Protocoles distribués de contrôle d'accès au médium pour réseaux ad hoc fortement chargés
Distributed MAC Protocols for Heavily Loaded Ad Hoc Networks

Marceau Coupechoux

Thèse présentée à

Directeur de thèse : Christian Bonnet
Jury :
Mario Gerla (rapporteur)
Xavier Lagrange (rapporteur)
Bruno Baynat
Jean-Claude Belfiore
Christian Bonnet
Vinod Kumar

25 juin 2004

Remerciements

Je remercie M. Jean-Claude Belfiore de l'Ecole Nationale Supérieure des Télécommunications de Paris (ENST) d'avoir accepté d'être le Président de mon jury.

Je remercie mes deux rapporteurs MM. Xavier Lagrange de l'Ecole Nationale Supérieure des Télécommunications de Bretagne (ENSTBr) et Mario Gerla de l'Université de Californie à Los Angeles (UCLA).

Je remercie mon directeur de thèse M. Christian Bonnet de l'Institut Eurecom pour avoir accepté d'encadrer mon travail, pour son aide précieuse et en souvenir de notre voyage à Madonna di Campliglio pour une conférence en Italie.

Je remercie M. Vinod Kumar d'Alcatel pour m'avoir embauché chez Alcatel, pour m'avoir transmis une part de ses connaissances étendues dans le domaine de la radio, pour sa relecture méticuleuse du manuscrit, pour son soutien sans faille à l'extérieur comme à l'intérieur de l'entreprise.

Je remercie M. Bruno Baynat de l'Université Pierre et Marie Curie de Paris pour sa participation à mon jury, pour son excellent cours sur la théorie des files d'attentes, pour ses relectures pointilleuses, pour son aide précieuse et ses remarques pertinentes qui ont servi à élaborer le protocole CROMA de manière rigoureuse.

Je remercie dans son ensemble le département radio du centre de recherche d'Alcatel à Marcoussis, et en particulier MM. Thierry Lestable pour ses compétences scientifiques, Luc Brignol pour son esprit critique et sa relecture du manuscrit, ainsi que MM. Jérôme Brouet et Denis Rouffet qui m'ont laissé une grande liberté dans mes recherches quelles qu'aient été les priorités du département.

Résumé

Les communications sans fil ont connu en moins de quinze ans un formidable succès. Dix années ont séparé la première génération de téléphonie mobile, analogique, de la seconde (2G), numérique. Dix ans après, le nombre de terminaux 2G a dépassé dans de nombreux pays celui des postes fixes. Aujourd'hui, la troisième génération (3G) ainsi que les réseaux locaux sans fil (RLAN) nous donnent accès à des contenus multimédia.

Dans ce contexte, quelle pourrait être la quatrième génération (4G) ? Probablement l'intégration de différents systèmes comme les évolutions de la 3G et des RLAN (IEEE 802.11n), les réseaux personnels et corporels sans fil, ou l'accès large bande Worldwide Interoperability for Microwave Access (WiMAX). Il s'agira avant tout de rendre transparente l'utilisation de ces technologies pour l'usager.

L'une de ces technologies pourrait être les réseaux ad hoc sans fil. Ils sont nés dans les années 1970, dans le domaine militaire, pour permettre à des terminaux de communiquer sans infrastructure a priori, par exemple sans les stations de base nécessaires aux réseaux cellulaires 2G et 3G. Le réseau est ainsi créé *pour l'occasion*. Depuis quelques années, la recherche dans ce domaine connaît une grande activité probablement liée au succès de la norme IEEE 802.11 qui permet de réaliser des réseaux ad hoc à moindre coût. L'augmentation des débits de la couche physique laisse également présager des déploiements commerciaux.

En l'absence d'infrastructure, les nœuds d'un réseau ad hoc ont besoin d'un protocole spécifique de contrôle d'accès au médium (MAC) qu'ils partagent. Cette thèse a pour objet l'étude de tels algorithmes et proposent des alternatives à IEEE 802.11 pour des réseaux fortement chargés. La couche MAC apparaît en effet comme un goulot d'étranglement à forte charge. Les difficultés proviennent principalement des variations du canal radio, des changements possibles de topologie dus à la mobilité des nœuds, ainsi que de la nécessité d'imaginer des solutions distribuées. A cela il faut ajouter que les communications entre terminaux peuvent être multi-bonds, c'est-à-dire que l'information peut être relayée, un nœud se comportant alors comme routeur (cf. figure 1).

