that you wish to transmit a 10 MB file through a loss-prone channel [152]. Assume that you can expect somewhat less than a 20% packet loss. Now process the file using Fountain encoding and decoding. The process works as follows:

- Start with a 10 MB file.
- Apply the encoding process, expanding the file to 15 MB.
- Send the encoded file through the channel.
- The receiver experiences a "worst case" loss of about 20%. Specifically, it receives 12 MB of the 15 MB file. The rest is corrupted or lost entirely.
- The receiver applies a decoding process to successfully recover the original 10 MB file.

The term "Fountain Coding" is quite appropriate here. If you were to top off an empty glass from a fountain, you would need to capture only enough drops to fill the glass. It doesn't matter if you spill some water, or pull the cup out of the fountain when it is halfway full, wait a few minutes, and then put the cup back under the fountain. Similarly, to receive a Fountain Coded file, one only need to receive only enough packets to "fill up" a quota. One may stop listening for a period of time, and then start capturing packets again. All that matters is that you receive enough packets.

For illustrative purpose and such as in the previous chapter, we reconsider the following interfering groups : $\{6\}$ and $\{5,7,8\}$. Node 6 is a one-hop neighbor of node $j_{i,s,d}$ which explain its strong impact and obvious collision when they both are transmitting. Whereas a failure may only occur when node of group $\{5,7,8\}$ are all transmitting simultaneously, see figure 7.1. However, combining FCPs that are correctly received with each other or with received FCPs from previous attempts will improve the capture as well as the decoding outcome.

7.2.2 Failure probability and virtual slot expressions.

In IEEE 802.11 DCF standard a transmission is successful when all blocs composing a packet are successfully received. The collision probability of a packet is then defined when either the data or the acknowledgment experiences a collision. Under our Fountain coding-based scheme, a transmission is successful when a M-length uncorrupted packet is totally recovered at the output of the decoder. This way, the FC decoder may recover, with probability $1-\delta(l)$, the original packet when at least l ($l \geq L$) FCPs are successfully received instead of $\bar{L}$ sent FCPs. This is why we need to define what we

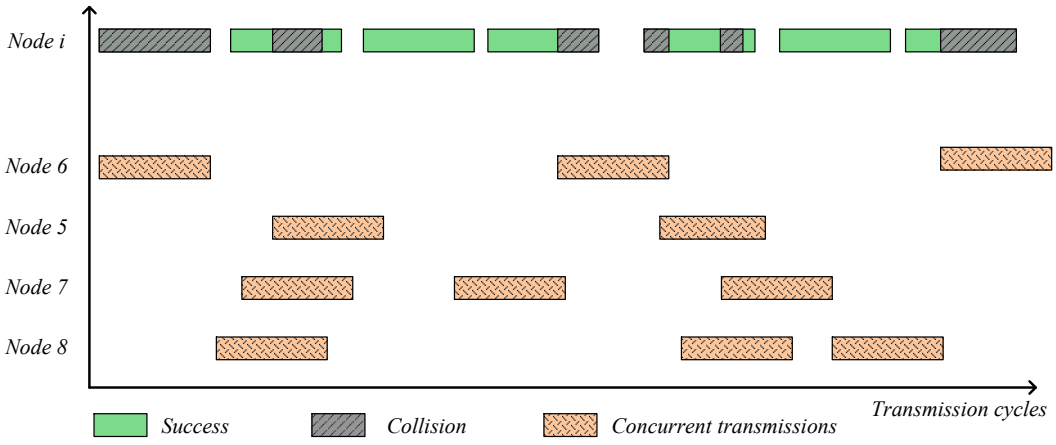


Figure 7.1: *Effect of accumulative interferences on transmission of node i to node $j_{i,s,d}$, see figure 6.2.*

call the transmission *failure probability*, it is given by

$$\gamma_{i,s,d} = 1 - (1 - \gamma^{D}_{i,s,d})(1 - \gamma^{A}_{j_{i,s,d},s,d}), \tag{7.1}$$

where $\gamma^{D}_{i,s,d}$ and $\gamma^{A}_{j_{i,s,d},s,d}$ are, respectively, the failure probability of a data packet and its acknowledgement collision probability. After the beginning of data transmission, nodes in H_i will defer their transmission to $EIFS$ (Extended Inter-Frame Space) duration, which would insure the good reception of the acknowledgment. In practice, acknowledgement are small packets and less vulnerable to loss, for that it is plausible to consider $\gamma^{A}_{j_{i,s,d},s,d} \simeq 0$. Then, we can write $\gamma_{i,s,d} = \gamma^{D}_{i,s,d}$. A tagged FCP experiences a collision when nodes group $\mathcal{Z} \in \mathcal{I}_{i,s,d}$ are transmitting all together during the corresponding time slot. Let the random variable $X_{\mathcal{Z}}$ denote the number of lost FCPs due to collision with nodes group $\mathcal{Z}$. Since the intermediate node i transmits a flow of $\bar{L}$ FCPs, then $X_{\mathcal{Z}} \in \{0, 1, \cdots, \bar{L}\}$, thus

$$P(X_{\mathcal{Z}} = t) = \binom{\bar{L}}{t} \hat{P}^{t}_{\mathcal{Z}} \left(1 - \hat{P}_{\mathcal{Z}}\right)^{\bar{L}-t}, \tag{7.2}$$

where $\hat{P}_{\mathcal{Z}}$ is the transmission probability of the nodes group $\mathcal{Z}$, it is given by $\hat{P}_{\mathcal{Z}} = \prod_{k \in \mathcal{Z}} \frac{P_k}{\Delta_k}$. We can then calculate the average number of slots that takes a transmission initiated by group $\mathcal{Z}$, it is given by $\sum_{t=1}^{\bar{L}} t \cdot P(X_{\mathcal{Z}} = t)$. Note that the distribution of the duration of a group transmission, seen by a given receiver, is arbitrary and is strongly depending on the local topology, nodes density and the individual transmit power.

Example : We consider in figure 7.2 a transmission attempt from node i to its next hop j. The virtual node $\mathcal{Z} = \{k, l\}$ may interfere with this latter communication when both node k and node l are transmitting simultaneously. In subfigure 7.2 (a), nodes

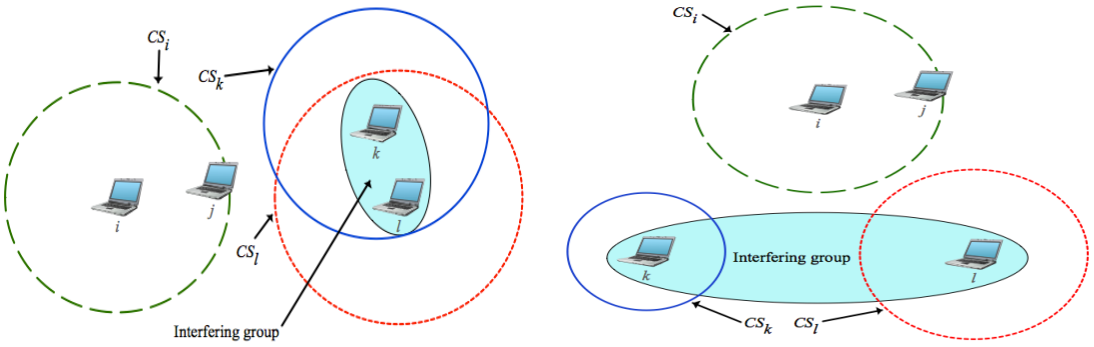


Figure 7.2: *Effect of carrier sensing among nodes of the interfering group {k, l} when intermediate node i transmits to node j. In scheme (a), nodes k and l hear each other, whereas they can't hear each other in scheme (b).*

k and l may detect transmissions of each other and therefore are able to avoid collisions induced by themselves. This way, node k will never attempt a transmission when he/she detects that the channel is being used by node l and vice-versa. Then $P(X_{\mathcal{Z}} = t) = 0, \forall t = 0, 1, \cdots, \bar{L} - 1$. However unfortunately, when the backoff counters of both nodes k and l expire at the same slot, a simultaneous transmission occurs and the remaining FCPs of node i are lost. Thus $P(X_{\mathcal{Z}} = \bar{L}) = 1)$.
In subfigure 7.2 (b), nodes k and l are unable to detect ongoing transmissions of each other. Then a simultaneous transmission of both of them may occur during $0, 1, \cdots$ or $\bar{L}$ slots which is different from previous scheme.

Define now $\eta_{i,s,d}(l)$ as the success probability experienced by a transmission originated from node i when the receiver $j_{i,s,d}$, in the path $R_{s,d}$, captures l ($l \geq L$) FCPs and decodes them to recover successfully the original M-length packet. When receiving l FCPs, the output of the Fountain decoder yields a corrupted packet with probability $\delta(l) \leq 2^{M-l}$. We consider here the worst case and stipulate that a success decoding occurs with probability $1 - \delta(l)$. Thus

$$\eta_{i,s,d}(l) = \prod_{k \in H_i \cap \mathcal{I}_{j_{i,s,d}}} (1 - P_k) \prod_{\mathcal{Z} \in \mathcal{I}_{j_{i,s,d}} \setminus H_i} P\left(X_{\mathcal{Z}} = \bar{L} - l\right), \tag{7.3}$$

Received FCPs may not always be decoded correctly by the FC decoder, thus

$$\gamma_{i,s,d} = 1 - \sum_{l=L}^{\bar{L}} (1 - \delta(l)) \cdot \eta_{i,s,d}(l). \tag{7.4}$$

Nodes in area $H_i \cap \mathcal{I}_{j_{i,s,d}}$ must be silent at the beginning of node i transmission. While nodes in $\mathcal{I}_{j_{i,s,d}} \setminus H_i$ are hidden from i and needs to be silent during at least the duration needed to receive correctly L-length FCPs. $\dfrac{l}{\Delta_{i,s,d}}$ is called the normalized vulnerable

time corresponding to transmission of l FCPs.

Now, the virtual slot, seen by an attempted packet in the path (s,d), as defined previously can be written as following:

$$\Delta_{i,s,d} = P^{succ}_{i,s,d}.T_{succ} + P^{col}_{i,s,d}.T_{col} + P^{idle}_{i}.T_{idle} + P^{busy}_{i}.T_{busy}, \tag{7.5}$$

where $P^{succ}_{i,s,d} = P_{i,s,d}(1-\gamma_{i,s,d})$ and $P^{col}_{i,s,d} = P_{i,s,d}\gamma_{i,s,d}$. A successful transmission takes $T_{succ} = \bar{L} + ACK + SIFS + DIFS$ slots to be transmitted. While a collision takes $T_{col} = \bar{L} + ACK + DIFS$ slots. The duration of an idle period is $T_{idle} = \tau$ and the probability that this idle time occurs is

$$P^{idle}_{i} = \prod_{\mathcal{C}\in CS_i\cup\{i\}} (1-P_{\mathcal{C}}). \tag{7.6}$$

Also, the duration of a busy period is approximated by $T_{busy} = \bar{L} + DIFS$ and the corresponding probability is

$$P^{busy}_{i} = (1-P_i)\left(1 - \prod_{\mathcal{C}\in CS_i} (1-P_{\mathcal{C}})\right). \tag{7.7}$$

Hence, we have

$$\begin{aligned}\Delta_{i,s,d} \quad = \quad & P_{i,s,d}\left(1-\gamma_{i,s,d}\right).T_{succ} + P_{i,s,d}\gamma_{i,s,d}.T_{col} + \\ & \prod_{\mathcal{C}\in CS_i\cup\{i\}} (1-P_{\mathcal{C}}).\tau + (1-P_i)\left(1 - \prod_{\mathcal{C}\in CS_i} (1-P_{\mathcal{C}})\right).\end{aligned} \tag{7.8}$$

Finally, let us denote the equations (6.5), (6.6), (7.1) and (7.8) by *system I*. The Network layer model is exactly the same as the one derived in Section 6.3. We follow the method we proposed in previous chapter to write the rate balance equations at each node, from which $\pi_{i,s,d}$ can be derived function of P_j and $\gamma_{j,s,d}$, for all j. Using algorithm 5, we solve the coupled systems and therefore compute the average throughput of each path and the load of each forwarding queue.

7.3 Fair Bandwidth Allocation

An important aspect related to the efficiency and resource sharing among the network is regarding network fairness, i.e., the fair share of channel access among competing nodes/paths, leading to overall throughput values much higher across all nodes. The IEEE 802.11 DCF/EDCF MAC, especially in ad hoc mode, is well-known for presenting serious fairness problems. Here, the fairness problem is directly related to the binary exponential backoff algorithm, asymmetric topology, node density and, consequently, to the selection of the minimum contention window size as well. To compute the level

of fairness with respect to e2e throughput, we use Jain's fairness index [52, 79, 80], which is defined as

$$\text{Fairness Index} = \frac{1}{n} \cdot \frac{\left(\sum_{s,d} thp_{s,d}\right)^2}{\sum_{s,d} thp_{s,d}^2} \tag{7.9}$$

where $thp_{s,d}$ is the throughput of the tagged active connection (s,d) in the network, and n is the number of active connections/paths[1]. The e2e throughput is given by equation (6.18) replacing collision probability the the failure probability given by (7.4). Regarding the fairness index, the smaller the value is, the more unfair the network becomes. A fairness index closer to 1 indicates almost equal bandwidth shares among the n active flows.

7.4 Simulation results

In contrast to their use in broadcast systems, the use of Fountain coding in multihop networks has not always a beneficial effect on network performances. We performed several simulations with various topologies and nodes densities. Indeed, we found that although our scheme provides a better channel sharing, it may provide less throughput, especially in small and sparse networks. However fortunately, Fountain coding is useful in large and dense networks, i.e., where the cumulative interferences effect becomes perturbing. Figure 7.3 shows the network used for performance evaluation of our Fountain code-based MAC layer. It is a collection of 30 nodes labeled by an integer from 1 to 30. The location of each node is defined by its Cartesian coordinates expressed in meters. Except contraindication, the main parameters are fixed to the following values : $CW_{min} = 32$, $CW_{max} = 1024$, $K_{i,s,d} \equiv K = 6$, $f_i \equiv f = 0.8$ (to insure operating in the stability region of all forwarding queues), $T_{i,s,d} \equiv T = 0.1W$ $(\forall i, s, d)$, $RX_{th} = 0$ dBm, $SIR_{th} = 10$ dB (target SIR), $\rho = 2$ Mbps (bit rate), $\tau = 20\mu s$ (physical slot duration), $DISF = 3\tau$ and $SIFS = \tau$. We consider that each packet payload (without coding redundancy) requires $M = 200$ slots to be transmitted. For sake of illustration, we assume that the signal attenuation is only due to the path-loss phenomenon, i.e., a tagged receiver experience a signal power of $c \cdot T_{i,s,d}^{-\alpha}$, where $\alpha = 2$ (path-loss exponent factor) and $c = 6$ dBi (antenna isotropic gain). Each result is obtained by averaging 10 independent simulation runs. Flows crossing user nodes labeled by integers 3, 4, 7, 11, 16, 23 and 30 suffer from very bad channel conditions. To improve their average throughput, these users encode their packets before transmitting them using a Fountain encoder with redundancy $M\epsilon$.

Fairness index and aggregate throughput are shown in table 7.1. The first row corresponds to the IEEE 802.11e EDCF, i.e., without Fountain encoding. We first note that our scheme has better fairness properties compared to IEEE 802.11e EDCF. Moreover,

[1] The tagged connection (s,d) is active iff $p_{s,d} \neq 0$

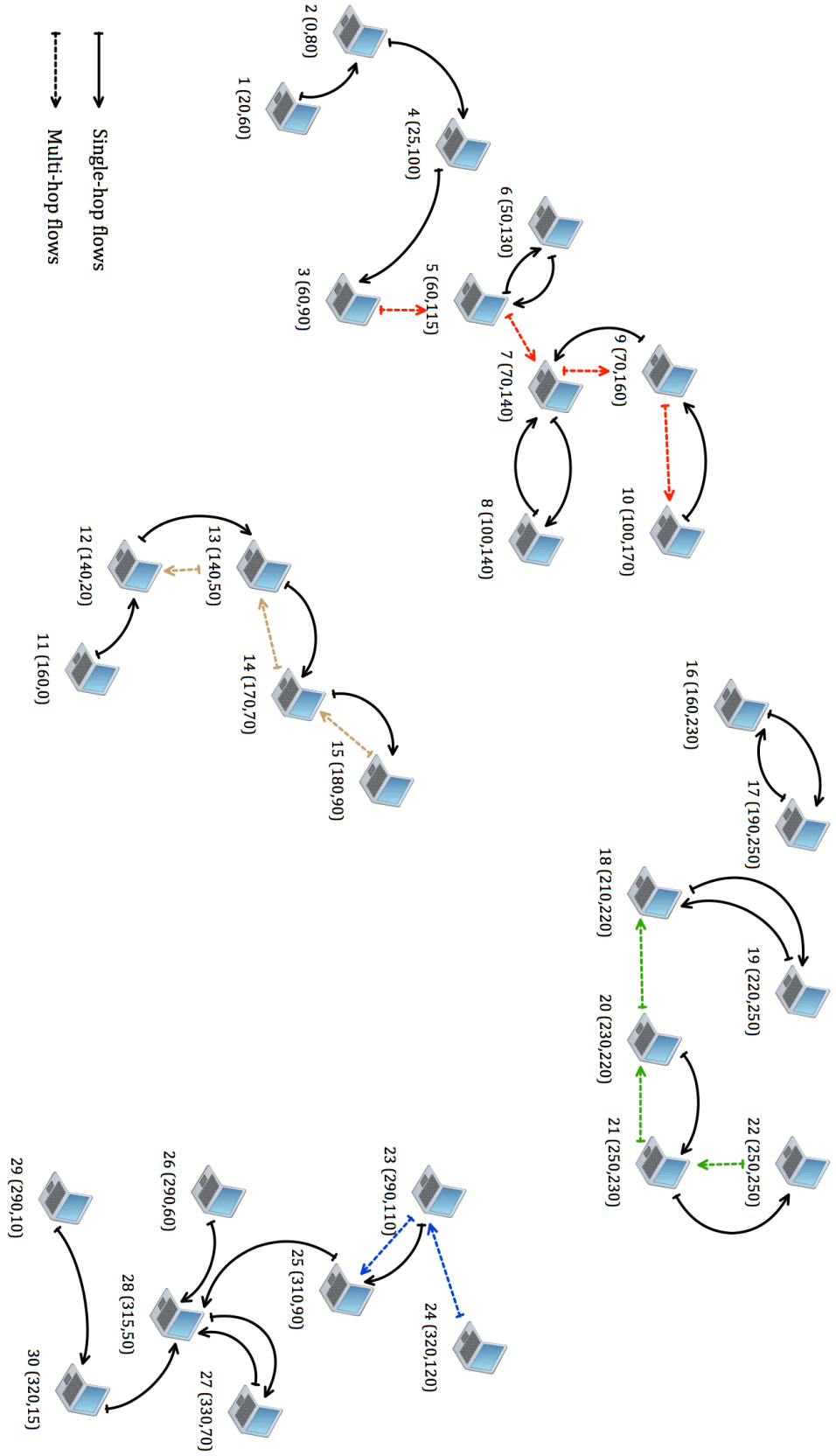


Figure 7.3: *Multihop ad hoc network used for simulations. Flows crossing users labeled by 3, 4, 7, 11, 16, 23 and 30 suffer from very low throughput. To improve their throughput, these users encode their packets before transmitting them.*

the fairness index increases as the redundancy size increases, but tends to decrease as the redundancy becomes larger than 200. Indeed, when the redundancy becomes large, the tagged user using Fountain encoding monopolizes the channel for long time. This results in a channel under-utilization and an average throughput experienced by concurrent links decreases. Henceforth, this implies a lower fairness index. The same explanation holds for aggregate throughput. Table 7.1 shows existence of an optimal redundancy that maximizes the aggregate throughput. We also observe that the maximizer of the aggregate throughput (around $\epsilon = 0.5$) and the maximizer of the fairness (around $\epsilon = 1$) do not coincide. A throughput/fairness tradeoff can be defined for better performance and fair channel sharing.

FC redundancy ϵM	Fairness Index	Aggregate throughput (Mbits/s)
IEEE 802.11e EDCF	0.3271	1.44
20	0.3884	1.59
40	0.3748	1.65
60	0.3823	1.76
80	0.3738	1.84
100	0.3911	1.96
120	0.3966	1.62
140	0.3940	1.62
160	0.3981	1.56
180	0.4107	1.52
200	0.4320	1.51
300	0.4230	1.52
400	0.4035	1.51
1000	0.3419	1.36

Table 7.1: *Fairness result of Fountain code-based MAC layer VS IEEE 802.11e EDCF for* $W_{min} = 32$ *and* $CS_{th} = 10^{-3}$.

CW_{min}	Fairness Index	Aggregate throughput (Mbits/s)
8	0.3062	1.02
16	0.3417	1.41
32	0.3884	1.59
64	0.4007	1.67
128	0.4210	1.75
256	0.4507	1.88
512	0.5415	1.95
1024	0.6270	2.06

Table 7.2: *Fairness result of Fountain code-based MAC layer versus the contention window* CW_{min}, *for* $\epsilon M = 20$ *and* $CS_{th} = 10^{-3}$.

Contention window and fairness : Next we set the redundancy $\epsilon M = 20$ and the carrier sense threshold $CS_{th} = 10^{-3}$, and experimente with the effect of varying the contention window, upon fairness and aggregate throughput (Table 7.2). We note clearly

that both the aggregate throughput and the fairness increase as the contention windows becomes larger.

Carrier sense threshold and fairness : We vary now the carrier sense threshold and check its impact on aggregate throughput and fairness index of the considered multi-hop ad hoc network. As we expected, the system becomes fairer as the carrier sense threshold becomes small. Yet, setting a small carrier sense threshold results in an accurate channel state detection, thus a high amount of collisions is avoided. However, the user nodes skip numerous transmission chances which results in low throughput. Fine-tuning the carrier sense threshold may then judiciously employed to obtain an efficient throughput/fairness tradeoff.

CS_{th} (Watt)	Fairness Index	Aggregate throughput (Mbits/s)
10^{-2}	0.5478	1.43
10^{-3}	0.5351	1.77
10^{-4}	0.5473	6.01
10^{-5}	0.5578	2.88
10^{-6}	0.8287	1.51
10^{-7}	0.8638	1.52
10^{-8}	0.8574	1.52
10^{-9}	0.8544	1.52
10^{-10}	0.8582	1.52

Table 7.3: *Fairness result of Fountain code-based MAC layer versus the carrier sense threshold CS_{th}, for $CW_{min} = 256$ and $\epsilon M = 100$.*

7.5 Concluding discussion

We modeled the MAC layer of IEEE 802.11e EDCF where packets are coded with a Fountain coding before transmission. The idea here is to use incremental redundancy to improve the success probability for links suffering from bad channel conditions. A detailed performance analysis study is then carried out to provide insights into the choice of various system parameters that can lead to optimal throughput performance. We concluded that the use of Fountain codes in ad hoc networking is not always beneficial, since it may decrease the aggregate throughput and is not suitable for sparse networks. In dense and large networks, our scheme outperforms IEEE 802.11e EDCF in terms of fairness and aggregate throughput. We also showed the existence of a throughput/fairness tradeoff. This solution is quite promising since it requires no change of the standard and only introduces a low amount of redundancy.

Conclusion

Summary and general discussion

For wireless next generation networks, simple (i.e., low overhead), scalable, distributed, load-balancing and link quality-aware routing protocols would be required for efficient multihop communications. Designing efficient routing protocols for multi-channel and multi-radio networks is a major research challenge. In particular, an integrated design of medium access control and channel allocation (or scheduling) may lead to an efficient solution. This is why it is important to design distributed algorithms which can be used by the mobiles to compute the equilibrium strategy and simultaneously achieve the optimal operation points. The obvious desirable features of such an algorithm are that it should be decentralized, distributed scalability and should be able to adapt to changes in network.

In the first part of this thesis , we addressed the under-utilization problem of wireless collision channels, e.g., slotted aloha and CSMA. On one hand, we considered the team problem where all users are assumed to maximize a common objective function. On the other hand, we analyzed the game problem where each user seeks to maximize its own objective. Based on a Markovian model and a stochastic game formulation, many performance metrics were derived. We then proposed an improved version with capture effect capability. Using several transmit power allocation policies, four priority-based schemes were discussed. Various different examples and analysis in this work illustrate the better performance of slotted aloha with random power level selection and capture effect. Next, we studied a hierarchical slotted aloha protocol. The main idea is to split the users into two sets, leaders and followers. We showed how hierarchy can improve the channel utilization without any change of the standard protocol. We also provided the protocol design and discussed how its can be implemented in realistic systems. We then introduced the notion of virtual controller that could be the base-station itself. Our hierarchical slotted aloha opens a very promising insights to design next generation MAC protocols. There are however many related issues to solve. For instance signaling issue or how to distribute the leader/follower roles. We provided in chapter 3 a trial to characterize the stability region of collision channels. We relaxed the saturation assumption and showed existence of a continuum of Nash equilibria. Based on energy consumption we discussed the existence of an efficient equilibrium point for all users. Consequently, we proposed two distributed algorithms to learn the desired throughput with a low amount of information (BRA, see algorithm 1) and without any coordination/information (FDTPA, see algorithm 2). Both algorithms converge to the desired equilibrium. Our proposed algorithms are highly scalable in the sense that they are fully distributed and do not require any global information.

Recent field trials and experiments on heterogeneous and wireless mesh networks in several academic research testbeds and commercial installations have shown that the performance is not quite satisfactory. This reflects the need for development of novel architectures and protocol suites to address the issues such as QoS, scalability, heterogeneity, self-reconfiguration, and security for next generation wireless networks. Our

work in chapters 4-7 aims to : 1) understand the behavior of heterogeneous wireless networks and multihop ad hoc networks, 2) improve the channel utilization, 3) propose a cross-layered architecture instead of the standard non communicating OSI layers, 4) provide some new insights to fine-tune network parameters to meet desired tradeoff or best performance and 5) to simply fill the lack of analytical models in the related literature.

Considering a WiMAX cellular network interconnected with a random access wireless ad hoc network, we obtained important insights into various tradeoffs that can be achieved by varying certain network parameters. Some of the important results are that

- As long as the intermediate queues in the network are stable, the end-to-end throughput of a connection does not depend on the load on the intermediate relay nodes.
- As long as the intermediate queues in the network are stable, the end-to-end throughput of a connection does not depend on the cooperation level of the intermediate relay nodes.
- Routing policy can be crucial in determining the stability properties of the network nodes. In a heterogeneous network context, we discussed and proposed a route selection scheme that maximizes the end-to-end throughput.
- The results of parts 2 can be extended in a straightforward manner to systems of weighted fair queues with coupled servers.
- Using the virtual slot definition, our results can be straightforwardly extended to CSMA and its variants.

We also provided various analyses and approximate model to show the validity of our Probability Generating Function approach to derive the distribution of delay in multihop ad hoc networks. Indeed, we have given an approximate and accurate analysis for end-to-end loss probability that we used to build a cross-layered packet control. Our analytical result is then straightforward applied to support conversational and streaming services. Moreover, we have driven extensive numerical and simulation results. We also have discussed how users intrinsic parameters should be set for better performance and delay/throughput/stability tradeoff.

The last part of this thesis highlights the importance of efficient use of spatial resource in multihop environment unlike in conventional single-hop wireless networks and proposes a APPLICATION/NETWORK/MAC/PHY cross-layer architecture instead of standard OSI model. Furthermore, the effect of interferences among all nodes in the network is shown to be crucial since it is responsible of queues dependencies and packets loss. Our main results in chapter 6 consisted on characterizing the forwarding queues stability, end-to-end throughput and delay metrics for performance evaluation. A direct application of our work is to find new distributed schemes for channel access and routing that work near optimal stability region of the network. Extensive simula-

tions were so important to understand the real impact of a non-ideal MAC/PHY layers on routing and end-to-end performances as well. It is worth mentioning here that many routing protocols in the literature were proposed by considering an ideal MAC/PHY layers. Our simulation results show that such latter approaches have to be revised to account for lower layers impact or at least for the network application constraints. This way, many tradeoffs are possible. In fact, one can guarantee a fixed throughput while the corresponding delay or stability does not exceed some threshold. We also provided a discussion on how to judiciously set nodes intrinsic parameters according the network topology and the local nodes density (number of neighbors, accumulative interferences, channel gain, etc.).