Dans le premier chapitre, nous dégageons de la littérature une classification des protocoles MAC en distinguant deux familles de solutions : les protocoles basés sur la compétition pour le canal et les protocoles sans conflit.

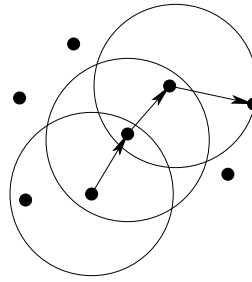La norme IEEE 802.11 est en grande partie l'héritière de la première famille.

Figure 1: Exemple de communication multi-bond dans un réseau ad hoc sans fil.

Elle est adaptée à la fois aux réseaux avec point d'accès (PA) et aux réseaux ad hoc. Dans le premier cas, le PA est un pont vers un réseau cœur et est nœud central pour le trafic. L'objet du chapitre 2 est l'évaluation des performances de IEEE 802.11 dans un réseau avec PA. C'est en effet le cas qui intéresse le plus le monde industriel. Nous montrons en particulier que les concepts du réseau ad hoc et multi-bond peuvent être utilisés pour étendre la couverture d'une cellule.

Nous montrons également que le standard offre de bonnes performances en simple-bond, mais qu'il présente des faiblesses dans un scénario complètement décentralisé, sans point d'accès, et de surcroît multi-bond, et ce, en terme de capacité et d'équité dans le partage de la ressource radio entre les nœuds. Les chapitres suivants proposent des alternatives à la norme IEEE.

Le chapitre 3 propose un nouveau protocole de la seconde famille, basé sur l'accès multiple par répartition dans le temps (AMRT) et appelé Protocole MAC sans Conflit Orienté Récepteur (CROMA). CROMA est une alternative à IEEE 802.11 principalement destiné aux réseaux fortement chargés avec une exigence d'équité. CROMA a été conçu et il est étudié dans ce chapitre pour des réseaux ad hoc sans fil.

Le chapitre 4 étudie des mécanismes additionnels pour l'augmentation de la capacité de ces réseaux. Trois concepts sont présentés dont le point commun est d'utiliser les interactions entre couches protocolaires pour atteindre de meilleures performances. Ils sont également le point de départ de futures versions de CROMA.

Chapitre 1 - Protocoles distribués de contrôle d'accès au médium pour réseaux sans fil

La littérature ne fournit pas de définition unifiée et claire des réseaux ad hoc. Ce chapitre propose donc un ensemble de définitions fondé sur les caractéristiques les plus souvent rencontrées.

Parmi les applications de cette technologie, nous trouvons historiquement les communications militaires sur champ de bataille. Aujourd'hui, les communications entre véhicules, entre équipements domestiques, les réseaux de capteurs et maillés, ou encore l'extension de couverture des réseaux cellulaires sont aussi envisagés.

Les domaines de recherche sont variés car la structure des réseaux ad hoc influe sur presque toutes les couches protocolaires. Parmi les champs les plus actifs, notons le routage, la théorie de l'information pour l'évaluation des capacités théoriques, et le contrôle d'accès au médium qui est l'objet de cette thèse.

La couche MAC, dont le but est de définir les règles de partage des ressources radio, est confrontée à quatre problèmes principaux : (i) Les antennes ne peuvent fonctionner qu'en mode semi-duplex, c'est-à-dire qu'elles ne sont pas capables de transmettre et recevoir au même moment. (ii) Les transmissions sont soumises à un canal radio fluctuant et éventuellement à une topologie de réseau variable. (iii) Les mécanismes d'écoute de porteuse sont fragilisés par le problème du terminal caché. Sur la figure 2, le nœud C n'est pas en mesure d'entendre la transmission de A vers B. (iv) La capacité peut être réduite à cause du problème du terminal exposé. Sur la figure 3, le nœud C, détectant la transmission de B vers A, peut croire que sa transmission vers D est impossible alors qu'elle peut avoir lieu car A est hors de portée.
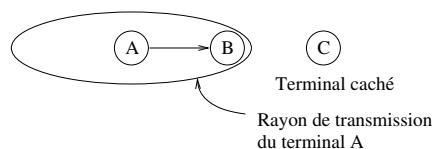

Figure 2: Problème du terminal caché.