A carrier sense control has also been built based on the assumption that the CS threshold is tunable within the detect sensitivity of the hardware. A higher CS threshold can encourage more concurrent transmissions but at the cost of more collisions. On the other hand, a lower CS threshold reduces the collision probability but it requires a larger spatial footprint and prevents simultaneous transmissions from occurring potentially limiting therefore the network throughput. Obviously there is a tradeoff between high spatial reuse and increased chances of collisions. Each node sets its transmit power to a level that allows to reach the farthest neighbor, i.e. the received power is at least equal to the receiver sensitivity. The interplay of carrier sense threshold is strongly impacted by the nodes local density.

From various test-beds and simulations, wireless ad hoc networks are usually unfair. Consequently, many paths are penalized and suffer from high amount of collision. This implies unstable forwarding queues and a throughput collapse which leads to several retransmission per packet. In standard IEEE 802.11, whenever a collision appears the whole packet is lost. We believed that coding packets at MAC layer, may aid to recover the original packet although it is involved in a collision thanks to incremental combination of previous copies. We decided to use a Fountain coder/decoder pair. We showed that the IEEE 802.11e DCF/EDCF achieves now more fair results in all the cases and better throughput values in most of the considered cases.

Future guidelines

Our ongoing works consist in the improvement of our hierarchical scheme. Indeed, we have seen that Stackelberg formulation leads, without any central entity that manages the transmission scheduling, to better performances, in particular for average and relatively high loads. In order to keep the load of the system reasonable and therefore take advantage from hierarchy, we propose to introduce an admission control such the one studied in [101]. Another solution to avoid the observed aggressiveness behavior of users, at heavy load, is to introduce a transmission cost. Furthermore, we are developing a power diversity solution to balance the leader/follower performance metrics. We are introducing a simple scheme with two power levels : A low power for leaders and

high power level for followers transmissions. We are also working on the proof of the FDTPA, see algorithm 2, convergence as well as characterizing its speed of convergence.

In order to maintain connectivity in an ad hoc or in heterogeneous wireless networks, mobile terminals should not only spend their resources (battery power) to send their own packets, but also for forwarding packets of other mobiles. Since ad hoc networks do not have a centralized base-station that coordinates between them, an important question that has been addressed is to know whether we may indeed expect mobiles to collaborate in such forwarding. If mobiles behave selfishly, they might not be interested in spending their precious transmission power in forwarding of other mobile's traffic. A natural framework to study this problem is noncooperative game theory. As already observed in many papers that consider noncooperative behavior in ad hoc networks, if we restrict to simplistic policies in which each mobile determines a fixed probability of forwarding a packet, then this gives rise to a non efficient equilibrium in which no one forwards packets, see e.g. [55, 111], thus preventing the system to behave as a connected network. The phenomenon of aggressive equilibrium that severely affects performance has also been reported in other noncooperative problems in networking, see e.g. [13, 42, 44, 81], chapter 1 and chapter 2 of this document for a flow control context (in which the aggressive equilibrium corresponds to all users sending at their maximum rate, i.e., transmit w.p. 1). As a possible extension of our work, we cite the hard issue of relaxing the saturation assumption of own queues and consider a realistic finite storage capacity. Now, the packets loss will also be caused by the buffers overhead. Another interesting future direction is to consider dynamic environments, in which the network topology and channels change with time.

Clearly, the researchers community has started to revisit the protocol design of existing wireless networks, especially of IEEE 802.11 networks, ad hoc networks and wireless sensor networks, from the perspective of heterogeneous networks and Wireless Mesh Networks. This research field is still non-mature and presents very interesting issues which need to be solved, especially the technology association and the scalability problems. Given that the PHY, the MAC and the NETWORK layers play such a fundamental role in the performance of any wireless ad hoc network -and because all other layers in the protocol stack rely on the NETWORK/MAC/PHY performance- the focus on the modeling of NETWORK/MAC/PHY-layer interactions deserves a study on its own, and it should be fully exploited in the design and optimization of wireless ad hoc networks context. In fact, contrary to wired networks and the Internet, it has become a common belief that cross-layer designs would largely benefit wireless networks, especially at the PHY and MAC layers [38, 138, 161]. In this respect, the modeling framework introduced in this dissertation constitutes a suitable platform for such studies, as it focuses on the interaction of the most crucial layers and represents their functionalities by taking into account the impact of all nodes in the network. Therefore, by using our analytical modeling framework, we can assess not only the performance of realistic NETWORK/MAC/PHY cross-layer designs, but also the impact that different parameters have on network performance. Chapters 4, 5, 6 and 7 illustrate well the

type of modeling and performance analysis that can be carried out with respect to different choices in the design of intended layers. A possible extension of this model is to study the cooperation paradox from a game theoretical perspective. Furthermore, one can also extend our results on end-to-end delay to explicitly incorporate the routing time and topology discovery time when considering mobile nodes. Our next problem to deal with is to provide a generic model for multihoming in heterogeneous wireless networks and exploit the macro-diversity to improve the capture effect. We are also interested in emerging technologies such as cognitive radio, delay tolerant networks and 4G/5G candidate technologies, more precisely femto-cell and Software Defined Networks also known as flexible radio.

Finally, we believe that researchers in wireless medium access and heterogeneous wireless networks domains or any related area will enjoy, help to improve and take benefits from all presented elements in this thesis.

Appendix A : The number of new arrivals

We are interested here to compute the number of arrivals during a service time and the number of arrivals during the residual time. Let $P\left(a^{i,s,d}=j\right)$, $j \geq 0$ denote the probability that j number of packets arrived in the queue F at node i during a service time of a packet (it may be F or Q depends on source s) on the path from s to d. Then, for $K_{i,s,d}=1$, we have

$$P(a^{i,s,d}=j)=\left(\frac{\tilde{\lambda}_i}{1-\tilde{\lambda}_i}\right)^j \sum_{t=j+1}^{\infty}\binom{t-1}{j}\left(1-\tilde{\lambda}_i\right)^{t-1}\bar{P}_i\left(1-\bar{P}_i\right)^{t-1},$$

and, for $K_{i,s,d}>1$, we have $P(a^{i,s,d}=j)=$

$$\left(\frac{\tilde{\lambda}_i}{1-\tilde{\lambda}_i}\right)^j\left[\sum_{k=1}^{\widetilde{K}-1}\sum_{t=j+k}^{\infty}\binom{t-k}{j}\left(1-\tilde{\lambda}_i\right)^{t-k}\binom{t-1}{k-1}\bar{P}_i^k\left(1-\bar{P}_i\right)^{t-k}P_{i,s,d}\left(1-P_{i,s,d}\right)^{k-1}\right.$$
$$\left.+\sum_{t=j+\widetilde{K}}^{\infty}\binom{t-\widetilde{K}}{j}\left(1-\tilde{\lambda}_i\right)^{t-\widetilde{K}}\binom{t-1}{\widetilde{K}-1}\bar{P}_i^{\widetilde{K}}\left(1-\bar{P}_i\right)^{t-\widetilde{K}}\left(1-P_{i,s,d}\right)^{\widetilde{K}-1}\right],$$

where $\tilde{\lambda}_i=\dfrac{\lambda_i}{1-\bar{P}_i}$ and $\widetilde{K}=K_{i,s,d}$. Let us define the random variable R as the residual service time of a Q packet in service. Then the probability mass function of R is

$$P\left(R=t\right) \;=\; \frac{1}{E[S]}\sum_{n=t}^{\infty}P\left(S=n\right),\;\; t\geq 1,$$

where $E[S]$ represents the mean service time. Let $P(r^{i,i,d}=j)$, $j\geq 0$ denote the probability that j number of packets arrived in the queue F at node i during the residual service time of a Q packet on the path from i to d. Then, for $K_{i,i,d}=1$, we have

$$P\left(r^{i,i,d}=j\right)=\frac{1}{E[S]}\left(\frac{\tilde{\lambda}_i}{1-\tilde{\lambda}_i}\right)^j\sum_{t=j+1}^{\infty}\binom{t-1}{j}\left(1-\tilde{\lambda}_i\right)^{t-1}\left(1-\bar{P}_i\right)^{t-1},$$

where $E[S] = \dfrac{1}{\bar{P}_i}$ and, for $K_{i,i,d} > 1$, we have

$$\begin{aligned} P(r^{i,i,d} = j) &= \sum_{t=j+1}^{\infty} P(r^{i,i,d} = j | R = t) P(R = t) \\ &= \frac{1}{E[S]} \sum_{t=j+1}^{\infty} \binom{t-1}{j} \tilde{\lambda}_i^j \left(1 - \tilde{\lambda}_i\right)^{t-1-j} \left[\sum_{k=1}^{\widetilde{K}-1} \sum_{n=t}^{\infty} \binom{n-1}{k-1} \bar{P}_i^k \left(1 - \bar{P}_i\right)^{n-k} \right. \\ &\quad \left. \times \left(1 - P_{i,i,d}\right)^{k-1} P_{i,i,d} + \sum_{n=t}^{\infty} \binom{n-1}{\widetilde{K}-1} \bar{P}_i^{\widetilde{K}} \left(1 - \bar{P}_i\right)^{n-\widetilde{K}} \left(1 - P_{i,i,d}\right)^{\widetilde{K}-1} \right], \end{aligned}$$

$$\begin{aligned} \text{where} \quad E[S] &= \sum_{k=1}^{\widetilde{K}-1} \sum_{n=k}^{\infty} n \binom{n-1}{k-1} \bar{P}_i^k \left(1 - \bar{P}_i\right)^{n-k} \left(1 - P_{i,i,d}\right)^{k-1} P_{i,i,d} \\ &\quad + \sum_{n=\widetilde{K}}^{\infty} n \binom{n-1}{\widetilde{K}-1} \bar{P}_i^{\widetilde{K}} \left(1 - \bar{P}_i\right)^{n-\widetilde{K}} \left(1 - P_{i,i,d}\right)^{\widetilde{K}-1}. \end{aligned}$$

Further, let $P(a^F = j)$ and $P(a^Q = j)$ denote the probability that j packets arrive in the queue F during a service time of a F and Q packet, respectively. Also let $P(r = j)$ denote the probability that j packets arrive in the queue F during the residual service time of a Q packet.

Then we have

$$\begin{aligned} P\left(a^Q = j\right) &= \sum_d P_{i,d} P\left(a^{i,i,d} = j\right), \\ P\left(a^F = j\right) &= \sum_{s,d} \frac{\pi_{i,s,d}}{\pi_i} P\left(a^{i,s,d} = j\right), \\ P\left(r = j\right) &= \sum_d P_{i,d} P\left(r^{i,i,d} = j\right). \end{aligned}$$

Appendix B : Proof of proposition 5.3.1

First of all, we consider the equality obtained by taking the exponentiation with base z at both sides of (5.4) and for any index i,

$$z^{n_{i+1}} = z^{n_i + rI(n_i) + \sum_{j=1}^{m} a_j^Q + a^F - 1 + I(n_i)}, \quad \forall i. \tag{7.10}$$

Then, we multiply both sides for the joint distribution $P_{h_1,h_2,h_3,h_4,h_5} \equiv P\{n_{i+1} = h_1, n_i = h_2, r = h_3, a^F = h_4, \sum_{j=1}^{m} a_j^Q = h_5\}$ and we sum over h_1, h_2, h_3, h_4, h_5. Note that on the left side the summations on h_2, h_3, h_4, h_5 can be exhausted; Whereas on the right side the

summation on h_1 can be exhausted. Therefore, we can write:

$$\sum_{h_1=0}^{\infty} z^{h_1} P_{h_1} = \sum z^{h_2+h_3 I(h_2)+h_5+h_4-1+I(h_2)} P_{h_2,h_3,h_4,h_5}. \tag{7.11}$$

Let $P(z)$ denote the PGF at regime of the state probability distribution at the imbedded instants. Then the left side of equation (7.11) can be written as:

$$\sum_{h_1=0}^{\infty} z^{h_1} P_{h_1} = P(z). \tag{7.12}$$

Further, let $A^F(z)$ and $A^Q(z)$ denote the PGF of the number of arrivals at regime during the service time of a packet which are from buffers F and Q, respectively. Also let $R(z)$ denote the PGF of the number of arrivals at regime during the residual service time of a packet from buffer Q. Then the right side of equation (7.11) can be written as:

$$\begin{aligned}
&\sum_{h_2=0}^{\infty}\sum_{h_3=0}^{\infty}\sum_{h_4=0}^{\infty}\sum_{h_5=0}^{\infty} z^{h_2+h_3 I(h_2)+h_5+h_4-1+I(h_2)} P_{h_2,h_3,h_4,h_5} \\
&= \sum_{h_4=0}^{\infty}\sum_{h_5=0}^{\infty} z^{h_5+h_4}\left[\sum_{h_3=0}^{\infty} z^{h_3} P_{h_3,h_4,h_5} P(h_2=0) + \sum_{h_2=1}^{\infty} z^{h_2-1} P_{h_2,h_4,h_5}\right] \\
&= \left(P_0 R(z) + \frac{P(z)-P_0}{z}\right)\sum_{h_4=0}^{\infty}\sum_{h_5=0}^{\infty} z^{h_5+h_4} P_{h_4,h_5}.
\end{aligned} \tag{7.13}$$

One may remark here that when buffer F has a packet to be sent, the node chooses to send it from buffer F with a probability f. Since $a_j^Q, j \geq 1$ and a^F are independent random variables, the term $\sum_{h_4=0}^{\infty}\sum_{h_5=0}^{\infty} z^{h_4+h_5} P_{h_4,h_5}$ appeared in (7.13) becomes

$$\sum_{h_4=0}^{\infty}\sum_{h_5=0}^{\infty} z^{h_4+h_5} P_{h_4,h_5} = \sum_{j=0}^{m}\left\{(1-f)\,A^Q(z)\right\}^j f A^F(z). \tag{7.14}$$

If the number of consecutive packets from buffer Q (remember that queue Q_i is saturated, meanwhile, queue F_i may be empty) becomes large enough, i.e., $m \to \infty$ then

$$\sum_{h_4=0}^{\infty}\sum_{h_5=0}^{\infty} z^{h_4+h_5} P_{h_4,h_5} = \frac{f A^F(z)}{1-(1-f)\,A^Q(z)}. \tag{7.15}$$

Then the right side of equation (7.13) can be written as:

$$\left(P_0 R(z) + \frac{P(z)-P_0}{z}\right)\frac{f A^F(z)}{1-(1-f)\,A^Q(z)}. \tag{7.16}$$

After some algebraic manipulations between (7.12) and (7.16), the proof follows.

Appendix C : Computations of P_0, $P'(1)$ and $P''(1)$

The first order derivative of equation (5.6) at $z = 1$ allows to calculate easily the probability that queue F_i is empty, i.e.,

$$P_0 = \frac{f - (1-f)\,A'^{Q}(1) - fA'^{F}(1)}{f\,(1 + R'(1))}. \tag{7.17}$$

The second order derivative of equation (5.6) at $z = 1$ yields

$$P'(1) = \frac{(1-f)\left\{2A'^{Q}(1) + A''^{Q}(1)\right\} + fA''^{F}(1) + P_0 f\left\{2\left(1 + R'(1)\right)A'^{F}(1) + 2R'(1) + R''(1)\right\}}{2\left\{f - (1-f)\,A'^{Q}(1) - fA'^{F}(1)\right\}}, \tag{7.18}$$

and, the third order derivative of equation (5.6) at $z = 1$ yields $P''(1) =$

$$\frac{\left\{\begin{array}{l}(1-f)\left[3A''^{Q}(1) + A'''^{Q}(1)\right] + 3\left[(1-f)\left(2A'^{Q}(1) + A''^{Q}(1)\right) + fA''^{F}(1)\right]P'(1) + \\ fA'''^{F}(1) + P_0 f\left[3\left(2R'(1) + R''(1)\right)A'^{F}(1) + 3\left(1 + R'(1)\right)A''^{F}(1) + 3R''(1) + R'''(1)\right]\end{array}\right\}}{3\left\{f - (1-f)\,A'^{Q}(1) - fA'^{F}(1)\right\}}, \tag{7.19}$$

where $\phi'(1)$, $\phi''(1)$ and $\phi'''(1)$, respectively, represent the first, second and third order derivatives of any probability generating function $\phi(z)$ at point $z = 1$.

Appendix D : Transition matrix of the hierarchical slotted aloha

The probability that the hierarchical aloha system transits from a state (n, n', i, a) to a state $(n + k, n' + k', j, b)$ is given by the following :

$\bullet P_{(n,n',i,a),(n-1,n',j,b)}\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right) =$

$$\left\{\begin{array}{ll}(1-q_a^l)(1-q_a^f)Z_1(0,0,1,0), & a = b = 0, i = j = 0\\ (1-q_a^l)(1-p^f)Z_1(0,0,1,0), & a = b = 1, i = j = 0\\ (1-p^l)(1-q_a^f)Z_1(0,0,1,0), & a = b = 0, i = j = 1\\ (1-p^l)(1-p^f)Z_1(0,0,1,0), & a = b = 1, i = j = 1\end{array}\right.$$

$\bullet P_{(n,n',i,a),(n,n'-1,j,b)}\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right) =$

$$\left\{\begin{array}{ll}(1-q_a^l)(1-q_a^f)Z_1(0,0,0,1), & a = b = 0, i = j = 0\\ (1-q_a^l)(1-p^f)Z_1(0,0,0,1), & a = b = 1, i = j = 0\\ (1-p^l)(1-q_a^f)Z_1(0,0,0,1), & a = b = 0, i = j = 1\\ (1-p^l)(1-p^f)Z_1(0,0,0,1), & a = b = 1, i = j = 1\end{array}\right.$$

$$\bullet P_{(n,n',i,a),(n,n',j,b)}\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right) =$$

$$\begin{cases}
Q_r^l(0,n)Q_r^f(0,n')[Z_2(0)+Z_2(1)] + (1-q_a^l)(1-q_a^f)Z_3(2,0,0), & a=b=0, i=j=0 \\
p^f(1-q_a^l)Z_1(0,0,0,0), & a=1, b=0, i=j=0 \\
q_a^f(1-q_a^l)Z_3(1,0,0), & a=0, b=1, i=j=0 \\
(1-q_a^l)\Big[p^f Z_3(1,0,0) + (1-p^f)\big[Z_3(2,0,0) + Z_1(0,0,0,0) + Z_1(1,0,0,0) + Z_1(0,1,0,0)\big]\Big] & \\
+q_a^l(1-pf)Z_1(0,0,0,0), & a=1, b=1, i=j=0 \\
q_a^l q_a^f Z_3(0,0,0), & a=0, b=1, i=0, j=1 \\
q_a^l(1-q_a^f)Z_3(1,0,0), & a=b=0, i=0, j=1 \\
q_a^l[p^f Z_3(0,0,0) + (1-p^f)Z_3(1,0,0)], & a=b=1, i=0, j=1 \\
p^l(1-q_a^f)Z_1(0,0,0,0), & a=b=0, i=1, j=0 \\
p^l(1-p^f)Z_1(0,0,0,0), & a=b=1, i=1, j=0 \\
(1-q_a^f)\Big[p^l Z_3(1,0,0) + (1-p^l)(Z_3(2,0,0) + Z_1(0,0,0,0) + Z_1(1,0,0,0) + Z_1(0,1,0,0)\Big] & \\
+q_a^f(1-pl)Z_1(0,0,0,0), & a=b=0, i=j=1 \\
q_a^f[p^l Z_3(0,0,0) + (1-p^l)Z_3(1,0,0)], & a=0, b=1, i=j=1 \\
p^f(1-p^l)Z_1(0,0,0,0), & a=1, b=0, i=j=1 \\
\big[Q_a^l(1,n)Q_a^f(0,n') + Q_a^l(0,n)Q_a^f(1,n')\big]Z_4(0) & \\
+Q_a^l(0,n)Q_a^f(0,n')\sum\limits_{x=0,x\neq 1}^{n+n'+2} Z_4(x), & a=b=1, i=j=1
\end{cases}$$

$$\bullet P_{(n,n',i,a),(n+1,n',j,b)}\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right) =$$

$$\begin{cases}
(1-q_a^l)(1-q_a^f)Z_3(1,1,0), & a=b=0, i=j=0 \\
q_a^f(1-q_a^l)Z_3(0,1,0), & a=0, b=1, i=j=0 \\
(1-q_a^l)[p^f Z_3(0,1,0) + (1-p^f)Z_3(1,1,0)], & a=1, b=1, i=j=0 \\
q_a^l(1-q_a^f)Z_3(0,1,0), & a=b=0, i=0, j=1 \\
q_a^l q_a^f Z_3(0,1,0), & a=0, b=1, i=0, j=1 \\
q_a^l Z_3(0,1,0), & a=b=1, i=0, j=1 \\
(1-q_a^f)[p^l Z_3(0,1,0) + (1-p^l)Z_3(1,1,0)], & a=b=0, i=j=1 \\
q_a^f Z_3(0,1,0), & a=0, b=1, i=j=1 \\
\big[1-Q_r^l(0,n)Q_r^f(0,n')(1-p^l)(1-p^f)\big]Q_a^l(1,n)Q_a^f(0,n'), & a=b=1, i=j=1
\end{cases}$$

$\bullet P_{(n,n',i,a),(n,n'+1,j,b)}\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right) =$

$$\begin{cases} (1-q_a^l)(1-q_a^f)Z_3(1,0,1), & a=b=0, i=j=0 \\ q_a^f(1-q_a^l)Z_3(0,0,1), & a=0, b=1, i=j=0 \\ (1-q_a^l)[p^f Z_3(0,0,1)+(1-p^f)Z_3(1,0,1)], & a=1, b=1, i=j=0 \\ q_a^l(1-q_a^f)Z_3(0,0,1), & a=b=0, i=0, j=1 \\ q_a^l q_a^f Z_3(0,0,1), & a=0, b=1, i=0, j=1 \\ q_a^l Z_3(0,0,1), & a=b=1, i=0, j=1 \\ (1-q_a^f)[p^l Z_3(0,0,1)+(1-p^l)Z_3(1,0,1)], & a=b=0, i=j=1 \\ q_a^f Z_3(0,0,1), & a=0, b=1, i=j=1 \\ \left[1-Q_r^l(0,n)Q_r^f(0,n')(1-p^l)(1-p^f)\right]Q_a^l(0,n)Q_a^f(1,n'), & a=b=1, i=j=1 \end{cases}$$

• And for cases where $2 \leq k+k' \leq m^l+m^f-n-n'-2$ we have
$P_{(n,n',i,a),(n+k,n'+k',j,b)}\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right) =$

$$\begin{cases} Q_a^l(k,n)Q_a^f(k',n'), & i=j=a=b=1 \\ Q_a^l(k,n)Q_a^f(k',n')(1-q_a^f), & a=b=0, i=j=1 \\ Q_a^l(k,n)Q_a^f(k',n')(1-q_a^l), & a=b=1, i=j=0 \\ Q_a^l(k,n)Q_a^f(k',n')(1-q_a^l)^{Z_5(i,j)}(1-q_a^f)^{Z_5(a,b)}q_a^{l\,(1-Z_5(i,j))}q_a^{f\,(1-Z_5(a,b))}, & \text{else.} \end{cases}$$

$\bullet P_{(n,n',i,a),(n+k,n'+k',j,b)}\left(\bar{q}^l, p^l, \bar{q}^f, p^f\right) = 0$ for all other case. This is due to the fact that no more than one (fixed-length) packet can be successfully sent in a slot.

The generic functions Z_i, for $i = 1, \ldots 5$ are given by

- $Z_1(s,t,u,v) = Q_a^l(s,n)Q_a^f(t,n')Q_r^l(u,n)Q_r^f(v,n')$ is the probability to have s new arrivals among the set of idle leaders (and respectively, t new arrival among the set of idle followers), given that u backlogged leaders and v backlogged followers users attempt retransmission on the current slot.

- $Z_2(x) = \sum_{\substack{s,t,u,v\in\{0,1\} \\ s+t+u+v=x}} Q_a^l(s,n)Q_a^f(t,n')(1-q_a^l)^{1-u}(1-q_a^f)^{1-v}(q_a^l)^u(q_a^f)^v$ is the probability that there is exactly x new packets those enter the system.

- $Z_3(x,u,v) = Q_a^l(u,n)Q_a^f(v,n')\sum_{s=0,t=0}^{n,n'} Q_r^l(s,n)Q_r^f(t,n')\delta_{\{t+s\geq x\}}$ gives the probability that the set of idle leaders (and, respectively, the set of idle followers) generate u (respectively v) new packets and, at least, x backlogged users attempt new retransmission.

- $Z_4(x) = \sum_{\substack{s=0,t=0 \\ u=0,v=0}}^{n,n',1,1} Q_r^l(s,n) Q_r^f(t,n') (1-p^l)^{1-u} (1-p^f)^{1-v} (p^l)^u (p^f)^v \delta_{\{s+t+u+v=x\}}$ is the probability that x backlogged users attempt new retransmission.

- $Z_5(x,y) = \begin{cases} 1 & \text{if} \quad x=0 \quad \text{and} \quad y=0, \\ 0 & \text{if} \quad x=0 \quad \text{and} \quad y=1. \end{cases}$

Appendix E : Transition matrix of the controlled hierarchical slotted aloha.

$$P_{(n,a,i),(n+k,b,j)}\left(\bar{q}^f, p^f, \gamma, \lambda\right) =$$

$$\begin{cases}
Q_a(k,n)\,(1-q_a)^{1-b}\,(1-\lambda)^{1-j}\,{q_a}^{b-a}\lambda^{j-i}, & a\le b, i\le j \;\Big\}\; 2\le k\le m-n-1 \\
\left.\begin{array}{ll}
Q_a(1,n)(1-q_a)(1-\lambda)\Big[1-Q_r(0,n)\Big], & a=b=0, i=j=0 \\
Q_a(1,n)(1-q_a)\lambda, & a=b=0, i=0, j=1 \\
Q_a(1,n)(1-\lambda)q_a, & a=0, b=1, i=j=0 \\
Q_a(1,n)(1-\lambda)\Big[1-(1-p^f)Q_r(0,n)\Big], & a=1, b=1, i=j=0 \\
Q_a(1,n)q_a\lambda, & a=0, b=1, i=0, j=1 \\
Q_a(1,n)(1-q_a)\Big[1-(1-\gamma)Q_r(0,n)\Big], & a=b=0, i=j=1 \\
Q_a(1,n)\lambda, & a=b=1, i=0, j=1 \\
Q_a(1,n)q_a, & a=0, b=1, i=j=1 \\
Q_a(1,n)\Big[1-(1-p^f)(1-\gamma)Q_r(0,n)\Big], & a=b=1, i=j=1
\end{array}\right\} k=1 \\
\left.\begin{array}{ll}
Q_r(0,n)\Big[Q_a(0,n)\,(1-q_a\lambda)+Q_a(1,n)(1-q_a)(1-\lambda)\Big], & a=b=0, i=j=0 \\
Q_a(0,n)Q_r(0,n)(1-\lambda)p^f, & a=1, b=0, i=j=0 \\
Q_a(0,n)(1-q_a)\left[1-Q_r(0,n)\right]\lambda, & a=b=0, i=0, j=1 \\
Q_a(0,n)(1-\lambda)\left[1-Q_r(0,n)\right]q_a, & a=0, b=1, i=j=0 \\
Q_a(0,n)\Big\{q_aQ_r(0,n)(1-p^f)+(1-\lambda)\Big[1-p^fQ_r(0,n)-(1-p^f)Q_r(1,n)\Big]\Big\} & \\
+Q_a(1,n)Q_r(0,n)\,(1-\lambda)\left(1-p^f\right), & a=1, b=1, i=j=0 \\
Q_a(0,n)q_a\lambda, & a=0, b=1, i=0, j=1 \\
Q_a(0,n)\left\{p^f+(1-p^f)[1-Q_r(0,n)]\right\}\lambda, & a=b=1, i=0, j=1 \\
Q_a(0,n)Q_r(0,n)(1-q_a)\gamma, & a=b=0, i=1, j=0 \\
Q_a(0,n)Q_r(0,n)(1-p^f)\gamma, & a=b=1, i=1, j=0 \\
Q_a(0,n)\Big\{q_aQ_r(0,n)(1-\gamma)+(1-q_a)\Big[1-\gamma Q_r(0,n)-(1-\gamma)Q_r(1,n)\Big]\Big\} & \\
+Q_a(1,n)Q_r(0,n)\,(1-q_a)\,(1-\gamma), & a=b=0, i=j=1 \\
Q_a(0,n)q_a\left\{\gamma+(1-\gamma)\Big[1-Q_r(0,n)\Big]\right\}, & a=0, b=1, i=j=1 \\
Q_a(0,n)Q_r(0,n)p^f(1-\gamma), & a=1, b=0, i=j=1 \\
Q_a(0,n)\displaystyle\sum_{j=0,s=0,t=0}^{n,1,1} Q_r(j,n)\,(\gamma)^s\left(p^f\right)^t(1-\gamma)^{1-s}\left(1-p^f\right)^{1-t}\mathbf{1}_{j+s+t\ge 2} & \\
+Q_a(1,n)Q_r(0,n)\,(1-\gamma)\left(1-p^f\right), & a=b=1, i=j=1
\end{array}\right\} k=0 \\
Q_a(0,n)Q_r(1,n)\,(1-q_a)^{1-a}\,(1-\lambda)^{1-i}\left(1-p^f\right)^a(1-\gamma)^i, \quad a=b, i=j \;\Big\} \quad k=-1 \\
0, \qquad \text{otherwise.}
\end{cases}$$