Sous ces contraintes, deux familles de solutions MAC apparaissent dans la littérature : (i) Les protocoles basés sur la compétition. Pour chaque transmission de paquet de données, les nœuds entrent en compétition pour accéder au canal. (ii) Les protocoles sans conflit, issus de l'accès multiple à répartition dans le temps, en fréquence
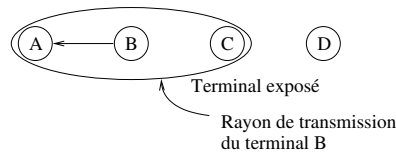
Figure 3: Problème du terminal exposé.

ou en code. Après une éventuelle phase de réservation, la transmission des données successives est sans collision.

IEEE 802.11 DCF (Distributed Coordination Function) rassemble de nombreux algorithmes utilisés par les protocoles de la première classe. La figure 4 montre un exemple de transmission de paquet de données. Après une phase d'écoute du canal, source et destination échangent de petits paquets de contrôle RTS/CTS afin de réserver le canal, de se prémunir du problème du terminal caché et de réduire le temps perdu en cas de collision. Les données sont ensuite transmises par la source et la destination en accuse réception. RTS et CTS contiennent dans leur en-tête l'indication de la durée de cet échange, de telle sorte que les autres nœuds se retiennent de transmettre pendant ce dialogue (grâce à la fonction NAV). Pour diminuer les chances de collision entre paquets RTS, les stations ne tentent à nouveau leur chance qu'à partir d'un temps aléatoire, c'est l'algorithme de retour aléatoire.
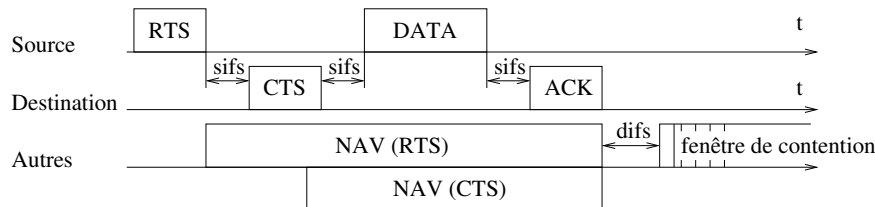

Figure 4: Accès IEEE 802.11 DCF avec RTS/CTS et fonction NAV.

Plusieurs articles contestent l'efficacité de IEEE 802.11 dans un environnement multi-bond. Il lui est reproché de ne pas complètement résoudre le problème du terminal caché, de ne pas tenir compte du problème du terminal exposé et de ne pas assurer un partage équitable du canal radio entre les flux de données. Cela est particulièrement vrai à forte charge.

C'est pourquoi, afin de traiter ces dysfonctionnements, il est possible de se tourner vers la seconde famille de protocoles MAC. Par exemple, les solutions AMRT sont réputées offrir de plus grandes capacités, malgré des temps d'accès parfois plus importants. Ce chapitre classe les contributions de la littérature dans ce do-

maine en trois sous-ensembles : (i) Les articles en relation avec la complexité des algorithmes AMRT. En fait, la plupart des problèmes d'ordonnancement, c'est-à-dire d'attribution optimale des intervalles de temps, sont NP-complets. Aucune recherche de solution en un temps polynomial n'est donc connue. (ii) Les algorithmes d'allocation de nœud. Un intervalle de temps élémentaire est attribué à un nœud. (iii) Les algorithmes d'allocation de lien. Les intervalles sont associés à des liens entre nœuds (cf. figure 5). Il est encore possible de distinguer les algorithmes centralisés, les protocoles à étalement dans le temps, à réservation déterministe, à accès aléatoire. Seuls ces derniers semblent en mesure de recouvrir un large éventail de scénarios.
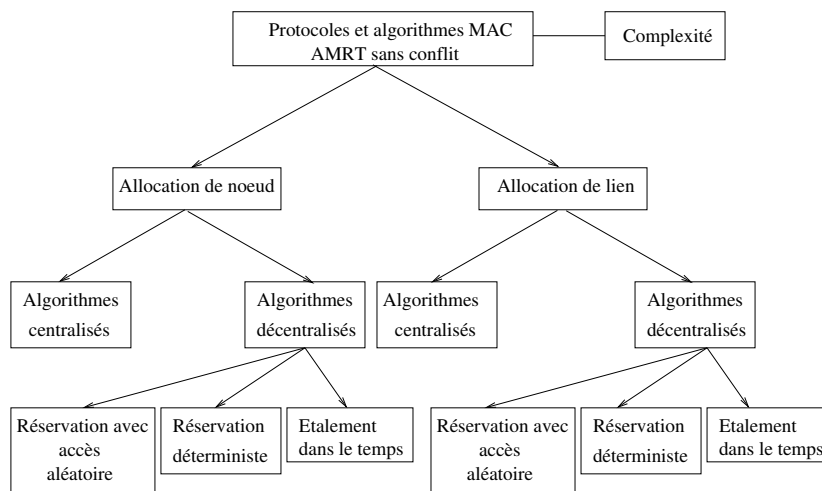

Figure 5: Une classification des protocoles et algorithmes MAC AMRT.