Appendix F : Transition matrix for the game problem under scheme 1, $P_{(n,a),(n+i,b)} =$

$$
\begin{cases}
\left.\begin{array}{ll}
q_r^{m+1} Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n) \frac{A_{j+1,i}}{j+i+1}, & a=1, b=0 \\
Q_a(i,n) q_a \sum_{j=0}^{n} Q_r(j,n)(1-A_{j,i+1}), & a=0, b=1 \\
Q_a(i,n)[(1-q_a) \sum_{j=0}^{n} Q_r(j,n)(1-A_{j,i}) + q_a \sum_{j=0}^{n} Q_r(j,n) \frac{A_{j,i+1}}{j+i+1}], & a=0, b=0 \\
Q_a(i,n)[(1-q_r^{m+1}) \sum_{j=0}^{n} Q_r(j,n)(1-A_{j,i}) + q_r^{m+1} \sum_{j=0}^{n} Q_r(j,n)(1-A_{j+1,i})], & a=1, b=1
\end{array}\right\} i = (m-n) \geq 2 \\
\left.\begin{array}{ll}
q_r^{m+1} Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n) \frac{A_{j+1,i}}{j+i+1}, & a=1, b=0 \\
q_a[Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n)(1-A_{j,i+1}) + Q_a(i+1,n) \sum_{j=0}^{n} \frac{j+i+1}{j+i+2} Q_r(j,n) A_{j,i+2}], & a=0, b=1 \\
(1-q_a)[Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n)(1-A_{j,i}) + Q_a(i+1,n) \sum_{j=0}^{n} Q_r(j,n) A_{j,i+1}] + & \\
q_a Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n) \frac{A_{j,i+1}}{j+i+1}, & a=0, b=0 \\
(1-q_r^{m+1})[Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n)(1-A_{j,i}) + Q_a(i+1,n) \sum_{j=0}^{n} Q_r(j,n) A_{j,i+1}] + & \\
q_r^{m+1}[Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n)(1-A_{j+1,i}) + Q_a(i+1,n) \sum_{j=0}^{n} \frac{j+i+1}{j+i+2} Q_r(j,n) A_{j+1,i+1}], & a=1, b=1
\end{array}\right\} 2 \leq i < m-n \\
\left.\begin{array}{ll}
q_r^{m+1} Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n) \frac{A_{j+1,1}}{j+2}, & a=1, b=0 \\
q_a[Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n)(1-A_{j,2}) + Q_a(2,n) \sum_{j=0}^{n} \frac{j+2}{j+3} Q_r(j,n) A_{j,3}], & a=0, b=1 \\
(1-q_a)[Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n)(1-A_{j,1}) + Q_a(2,n) \sum_{j=0}^{n} Q_r(j,n) A_{j,2}] + q_a Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n) \frac{A_{j,2}}{j+2}, & a=0, b=0 \\
(1-q_r^{m+1})[Q_a(1,n) \sum_{j=1}^{n} Q_r(j,n)(1-A_{j,1}) + Q_a(2,n) \sum_{j=0}^{n} Q_r(j,n) A_{j,2}] + & \\
q_r^{m+1}[Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n)(1-A_{j+1,1}) + Q_a(2,n) \sum_{j=0}^{n} \frac{j+2}{j+3} Q_r(j,n) A_{j+1,2}], & a=1, b=1
\end{array}\right\} i = 1 \\
\left.\begin{array}{ll}
q_r^{m+1} Q_a(0,n) \sum_{j=0}^{n} Q_r(j,n) \frac{A_{j+1,0}}{j+1}, & a=1, b=0 \\
q_a[Q_a(0,n) \sum_{j=0}^{n} Q_r(j,n)(1-A_{j,1}) + Q_a(1,n) \sum_{j=0}^{n} \frac{j+1}{j+2} Q_r(j,n) A_{j,2}], & a=0, b=1 \\
(1-q_a)[Q_a(0,n) \sum_{j=0, j\neq 1}^{n} Q_r(j,n)(1-A_{j,0}) + Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n) A_{j,1}] + q_a Q_a(0,n) \sum_{j=0}^{n} Q_r(j,n) \frac{A_{j,1}}{j+1}, & a=0, b=0 \\
(1-q_r^{m+1})[Q_a(0,n) \sum_{j=0, j\neq 1}^{n} Q_r(j,n)(1-A_{j,0}) + Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n) A_{j,1}] + & \\
q_r^{m+1}[Q_a(0,n) \sum_{j=1}^{n} Q_r(j,n)(1-A_{j+1,0}) + Q_a(1,n) \sum_{j=0}^{n} \frac{j+1}{j+2} Q_r(j,n) A_{j+1,1}], & a=1, b=1
\end{array}\right\} i = 0 \\
\left.\begin{array}{ll}
q_a Q_a(0,n) \sum_{j=1}^{n} \frac{j}{j+1} Q_r(j,n) A_{j+1}, & a=0, b=1 \\
Q_a(0,n)[(1-q_r^{m+1}) \sum_{j=1}^{n} Q_r(j,n) A_{j,0} + q_r^{m+1} \sum_{j=1}^{n} \frac{j}{j+1} Q_r(j,n) A_{j+1}], & a=1, b=1 \\
(1-q_a) Q_a(0,n) \sum_{j=1}^{n} Q_r(j,n) A_{j,0}, & a=0, b=0 \\
0, & a=1, b=0
\end{array}\right\} i = -1
\end{cases}
$$

Appendix G : Transition matrix for the game problem under scheme 2, $P_{(n,a),(n+i,b)} =$

$$
\begin{cases}
\left.\begin{array}{ll}
q_r^{m+1} Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n) B_{j+1,i}, & a=1, b=0 \\
Q_a(i,n) q_a \sum_{j=0}^{n} Q_r(j,n)(1-B_{j,i+1}), & a=0, b=1 \\
Q_a(i,n)(1-q_a) \sum_{j=0}^{n} Q_r(j,n)(1-B_{j,i}), & a=0, b=0 \\
Q_a(i,n)[(1-q_r^{m+1}) \sum_{j=0}^{n} Q_r(j,n)(1-B_{j,i}) + q_r^{m+1} \sum_{j=0}^{n} Q_r(j,n)(1-B_{j+1,i})], & a=1, b=1
\end{array}\right\} i=(m-n)\geq 2 \\
\left.\begin{array}{ll}
q_r^{m+1} Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n) B_{j+1,i} & a=1, b=0 \\
q_a[Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n)(1-B_{j,i+1}) + Q_a(i+1,n) \sum_{j=1}^{n} Q_r(j,n) B_{j,i+2}], & a=0, b=1 \\
(1-q_a)[Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n)(1-B_{j,i}) + Q_a(i+1,n) \sum_{j=0}^{n} Q_r(j,n) B_{j,i+1}], & a=0, b=0 \\
(1-q_r^{m+1})[Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n)(1-B_{j,i}) + Q_a(i+1,n) \sum_{j=1}^{n} Q_r(j,n) B_{j,i+1}] + & \\
q_r^{m+1} Q_a(i,n) \sum_{j=0}^{n} Q_r(j,n)(1-B_{j+1,i}), & a=1, b=1
\end{array}\right\} 2 \leq i < m-n \\
\left.\begin{array}{ll}
q_r^{m+1} Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n) B_{j+1,1}, & a=1, b=0 \\
q_a[Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n)(1-B_{j,2}) + Q_a(2,n) \sum_{j=0}^{n} Q_r(j,n) B_{j,3}], & a=0, b=1 \\
(1-q_a)[Q_a(1,n) \sum_{j=1}^{n} Q_r(j,n)(1-B_{j,1}) + Q_a(2,n) \sum_{j=1}^{n} Q_r(j,n) B_{j,2}], & a=0, b=0 \\
(1-q_r^{m+1})[Q_a(1,n) \sum_{j=1}^{n} Q_r(j,n)(1-B_{j,1}) + Q_a(2,n) \sum_{j=1}^{n} Q_r(j,n) B_{j,2}] + & \\
q_r^{m+1} Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n)(1-B_{j+1,1}), & a=1, b=1
\end{array}\right\} i=1 \\
\left.\begin{array}{ll}
q_r^{m+1} Q_a(0,n) \sum_{j=0}^{n} Q_r(j,n) B_{j+1,0}, & a=1, b=0 \\
q_a[Q_a(0,n) \sum_{j=1}^{n} Q_r(j,n)(1-B_{j,1}) + Q_a(1,n) \sum_{j=1}^{n} Q_r(j,n) B_{j,2}], & a=0, b=1 \\
(1-q_a)[Q_a(0,n) \sum_{j=0,j\neq 1}^{n} Q_r(j,n)(1-B_{j,0}) + Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n) B_{j,1}] + & \\
q_a Q_a(0,n) B_{0,1}, & a=0, b=0 \\
(1-q_r^{m+1})[Q_a(0,n) \sum_{j=0,j\neq 1}^{n} Q_r(j,n)(1-B_{j,0}) + Q_a(1,n) \sum_{j=0}^{n} Q_r(j,n) B_{j,1}] + & \\
q_r^{m+1} Q_a(0,n) \sum_{j=1}^{n} Q_r(j,n)(1-B_{j+1,0}), & a=1, b=1
\end{array}\right\} i=0 \\
\left.\begin{array}{ll}
q_a Q_a(0,n) \sum_{j=1}^{n} Q_r(j,n) B_{j,1}, & a=0, b=1 \\
Q_a(0,n)(1-q_r^{m+1}) \sum_{j=1}^{n} Q_r(j,n) B_{j,0}), & a=1, b=1 \\
(1-q_a) Q_a(0,n) \sum_{j=1}^{n} Q_r(j,n) B_{j,0}, & a=0, b=0 \\
0, & a=1, b=0
\end{array}\right\} i=-1
\end{cases}
$$

Appendix H : Transition matrix for the game problem under scheme 3, $P_{(n,a),(n+i,b)} =$

$$
\begin{cases}
\left.\begin{array}{ll} Q_a(i,n)(1-q_a), & a=0,b=0 \\ Q_a(i,n), & a=1,b=1 \\ Q_a(i,n)q_a, & a=0,b=1 \end{array}\right\} 2 \le i \\
\left.\begin{array}{ll} Q_a(1,n)(1-q_a)\sum_{j=1}^{n} Q_r(j,n)(1-C_{j,1}), & a=0,b=0 \\ Q_a(1,n)((1-q_r^{m+1})\sum_{j=1}^{n} Q_r(j,n)(1-C_{j,1})+q_r^{m+1}\sum_{j=0}^{n} Q_r(j,n)(1-C_{j+1,1})), & a=1,b=1 \\ Q_a(i,n)q_a, & a=0,b=1 \end{array}\right\} i=1 \\
\left.\begin{array}{ll} Q_a(0,n)[(1-q_a)\sum_{j=0,j\neq 1}^{n} Q_r(j,n)(1-C_{j,0})+q_a\sum_{j=0}^{n} Q_r(j,n)C_{j,1}]+ & \\ \quad Q_a(1,n)(1-q_a)\sum_{j=0}^{n} Q_r(j,n)C_{j,1}, & a=0,b=0 \\ Q_a(0,n)[(1-q_r^{m+1})\sum_{j=0,j\neq 1}^{n} Q_r(j,n)(1-C_{j,0})+q_r^{m+1}\sum_{j=1}^{n} Q_r(j,n)C_{j+1,0}]+ & \\ \quad Q_a(1,n)[(1-q_r^{m+1})\sum_{j=0,j\neq 1}^{n} Q_r(j,n)C_{j,1}+q_r^{m+1}\sum_{j=0}^{n} Q_r(j,n)C_{j+1,1}], & a=1,b=1 \\ Q_a(0,n)q_a\sum_{j=1}^{n} Q_r(j,n)(1-C_{j,1}), & a=0,b=1 \\ Q_a(0,n)q_r^{m+1}\sum_{j=0}^{n}\frac{1}{j+1}Q_r(j,n), & a=1,b=0 \end{array}\right\} i=0 \\
\left.\begin{array}{ll} Q_a(0,n)(1-q_a)\sum_{j=1}^{n} Q_r(j,n)C_{j,0}, & a=0,b=0 \\ Q_a(0,n)((1-q_r^{m+1})\sum_{j=1}^{n} Q_r(j,n)C_{j,0}+q_r^{m+1}\sum_{j=1}^{n}\frac{j}{j+1}Q_r(j,n)C_{j+1,0}), & a=1,b=1 \end{array}\right\} i=-1 \\
0, \quad \text{otherwise}
\end{cases}
$$