Chapitre 2 - Évaluation des performances de IEEE 802.11 DCF

La disponibilité à faible coût de cartes à la norme IEEE 802.11b est l'une des raisons du regain d'intérêt de la recherche pour les réseaux ad hoc. La norme est donc susceptible d'être utilisé à la fois dans un réseau avec PA et dans un réseau ad hoc voir multi-bond. Pour un manufacturier d'équipements de télécommunications, il est difficile d'imaginer un réseau qui ne soit pas relié à un cœur capable de fournir des services. C'est pourquoi, dans ce chapitre, nous privilégions les configurations de la norme IEEE 802.11b avec PA. D'abord dans un réseau simple-bond (PA-terminal), puis multi-bond (PA-terminal-terminal ou PA-relais-terminal).

Résultats de capacité pour le trafic TCP

IEEE 802.11 a d'abord été conçu pour le transport de trafic de type TCP, c'est-à-dire essentiellement la navigation web (WWW) et le transfert de fichiers. L'étude des débits maximum accessibles montre que le standard souffre d'un sur-débit protocolaire important. A titre d'exemple, le débit maximal au-dessus de TCP avec le mécanisme de RTS/CTS atteint environ $3,3$ Mbps alors que la couche physique offre 11 Mbps.

Les capacités observées restent néanmoins importantes. Le téléchargement d'un fichier de 10 Mo par dix utilisateurs simultanés situés à dix mètres du PA s'effectue en moins de 400 s. A cette même distance, plus de cinquante usagers peuvent naviguer sur le web avec un débit de 450 Kbps.

Voix sur IP

Les opérateurs de télécommunications et les entreprises sont de plus en plus intéressés par la possibilité de transporter de la voix sur les RLAN. Pour les uns, les RLAN pourraient venir en complément des réseaux cellulaires. Pour les autres, il s'agirait de faire converger réseaux de voix et de données.

Dans cette partie, nous évaluons la capacité de IEEE 802.11b en terme de nombre simultané d'appels voix pour différents codecs. Contrairement aux autres contributions sur le sujet, nous fondons notre analyse sur le critère de qualité du modèle E, proposé par l'UIT. Les résultats de simulation, qui prennent en compte différentes stratégies d'absorption de la gigue de délai, montrent que la capacité est de sept appels avec G711, douze avec GSM-EFR et dix-huit avec G723.1.

Effet d'éblouissement

Les RLAN avec adaptation de liens connaissent un effet d'éblouissement d'un genre particulier : la présence d'un usager dans une zone de faible débit physique, c'est-à-dire éloignée du PA, fait chuter le débit d'un utilisateur proche du PA, et par conséquent le débit agrégé de la cellule. Cet effet est illustré tant pour le lien montant que pour le lien descendant, avec un trafic TCP ou UDP.

Cette partie propose quatre solutions pour atténuer l'influence d'un utilisateur éloigné. La première, pour le lien descendant, est de diminuer la bande passante

attribuée à cet utilisateur par une politique d'ordonnancement des paquets transmis par le PA. La seconde est d'utiliser un nœud relais capable de substituer deux liens haut débit à un lien faible débit. La troisième, pour un trafic TCP, est d'adapter la borne maximale de la fenêtre de transmission au débit physique observé. Enfin, faire varier la taille des paquets sur le lien montant atténue l'effet d'éblouissement.

Internet rapide pour zones peu denses

Cette partie donne un exemple pratique et original de déploiement de RLAN. Alors que le haut débit filaire connaît un grand succès grâce à l'Asymmetric Digital Subscriber Line (ADSL) et que l'opérateur historique abandonne ses obligations de service public, de grandes parties du territoire restent à l'écart de l'accès rapide à Internet. Ceci est essentiellement dû aux faibles densités de population qui ne permettent pas de rentabiliser l'investissement d'un centre de concentration (DSLAM).

Ici, une alternative qui couple un accès satellite à un réseau RLAN est étudiée. En prenant en compte un scénario fourni par le manufacturier, nous montrons qu'une telle solution peut servir jusqu'à soixante abonnés avec une qualité de service acceptable.