Appendix I : Transition matrix for the game problem under scheme 4, $P_{(n,a),(n+i,b)} =$

$$
\begin{cases}
\left.\begin{array}{ll}
Q_a(i,n)q_a\sum_{j=0}^{n}Q_r(j,n)(1-C_{j,i+1}), & a=0,b=1\\
Q_a(i,n)[(1-q_a)\sum_{j=0}^{n}Q_r(j,n)(1-C_{j,i})+q_a\sum_{j=0}^{n}Q_r(j,n)\frac{C_{j,i+1}}{i+1}], & a=0,b=0\\
Q_a(i,n)[(1-q_r^{m+1})\sum_{j=0}^{n}Q_r(j,n)(1-C_{j,i})+q_r^{m+1}\sum_{j=0}^{n}Q_r(j,n)(1-C_{j+1,i})], & a=1,b=1
\end{array}\right\} i=(m-n)\geq 2\\
\left.\begin{array}{ll}
q_a[Q_a(i,n)\sum_{j=0}^{n}Q_r(j,n)(1-C_{j,i+1})+Q_a(i+1,n)\sum_{j=0}^{n}\frac{i+1}{i+2}Q_r(j,n)C_{j,i+2}], & a=0,b=1\\
Q_a(i,n)[(1-q_a)\sum_{j=0}^{n}Q_r(j,n)(1-C_{j,i})+q_a\sum_{j=0}^{n}Q_r(j,n)\frac{C_{j,i+1}}{i+1}]+ & \\
(1-q_a)Q_a(i+1,n)\sum_{j=0}^{n}Q_r(j,n)C_{j,i+1}, & a=0,b=0\\
Q_a(i,n)[(1-q_r^{m+1})\sum_{j=0}^{n}Q_r(j,n)(1-C_{j,i})+q_r^{m+1}\sum_{j=0}^{n}Q_r(j,n)(1-C_{j+1,i})]+ & \\
Q_a(i+1,n)[(1-q_r^{m+1})\sum_{j=0}^{n}Q_r(j,n)C_{j,i+1}+q_r^{m+1}\sum_{j=0}^{n}Q_r(j,n)C_{j+1,i+1}], & a=1,b=1
\end{array}\right\} 2\leq i<m-n\\
\left.\begin{array}{ll}
q_a[Q_a(1,n)\sum_{j=0}^{n}Q_r(j,n)(1-C_{j,2})+Q_a(2,n)\sum_{j=0}^{n}\frac{2}{3}Q_r(j,n)C_{j,3}], & a=0,b=1\\
Q_a(1,n)[(1-q_a)\sum_{j=1}^{n}Q_r(j,n)(1-C_{j,1})+q_a\sum_{j=0}^{n}Q_r(j,n)\frac{C_{j,2}}{2}]+ & \\
(1-q_a)Q_a(2,n)\sum_{j=0}^{n}Q_r(j,n)C_{j,2}, & a=0,b=0\\
Q_a(1,n)[(1-q_r^{m+1})\sum_{j=1}^{n}Q_r(j,n)(1-C_{j,1})+q_r^{m+1}\sum_{j=0}^{n}Q_r(j,n)(1-C_{j+1,1})]+ & \\
Q_a(2,n)[(1-q_r^{m+1})\sum_{j=0}^{n}Q_r(j,n)C_{j,2}+q_r^{m+1}\sum_{j=0}^{n}Q_r(j,n)C_{j+1,2}], & a=1,b=1
\end{array}\right\} i=1\\
\left.\begin{array}{ll}
q_a[Q_a(0,n)\sum_{j=1}^{n}Q_r(j,n)(1-C_{j,1})+Q_a(1,n)\sum_{j=0}^{n}\frac{1}{2}Q_r(j,n)C_{j,2}], & a=0,b=1\\
Q_a(0,n)q_r^{m+1}Q_r(0,n), & a=1,b=0\\
Q_a(0,n)[(1-q_a)\sum_{j=0,j\neq 1}^{n}Q_r(j,n)(1-C_{j,0})+q_a\sum_{j=0}^{n}Q_r(j,n)C_{j,1}]+ & \\
Q_a(1,n)(1-q_a)\sum_{j=0}^{n}Q_r(j,n)C_{j,1}, & a=0,b=0\\
Q_a(0,n)[(1-q_r^{m+1})\sum_{j=0,j\neq 1}^{n}Q_r(j,n)(1-C_{j,0})+q_r^{m+1}\sum_{j=1}^{n}Q_r(j,n)(1-C_{j+1,0})]+ & \\
Q_a(1,n)[(1-q_r^{m+1})\sum_{j=0}^{n}Q_r(j,n)C_{j,1}+q_r^{m+1}\sum_{j=0}^{n}Q_r(j,n)C_{j+1,1}], & a=1,b=1
\end{array}\right\} i=0\\
\left.\begin{array}{ll}
(1-q_a)Q_a(0,n)Q_r(1,n), & a=0,b=0\\
(1-q_r^{m+1})Q_a(0,n)Q_r(1,n), & a=1,b=1
\end{array}\right\} i=-1\\
0, \qquad \text{otherwise}
\end{cases}
$$

List of publications

International Journals

1. Rachid El-Azouzi, **Essaïd Sabir**, Sujit K. Samanta and Ralph El-Koury. "Asymptotic delay analysis and timeout-based admission control for ad-hoc wireless networks and Asymmetric users". To appear in the Elsiever International Journal of Computer Communications, 2010.
2. Rachid El-Azouzi, **Essaïd Sabir**, Tania Jiminez and El-Houssine bouyakhf. "Modeling Slotted Aloha as a Stochastic Game with Random Discrete Power Selection Algorithms". International Journal of Computer Systems, Networks, and Communications, 2009.
3. **Essaïd Sabir**, Rachid El-Azouzi, and Yezekael Hayel. "Towards Sustaining Partial Cooperation in Slotted Aloha-like Protocols Using Hierarchy and Lossy Channel". Submitted.
4. **Essaïd Sabir**, Rachid El-Azouzi, Verraruna Kavitha, Yezekael Hayel and El-Houssine bouyakhf. "Stochastic Learning Solution for Constrained Nash Equilibrium Throughput in Non Saturated Wireless Collision Channels". Submitted.

International Conferences

1. Rachid El-Azouzi, **Essaïd Sabir**, Mohamed Raiss and Sujit Samanta. "Understanding and Analyzing Performances of the IEEE-802.11e in Asymmetric multihop Ad hoc Networks". Submitted.
2. **Essaïd Sabir** and Rachid El-Azouzi. "A Fountain Code-based Fair Scheme for Multi-hop Ad hoc Networks". To appear in IEEE ISABEL, 2010, Italy.
3. **Essaïd Sabir**, Rachid El-Azouzi, Verraruna Kavitha, Yezekael Hayel and El-Houssine bouyakhf. "Stochastic Learning Solution for Constrained Nash Equilibrium Throughput in Non Saturated Wireless Collision Channels". Proceedings of ICST/ACM Valuetools (GameComm), October 23, 2009, Pisa, Italy.
4. Rachid El-Azouzi, Sujit K. Samanta, **Essaïd Sabir** and Ralph El-Koury. "Asymptotic delay analysis and timeout-based admission control for ad-hoc wireless net-

works". Proceedings of the eighth International Conference on Ad-Hoc Networks and Wireless (AdHoc-Now'09), September 22-25, 2009, Murcia, Spain.

5. Rachid El-Azouzi, **Essaïd Sabir**, Ralph El-Khoury, Sujit Kumar Samanta and El-Houssine Bouyakhf. "An End-to-End QoS Framework for IEEE 802.16 and Ad-hoc Integrated Networks". Proceedings of the ACM International Conference on Mobile Technology, Applications and Systems, September 02-04, 2009, France.
6. **Essaïd Sabir**, Rachid El-Azouzi and Yezekael Hayel. "A Hierarchical Slotted Aloha Game". Proceedings of the IEEE/ICST International Conference on Game Theory for Networks (GameNets'09), 13-15 Mai 2009, Istanbul, Turkey.
7. Rachid El-Azouzi, Ralph El-Khoury, Abdellatif Kobbane and **Essaïd Sabir**. "On extending coverage of UMTS networks using an Ad-Hoc network with weighted fair queueing". Proceedings of IFIP Networking 2008, 4-7 Mai 2008, Singapore.
8. **Essaïd Sabir** and Rachid El-Azouzi. "Achieving Constrained Nash Equilibrium in Wireless Collision Channels". Proceedings of ROADEF'08, February 25-27, 2008, Clérmont-Ferrand, France.
9. Rachid El-Azouzi, **Essaïd Sabir**, Tania Jimenez, Souad Benarfa and El-Houssine Bouyakhf. "Cooperative and Non-cooperative control for Slotted Aloha with random power level selections algorithms". Proceedings of the 2nd ICST/ACM International Conference on Performance Evaluation Methodologies and Tools (Valuetools), 24-26 octobre 2007, Nantes, France.

National Conferences

1. **Essaïd Sabir**, Mohammed Raiss and Mohamed Elkamili. "Slotted Aloha à Première Transmission Différée : Une Nouvelle Solution pour Supporter les Applications Sensibles au Délai". Proceedings of MajecSTIC 2009, November 16-18, 2009, Avignon, France.
2. **Essaïd Sabir** and El-Houssine Bouyakhf. "L'impact des niveaux de puissance sur les performances de Slotted Aloha". Journées Scientifiques en Technologies de l'Information et de la Communications (JOSTIC'07), February 24, 2007,Rabat, Morocco.

Technical Reports

1. El-Houssine Bouyakhf, Rachid El-Azouzi, Khalil Ibrahimi, **Essaïd Sabir** et Sidi Habib. "Capacité des réseaux cellulaires avec des applications multimédias", Maroc Telecom R&D Project, October, 2009.
2. El-Houssine Bouyakhf, Rachid El-Azouzi, Khalil Ibrahimi, **Essaïd Sabir** et Ralph El-Khoury. "Capacité des réseaux UMTS et HSDPA avec des applications multimédias", Maroc Telecom R&D Project, October, 2008.

Bibliography

[1] IEEE Std 802.11-2007. Part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications.

[2] N. Abramson. The aloha system – another alternative for computer communications. In *Proceedings of the AFIPS Conference*, volume 36, pages 295–298, 1970.

[3] N. Abramson. The throughput of packet broadcasting channels. *IEEE Transactions on Communications*, 25:117–128, Juanary 1977.

[4] Ian F. Akyildiz and Xudong Wang. A survey on wireless mesh networks. *EEE Radio Communications*, 46(101 112 v3.2.0):23–30, September 2005.

[5] Ian F. Akyildiz and Xudong Wang. *Wireless Mesh Networks*. John Wiley & Sons, Ltd, Publication. ISBN 9780470032565, 2009.

[6] B. Alawieh, C. Assi, and H. T. Mouftah. Investigation of power-aware ieee 802.11 performance in multi-hop ad hoc networks. In *Proceedings of the Third international conference on Mobile ad-hoc and sensor networks (MSN)*, pages 409–420, Beijing, China, December 2007.

[7] F. Alizadeh-Shabdiz and S. Subramaniam. Analytical models for single-hop and multi-hop ad hoc networks. *Mobile Networks Application*, 11:75–90, 2006.

[8] E. Altman, D. Barman, A. Benslimane, and R. El-Azouzi. Slotted aloha with priorities and random power. In *Proceedings of the IFIP Networking*, Ontario, Canada, 2006.

[9] E. Altman, D. Barman, R El-Azouzi, and T. Jimenez. A game theoretic approach for delay minimization in slotted aloha. In *Proceedings of the IEEE International Conference on Communications*, Paris, France, june 2004.

[10] E. Altman, N. Bonneau, and M. Debbah. Correlated equilibrium in access control for wireless communications. In *Proceedings of the IFIP Networking*, Coimbra, Portugal, 2006.

[11] E. Altman, T. Boulogne, R. El Azouzi, T. Jiménez, and L. Wynter. A survey on networking games in telecommunications. *Computers and Operations Research*, 33(2):286–311, 2006.

[12] E. Altman, R El-Azouzi, and T. Jimenez. Slotted aloha as a stochastic game with partial information. In *Proceedings of the Modeling and Optimization in Mobile, Ad-Hoc and Wireless Networks (WiOpt'03)*, Sophia Antipolis, France, March 2003.

[13] E. Altman, R El-Azouzi, and T. Jimenez. Slotted aloha as a stochastic game with partial information. *Computer Networks*, 45, 2004.

[14] E. Altman, S. Shakkottai, and A. Kumar. The case for non-cooperative multihoming of users to access points in ieee 802.11 wlans. In *Proceedings of the IEEE Infocom*, pages 1–12, Spain, 2006.

[15] V. Anantharam. The stability region of the finite-user slotted aloha protocol. *IEEE Transactions on Information Theory*, 37:535–540, May 1991.

[16] G. E. Andrews. *The theory of partitions*. Addison-Wesley Pub. Co., Advanced Book Program (Reading, Mass), 1976.

[17] Jeffrey G. Andrews. *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*. ISBN: 0132225522, Prentice Hall PTR, Pages: 496, 2007.

[18] J. C. Arnbak and W. Blitterswijk. Capacity of slotted aloha in rayleigh-fading channels. *IEEE Journal on Selected Areas in Communications*, 5:261–269, February 1987.

[19] G. Arslan, M. Fatih Demirkol, and Y. Song. Equilibrium efficiency improvement in mimo interference systems: a decentralized stream control approach. *IEEE Transaction on Wireless Communications*, 6:2984–2993, 2007.

[20] A. Banchs and X. Perez. Providing throughput guarantees in ieee 802.11 wireless lan. In *Proceedings of the IEEE WCNC*, volume 1, pages 130–8, 2002.

[21] N. Bansal and Z. Liu. Capacity, delay and mobility in wireless ad-hoc networks. In *Proceedings of the IEEE Infocom*, pages 1553–1563, 2003.

[22] J. S. Baras, V. Tabatabaee, G. Papageorgiou, and N. Rentz. Modelling and optimization for multi-hop wireless networks using fixed point and automatic differentiation. In *Proceedings of the 6th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Berlin, Germany, March 2008.

[23] Y. Barowski, S. Biaz, and P. Agrawal. Towards the performance analysis of ieee 802.11 in multi-hop ad-hoc networks. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, pages 100–106, New Orleans, March 2005.

[24] B. Baynat, G. Nogueira, M. Maqbool, and M. Coupechoux. An efficient analytical model for the dimensioning of wimax networks. In *Proceedings of the IFIP Networking*, pages 521–534, 2009.

[25] D. Bertsekas and R. Gallager. *Data Networks*. PRENTICE HALL, Englewood Cliffs, New Jersey, 1987.

[26] G. Bianchi. Performance analysis of the ieee 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18:535–547, 2000.

[27] N. Bisnik and A.A. Abouzeid. Queuing network models for delay analysis of multihop wireless ad hoc networks. *Ad Hoc Networks*, 7:79–97, 2009.

[28] M. Bloem, T. Alpcan, and T. Basar. A stackelberg game for power control and channel allocation in cognitive radio networks. In *Proceedings of the ACM/ICST GameComm workshop*, Nantes, France, October 2007.

[29] A. Bria. Joint resource management of cellular and broadcasting systems-research challenges. In *Proceedings of the RadioVetenskap och Kommunikation (RVK)*, 2005.

[30] R. Bruno, M. Conti, and E. Gregori. Mesh networks: Commodity multi-hop ad hoc networks. *IEEE Communications Magazine*, pages 123–131, 2005.

[31] J. W. Byers, M. Luby, and W. Mitzenmacher. A digital fountain approach to asynchronous reliable multicast. *IEEE Journal on Selected Area in Communications*, 20:1528–1540, 2002.

[32] J. Camp, E. Aryafar, and E. Knightly. Coupled 802.11 flows in urban channels: Model and experimental evaluation. In *Proceedings of the IEEE Infocom*, San Diego, CA, March 2010.

[33] P. Casari, M. Rossi, and M Zorzi. Fountain codes and their application to broadcasting in underwater networks: Performance modeling and relevant tradeoffs. In *Proceedings of the ACM WUWNet'08*, San Francisco, California, USA., September 2008.

[34] Guner Dincer CELIK. Master thesis dissertation: Distributed mac protocol for networks with multipacket reception capability and spatially distributed nodes, May 2007.

[35] P. Chatzimisios, A. Boucouvalas, and V. Vitsas. Ieee 802.11 packet delay: A finite retry limit analysis. In *Proceedings of the IEEE Globecom*, volume 2, pages 950–954, San Francisco, USA, December 2003.

[36] Kwang-Cheng Chen and J. Roberto B. de Marca. *Mobile WiMAX*. John Wiley & Sons, Ltd, Publication. ISBN 978-0-470-51941-7, 2008.

[37] I. Cidon, H. Kodesh, and M. Sidi. Erasure, capture and random power level selection in multi-access systems. *IEEE Transactions in Commununications*, 36:263–271, March 1988.

[38] M. Conti, G. Maselli, G. Turi, and S. Giordano. Cross layering in mobile ad hoc network design. *IEEE Computer*, 37:48–51, February 2004.

[39] S. Conti and M. Giordano. Multihop ad hoc networking: The reality. *IEEE Communications Magazine*, 45:88–95, 2007.

[40] R. B. Cooper. *Introduction to Queueing Theory*. North Holland, 2nd edition, 1981.

[41] Marcelo Menezes de Carvalho. Ph.D. dissertation: Analytical modeling of medium access control protocols in wireless networks, March 2006.

[42] D. Dutta, A. Goel, and J. Heidemann. Oblivious aqm and nash equilibria. In *Proceedings of the IEEE Infocom*, April 2003.

[43] R. El-Azouzi, R. El-Khoury, A. Kobbane, and E. Sabir. On extending coverage of umts networks using an ad-hoc network with weighted fair queueing. In *Proceedings of the IFIP Networking*, volume 4982, pages 135–148, Singapore, May 2008.

[44] R. El-Azouzi, E. Sabir, S. Benarfa, T. Jimenez, and E. H. Bouyakhf. Cooperative and noncooperative control for slotted aloha with random power level selection algorithms. In *Proceedings of the Second ACM/ICST International Conference on Performance Evaluation Methodologies and Tools (valuetools'07)*, Nantes, France, October 2007.

[45] R. El-Azouzi, E. Sabir, S.K. Samanta R. El-Khoury, and E.H. Bouyakhf. An end-to-end qos framework for ieee 802.16 and ad-hoc integrated networks. In *Proceedings of the ACM Mobility*, volume 4982, Nice, France, September 2009.

[46] R. El-Azouzi, E. Sabir, M. Raiss-El-Fenni, and S. K. Samanta. Understanding and analyzing performances of the ieee-802.11e in asymmetric multi-hop ad hoc networks. technical report, http://lia.univ-avignon.fr/fileadmin/documents/users/intranet/chercheurs/sabir/802etechreport.pdf, 2010.

[47] R. El-Azouzi, S.K. Samanta, E. Sabir, and R. El-Koury. Asymptotic delay analysis and timeout-based admission control for ad-hoc wireless networks. In *Proceedings of the eighth International Conference on Ad-Hoc Networks and Wireless (AdHoc-Now'09)*, volume 5793, pages 83–97, Murcia, Spain, 2009.

[48] R. El-Khoury and R. El-Azouzi. Dynamic retransmission limit scheme for routing in multi-hop ad hoc networks. In *Proceedings of the ICST/ACM Inter-Perf'07 Workshop*, Nantes, France, 2007.

[49] R. El-Khoury, R. El-Azouzi, and E. Altman. Delay analysis for real-time streaming media in multi-hop ad-hoc networks. In *Proceedings of the Modeling and Optimization in Mobile, Ad-Hoc and Wireless Networks (WiOpt'08)*, pages 419–428, Germany, 2008.

[50] Ralph El-Khoury. Ph.D. dissertation: Evaluation des performances dans les réseaux ad-hoc sans fil, September 2008.

[51] Byers et al. A digital fountain approach to the reliable distribution of bulk data. *IEEE Journal on Selected Areas in Communications*, 20:1528–1540, 2002.

[52] E. Park et al. Improving quality of service and assuring fairness in wlan access network. *IEEE Transaction on mobile computing*, 6, April 2007.

[53] Kumar et al. Fountain broadcast for wireless networks. In *Second International Workshop on Networked Sensing Systems (INSS)*, 2005.

[54] R. Fantacci and D. Tarchi. Bridging solutions for a heterogeneous wimax-wifi scenario. *Journal of Communications and Networks*, 8:1–9, 2006.

[55] M. Felegyhazi, L. Buttyan, and J. P. Hubaux. Equilibrium analysis of packet forwarding strategies in wireless ad hoc networks the static case. In *Proceedings of the Personal Wireless Communications Conference (PWC)*, Venice, Italy, September 2003.

[56] G. Fodor, A. Eriksson, and A. Tuoriniemi. Providing quality of service in always best connected networks. *IEEE Communications Magazine*, 41:154–163, June 2003.

[57] The Working Group for WLAN Standards. Ieee draft standard 802.11e/ d4.4, june 2003.

[58] J. Freebersyser and B. Leiner. A dod perspective on mobile ad hoc networks. *In C. Perkins, editor, Ad Hoc Networking. Addison Wesley*, pages 29–51, 2001.

[59] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah. Optimal throughput-delay scaling in wireless networks: part i: the fluid model. *IEEE/ACM Transactions Networking*, 14:2568–2592, 2006.

[60] A.E. Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-delay trade-off in wireless networks. In *Proceedings of the IEEE Infocom*, March 2004.

[61] V. Gazis, N. Alonistioti, and L. Merakos. Toward a generic "always best connected" capability in integrated wlan/umts cellular mobile networks (and beyond). *IEEE Wireless Communications*, 12:20–29, 2005.

[62] L. Georgiadis, M. Neely, and L. Tassiulas. Resource allocation and cross layer control in wireless networks. *Foundations and Trends in Networking*, 1, 2006.

[63] S. Ghez, S. Verdu, and S. C. Schwartz. Stability properties of slotted aloha with multipacket reception capability. *IEEE Transactions on Automatic Control*, 33:640–649, July 1988.

[64] S. Ghez, S. Verdu, and S. C. Schwartz. Optimal decentralized control in random-access multipacket channel. *IEEE Transactions on Automatic Control*, 34:1153–1163, November 1989.

[65] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[66] M. Grossglauser and D. Tse. Mobility increases the capacity of ad hoc wireless networks. *IEEE/ACM Transactions on Networking*, 10:477–486, 2002.

[67] Mobile Ad Hoc Network (MANET) Working Group. Internet Engineering Task Force. http://www.ietf.org/html.charters/manet-charter.html.

[68] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46:388–404, 2000.

[69] Z. Hadzi-Velkov and B. Spasenovski. Capture effect in ieee 802.11 basic service area under influence of rayleigh fading and near far effect. In *Proceedings of the IEEE PIMRC*, September 2002.

[70] B. Hajek, A. Krishna, and R. O. LaMaire. On the capture probability for a large number of stations. *IEEE Transactions on Communications*, 45(2):254–260, February 1997.

[71] B. Hajek and T. van Loon. Decentralized dynamic control of a multiaccess broadcast channel. *IEEE Transactions on Automatic Control*, 27(3):559–569, 1982.

[72] Y. Hayel, S. Lasaulce, R. El-Azouzi, and M. Debbah. Introducing hierarchy in energy efficient power control games. In *Proceedings of the ACM/ICST International Conference on Performance Evaluation Methodologies and Tools (valuetools'08, GameComm)*, Athens, Greece, October 2008.

[73] J. He and H. K. Pung. Performance modelling and evaluation of ieee 802.11 distributed coordination function in multihop wireless networks. *Computer Communications*, 29:1300–1308, 2006.

[74] C.V. Hollot, R. Srikant, H. Han, S. Shakkotai, and D. Townsley. Overlay tcp for multi-path routing and congestion control. In *Proceedings of the ENS-INRIA ARC-TCP Workshop*, 2004.

[75] Ekram Hossain. *Heterogeneous Wireless Access Networks Architectures and Protocols*. Springer, ISBN: 978-0-387-09776-3, 2009.

[76] Ekram Hossain and Kin Leung. *Wireless Mesh Networks: Architectures and Protocols*. Springer, ISBN 978-0-387-68838-1, 2008.

[77] C. Hu, H. Kim, and J. C. Hou. Short-term non-uniform access in ieee 802.11-compliant wlans: Astudy on its impact on the saturation performance. *Computer Networks*, 52:61–76, 2008.

[78] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu. Impact of interference on multihop wireless network performance. In *Proceedings of ACM Mobicom*, 2003.

[79] R. Jain. *The Art of Computer Systems Performance Analysis*. John Wiley & Sons, 1991.

[80] R. Jain, D. W. Chiu, and W. R. Hawe. A quantitative measure of fairness and discrimination for resource allocation in shared computer system. dEC research report tr-301., September 1984.

[81] Y. Jin and G. Kesidis. Equilibria of a noncooperative game for heterogeneous users of an aloha network. *IEEE Communications Letters*, 6(7):282–284, 2002.

[82] A. Kherani, R. El-Azouzi, and E. Altman. Stability-throughput tradeoff and routing in multi-hop wireless ad-hoc networks. In *Proceedings of the IFIP Networking*, Coimbra, Portugal, 2006.

[83] A. Kherani, R. El-Khoury, R. El-Azouzi, and E. Altman. Stability-throughput tradeoff and routing in multi-hop wireless ad-hoc networks. *Computer Networks*, 52:1365–1389, 2008.

[84] S.R. Kulkarni and P. Viswanath. A deterministic approach to throughput scaling in wireless networks. *IEEE Transaction on Information Theory*, 50:1041–1049, 2004.

[85] A. Kumar, E. Altman, D. Miorandi, and M. Goyal. New insights from a fixed point analysis of single cell ieee 802.11 wlans. In *Proceedings of the IEEE Infocom*, Miami, March 2005.

[86] D. Kumar, T. Chahed, and E. Altman. Analysis of a fountain codes based transport in an 802.11 wlan cell. In *Proceedings of the 21th International Teletraffic Congress (ITC)*, pages 1–8, Paris, France, 2009.

[87] R. O. LaMaire, A. Krishna, and M. Zorzi. On the randomization of transmitter power levels to increase throughput in multiple access radio systems. *Wireless Networks*, 4:263–277, 1998.

[88] S. Lasaulce, M. Debbah, and E. Altman. Methodologies for analyzing equilibria in wireless games. *IEEE Signal Processing Magazine, Special issue on Game Theory*, 26:51–52, 2009.

[89] C. T. Lau and C. Leung. Capture models for model packet radio networks. *IEEE Transactions in Communications*, 40:917–925, May 1992.

[90] N. Li, J. C. Hou, , and L. Sha. Design and analysis of a mst-based distributed topology control algorithm for wireless ad-hoc networks. *IEEE Transactions on Wireless Communications*, 4:1195–1207, 2005.

[91] X. Lin and N. B. Shroff. The impact of imperfect scheduling on cross-layer congestion control in wireless networks. *IEEE/ACM Transactions Networking*, 14:302–315, 2006.

[92] X. Lin and N.B. Shroff. The multi-path utility maximization problem. In *Proceedings of the 41st Annual Allerton Conference on Communication, Control, and Computing*, October 2003.

[93] J. D. C. Little. A proof for the queueing formula: L $= \lambda$W. *Operations Research*, 9:383–387, May-June 1961.

[94] K. Lu, Y. Qian, and H. Chen. A secure and service-oriented network control framework for wimax networks. *IEEE Communications Magazine*, pages 124–130, 2007.

[95] M. Luby. Lt codes. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 271–282, 2002.

[96] R. T.B. Ma, V. Misra, and D. Rubenstein. Modeling and analysis of generalized slotted-aloha mac protocols in cooperative, competitive and adversarial environments. In *Proceedings of the 26th International Conference on Distributed Computing Systems (ICDCS'06)*, Lisbon, Portugal, 2006.

[97] D. J. C. Mackay. Fountain codes. *IEE Proceedings - Communications*, 152:1062–1068, December 2005.

[98] A. B. MacKenzie and S. B. Wicker. Selfish users in aloha: A game theoretic approach. In *Proceedings of the 54th International Conference on Vehicular Technology (VTC'01)*, volume 3, pages 1354–1357, Atlantic City, NJ, October 2001.

[99] A. B. MacKenzie and S. B. Wicker. Stability of slotted aloha with multipacket reception and selfish users. In *Proceedings of the IEEE Infocom*, volume 3, pages 1583–1590, April 2003.

[100] M.H. Manshaei, G. R. Cantieni, C. Barakat, and T. Turletti. Performance analysis of the ieee 802.11 mac and physical layer protocol. In *Proceedings of the IEEE WoWMoM*, 2005.

[101] P. Marbach and R. Pang. Transmission costs, selfish nodes, and protocol design. In *Proceedings of the Third International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt'05)*, pages 31–40, 2005.

[102] K. Medepalli and F. A. Tobagi. Towards performance modeling of ieee 802.11 based wireless networks: A unified framework and its applications. In *Proceedings of the IEEE Infocom*, volume 2, pages 1–12, Barcelona, Spain, April 2006.

[103] I. Menashe and N. Shimkin. Fixed-rate equilibrium in wireless collision channels. In *Proceedings of the Net-Coop*, pages 23–32, Avignon, France, June 2007.

[104] I. Menashe and N. Shimkin. Efficient rate-constrained nash equilibrium in collision channels with state information. In *Proceedings of the IEEE Infocom*, pages 1076–1084, 2008.

[105] A. Mendez and D. Covarrubias. Stability and optimal retransmission control of s-aloha as a rach channel on wireless networks. In *Proceedings of the 54th International Conference on Vehicular Technology (VTC'01)*, volume 3, pages 1368–1372, Atlantic City, NJ, October 2001.

[106] G. Mergen and L. Tong. Receiver controlled medium access in multihop ad hoc networks with multipacket reception. In *Proceedings of IEEE MILCOM'01*, October 2001.

[107] G. Mergen and L. Tong. Random scheduling medium access for wireless ad hoc networks. In *Proceedings of IEEE MILCOM'02*, October 2002.

[108] G. Mergen and L. Tong. Stability and capacity of regular wireless networks. *IEEE Transactions on Information Theory*, 51:1938–1953, 2005.

[109] F. Meshkati, H.V. Poor, S.C. Schwartz, and N.B. Mandayam. An energy efficient approach to power control and receiver design in wireless data networks. *IEEE Transactions on Communications*, 53:1885–1894, November 2005.

[110] J. Metzner. On improving utilization in aloha networks. *IEEE Transaction on Communications*, 24(4):447–448, 1976.

[111] P. Michiardi and R. Molva. A game theoretical approach to evaluate cooperation enforcement mechanisms in mobile ad hoc networks. In *Proceedings of the International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Sophia Antipolis, France, March 2003.

[112] G.D. Nguyen, A. Ephremides, and J.E. Wieselthier. Comments on "capture and retransmission control in mobile radio". *IEEE Journal on Selected Area in Communications*, 24:2340–2341, December 2006.

[113] Q.T. Nguyen-Vuong, L. Fiat, and N. Agoulmine. An architecture for umts-wimax interworking. In *Proceedings of the IEEE Symposium on Broadband Convergence Networks*, April 2006.

[114] A. Nyandoro, L. Libman, and M. Hassan. Service differentiation in wireless lans based on capture. In *Proceedings of the IEEE Globecom*, December 2005.

[115] J. Park and M. van der Schaar. Stackelberg contention games in multi-user networks. *EURASIP Journal on Advances in Signal Processing*, 2009.

[116] J. Pérez-Romero, O. Sallent, and R. Agustí. Policy-based initial rat selection algorithms in heterogeneous networks. In *Proceedings of the 7th IFIP International Conference on Mobile and Wireless Communication Networks (MWCN'05)*, Marrakech, Morocco, September 2005.

[117] B. Radunovic and J. Y. Le Boudec. Joint scheduling, power control and routing in symmetric, one-dimensional, multi-hop wireless networks. In *Proceedings of the International conference on Modeling and Optimization in Mobile, Ad-Hoc and Wireless Networks (WiOpt'03)*, Sophia-Antipolis, France, March 2003.

[118] K. Ramadas and R. Jain. Wimax system level evaluation methodology. technical report, wimax forum, January 2007.

[119] V. Ramaiyan, A. Kumar, and E. Altman. Fixed point analysis of single cell ieee 802.11e wlans: uniqueness, multistability and throughput differentiation. In *Proceedings of the ACM SIGMETRICS, Performance Evaluation Rev.*, volume 33, pages 109–120, Barcelona, Spain, April 2005.

[120] R. Rao and A. Ephremides. On the stability of interacting queues in a multiple-access system. *IEEE Transactions on Information Theory*, 34:918–930, September 1988.

[121] S. Ray, D. Starobinski, and J.B. Carruthers. Performance of wireless networks with hidden nodes: A queuing-theoretic analysis. *Computer Communications*, 28:1179–1192, 2005.

[122] M. Realp and A.I. Perez-Neira. Phy-mac dialogue with multipacket reception. In *Proceedings of the ETSI Workshop on Broadband Wireless Ad-hoc Networks and Services*, September 2002.

[123] ITU-T Recommendation. *H.264/ISO/IEC 14496-10(AVC)*. Advanced Video Coding for Generic Audiovisual Services, 2003.

[124] E . M. Reingold, J. Nievergelt, and N. Deo. *Combinatorial algorithms*. Prentice Hall, Englewood Cliffs, New Jersey, 1977.

[125] L. G. Roberts. *Aloha packet system with and without slots and capture. Tech. Rep. Ass Note 8*, Stanford Research Institute, Advance Research Projects Agency, Network Information Center, 1972.

[126] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, 33:153–163, 1965.

[127] E. Sabir and R. El-Azouzi. Achieving constrained nash equilibrium in wireless collision channels. In *Proceedings of ROADEF'08*, Clérmont Ferrand, France, February 2008.

[128] E. Sabir, R. El-Azouzi, and Y. Hayel. A hierarchical slotted aloha game. In *Proceedings of the IEEE/ICST international conference on game theory for networks (GameNets'09)*, pages 222–231, Istanbul, Turkey, May 2009.

[129] E. Sabir, R. El-Azouzi, V. Kavitha, Y. Hayel, and E. H. Bouyakhf. Stochastic learning solution for constrained nash equilibrium throughput in non saturated wireless collision channels. In *Proceedings of the fourth ACM/ICST International Conference on Performance Evaluation Methodologies and Tools (valuetools'09, GameComm)*, pages 1–10, Pisa, Italy, October 2009.

[130] E. Sabir, M. Raiss-El-Fenni, and M. Elkamili. Slotted aloha à première transmission différée : Une nouvelle solution pour supporter les applications sensibles au délai. In *Proceedings of MAJEcSTIC*, Avignon, France, November 2009.

[131] T. Sakurai and H. L. Vu. Mac access delay of ieee 802.11 dcf. *IEEE Transactions on Wireless Communications*, 6:1702–1710, 2007.

[132] J. Sant and V. Sharma. Performance analysis of a slotted-aloha protocol on a capture channel with fading. *Queueing Systems Theory Applications*, 34:1–35, 2000.

[133] J. H. Sarker, M. Hassan, and S. Halme. Power level selection schemes to improve throughput and stability of slotted aloha under heavy load. *Computer Communications*, 25(18):1719–1726, 2002.

[134] P. S. Sastry, V. V. Phansalkar, and M. A. L. Thathachar. Decentralized learning of nash equilibria in multi-person stochastic games with incomplete information. *IEEE Transactions on Systems, Manufacturing and Cybernetics.*, 24:769–777, May 1994.

[135] M. Schwartz. *Information, Transmission, Modulation and Noise*. Third edition, McGraw-Hill, New York, 1980.

[136] N. Shacham. Throughput-delay performance of packet-switching multiple access channel with power capture. *Journal of Performance Evaluation*, 4:153–170, August 1984.

[137] A. Shaikh, R. Sitaraman, A. Akella, and B. Maggs. A measurement-based analysis of multihoming. In *Proceedings of the ACM SIGCOMM*, pages 353–364, Karlsruhe, Germany, 2003.

[138] S. Shakkottai, T. Rappaport, and P. Karlsson. Cross-layer design for wireless networks. *IEEE Communications Magazine*, pages 74–80, October 2003.

[139] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, July and October 1948.

[140] A. Shokrollahi. Raptor codes. *IEEE Transactions on Information Theory*, 52:2551–2567, June 2006.

[141] C. Smith and J. Meyer. *3G Wireless with WiMAX and WiFi, 802.16 and 802.11*. McGraw Hill, 2004.

[142] P. Smith. Bgp multihoming techniques. In *Proceedings of the North American Network Operator's Group (NANOG)*, june 2003.

[143] G. Song and Y. Li. Cross-layer optimization for ofdm wirless networks-part i: theoritical framework. *IEEE transactions on wireless communications*, 4:614–624, March 2005.

[144] Lubo Song and Chansu Yu. Minimizing spatial and time reservation with collision-aware dcf in mobile ad hoc networks. *Ad Hoc Networks*, 7:230–247, 2009.

[145] V. Srinivasan, P. Nuggehalli, C.F. Chiasserini, and R.R. Rao. Cooperation in wireless ad hoc networks. In *Proceedings of the IEEE Infocom*, volume 2, pages 808–817, San Francisco, March 2003.

[146] IEEE standard for local and metropolitan area networks. Part 16: Air interface for fixed broadband wireless access systems, April 2002.

[147] IEEE standard for local and metropolitan area networks. Part 16: Air interface for fixed and mobile broadband wireless access systems, February 2006.

[148] IEEE 802.11 standards. available from http://standards.ieee.org/getieee802/802.11.html.

[149] R. Stemm and R. Katz. Vertical handoffs in wireless overlay networks. *Journal on Mobile Networks and applications*, 3:335–350, 1998.

[150] Zhuo Sun and Wenbo Wang. Investigation of cooperation technologies in heterogeneous wireless networks. *Hindawi Publishing Corporation Journal of Computer Systems, Networks, and Communications, 12 pages*, 2010, February 2010.

[151] W. Szpankowski. Stability conditions for some distributed systems. *Buffered random access systems, Advances in Applied Probability*, 26:498–515, 1994.

[152] Penthera Technologies. White paper on fountain codes for mobile broadcast. penthera technologies, inc. all rights reserved., 2006.

[153] O. Tickoo and B. Sikdar. A queueing model for finite load ieee 802.11 random access mac. In *Proceedings of the IEEE International Conference on Communications (ICC)*, volume 1, pages 175–179, Paris, France, 2004.

[154] C. K. Toh, W. K. Tsai, V. O. L. Li, and G. Guichai. Transporting audio over wireless ad hoc networks: Experiments and new insights. In *Proceedings of the 14th Personal, Indoor and Mobile Radio Communications Symposium (PIMRC)*, volume 1, pages 772–777, China, 2003.

[155] B. Tsybakov and W. Mikhailov. Ergodicity of slotted aloha systems. *Problems of Information Transmission*, 15:73–87, 1979.

[156] A. Urpi, M. A. Bonuccelli, and S. Giordano. Modeling cooperation in mobile ad hoc networks: a formal description of selfishness. In *Proceedings of the International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Sophia Antipolis, France, 2003.

[157] D. Vassis and G. Kormentzas. Performance analysis of ieee 802.11 ad hoc networks in the presence of hidden terminals. *Computer Networks*, 51:2345–2352, 2007.

[158] D. Vassis and G. Kormentzas. Performance analysis of ieee 802.11 ad hoc networks in the presence of exposed terminals. *Ad Hoc Networks*, 6:474–482, 2008.

[159] T. Wan and A. U.H. Sheikh. Analysis of finite population buffered slotted aloha protocols using tagged user analysis (tua). In *Proceedings of the IEEE Globecom*, volume 3, pages 1652–1657, Avignon, France, June 1998.

[160] J.G. Wardrop. Some theoretical aspects of road traffic research. In *Proceedings of the Institution of Civil Engineers*, volume 2, pages 325–378, 1952.

[161] Y. Wu, P. Chou, Q. Zhang, K. Jain, W. Zhu, and S. Kung. Network planning in wireless ad hoc networks: A cross-layer approach. *IEEE Journal on Selected Areas in Communications*, 23:136–150, January 2005.

[162] L. L. Xie and P. R. Kumar. A network information theory for wireless communication: scaling laws and optimal operation. *IEEE Transactions on Information Theory*, 50:748–767, 2004.

[163] M. Xie and M. Haenggi. Towards an end-to-end delay analysis of wireless multihop networks. *Ad Hoc Networks*, 7:849–861, 2009.

[164] Yiping Xing and Rajarathnam Chandramouli. Stochastic learning solution for distributed discrete power control game in wireless data networks. *IEEE/ACM Transactions on Networking*, 14:932–944, 2008.

[165] F. Xu, L. Zhang, and Z. Zhou. Interworking of wimax and 3gpp networks based on ims. *IEEE Communications Magazine*, pages 144–150, 2007.

[166] Y. Yang, J.C. Hou, and L. Kung. Modeling the effect of transmit power and physical carrier sense in multi-hop wireless networks. In *Proceedings of the IEEE Infocom*, Alaska, 2007.

[167] Q. Zhao and L. Tong. A multiqueue service room mac protocol for wireless networks with multipacket reception. *IEEE/ACM Transactions on Networking*, 11:125–137, February 2003.

[168] Q. Zhao and L. Tong. A dynamic queue protocol for multiaccess wireless networks with multipacket reception. *IEEE Transactions on Wireless Commununications*, 3:2221–2231, November 2004.

[169] M. Zorzi. Mobile radio slotted aloha with capture, diversity and retransmission control in the presence of shadowing. *Wireless Networks*, 4:379–388, 1998.

[170] M. Zorzi and R.R. Rao. Comments on "capture and retransmission control in mobile radio". *IEEE Journal on Selected Area in Communications*, 24:2341–2342, December 2006.