Extension de couverture

Après avoir étudié les RLAN au sein d'une topologie simple-bond, nous étudions la possibilité d'étendre la couverture d'un PA grâce aux multi-bonds : un terminal hors de portée du PA peut néanmoins s'y rattacher s'il trouve d'autres nœuds capables de relayer ses données jusqu'à lui.

L'étude évalue d'abord la probabilité de connexion d'une station en fonction de sa distance au PA ainsi que le nombre moyen de bonds nécessaires pour atteindre le réseau filaire. Les résultats sont très dépendants de la densité des nœuds et suggère l'utilisation de relais fixes appartenant à l'infrastructure de l'opérateur.

L'extension de couverture souffre néanmoins de la dégradation du débit en fonction du nombre de bonds (cf. figure 6). Une des pistes à explorer pour gagner en capacité est la réutilisation spatiale des ressources : deux transmissions peuvent être simultanées si elles sont suffisamment éloignées pour ne pas s'interférer.

Enfin, cette partie propose un algorithme permettant d'obtenir des cartes de couverture dans le cas d'un déploiement avec extension de couverture. Cet algorithme
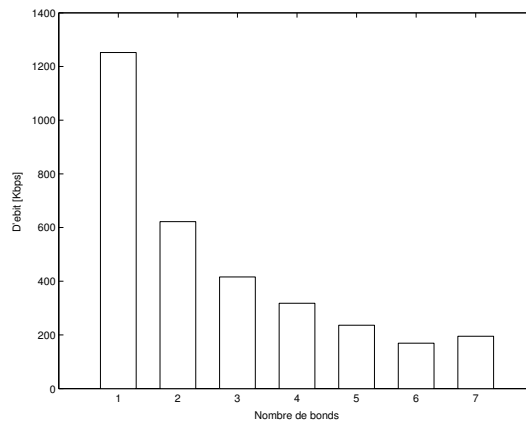
Figure 6: Débit en fonction du nombre de bonds (débit physique 2 Mbps, RTS/CTS, nœuds en ligne).

exploite la possibilité de réutilisation spatiale. Nous montrons également l'exemple d'une couverture de bâtiment et comparons les solutions avec ou sans nœuds relais.

En conclusion, ce chapitre a mis en valeur les bonnes performances de IEEE 802.11 dans un réseau simple-bond, malgré un sur-débit protocolaire important. L'utilisation des concepts des réseaux ad hoc et multi-bonds s'est avérée utile pour atténuer l'effet d'éblouissement et étendre la couverture d'un PA. Cette extension est obtenue au prix d'une dégradation du débit en multi-bonds et de problèmes d'équité.

Chapitre 3 - CROMA : description du protocole et évaluation de ses performances

Dans ce chapitre, nous étudions une alternative à IEEE 802.11 pour les réseaux ad hoc multi-bonds fortement chargés. Pour cela, nous nous tournons vers la famille des protocoles sans conflit basés sur de l'AMRT et proposons un nouveau protocole, appelé CROMA. Il repose sur deux principes : la synchronisation des nœuds et le rôle central donné aux récepteurs d'une transmission. La première idée nous laisse espérer une meilleure utilisation du canal radio. La seconde tire sa justification du fait que la zone de réception est le lieu le plus approprié de compétition pour le canal.

Description du protocole

Dans CROMA, le temps est divisé en trames, elle-mêmes subdivisées en intervalles de temps (figure 7). Chacun d'entre eux se compose de temps de requête (REQ), d'invitation (RTR) et de transmission de données (DATA). Un intervalle de temps peut être soit dans un état *libre* si aucune communication n'y a été établie, soit dans un état *occupé* s'il est associé à un récepteur. Dans ce cas, plusieurs communications peuvent être en cours avec ce récepteur qui reçoit alors des paquets de plusieurs émetteurs.
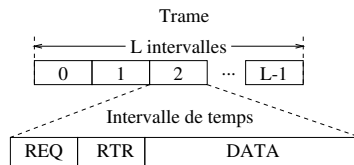

Figure 7: La structure de trame de CROMA.

La phase d'accès pour un émetteur commence par une écoute de la trame pendant laquelle l'état des intervalles de temps (libre ou occupé) est enregistré. Il choisit ensuite un intervalle sur lequel sera envoyée sa requête. En cas de collision avec d'autres requêtes, l'algorithme de retour aléatoire est appliqué. Sinon, le récepteur répond grâce à un paquet RTR.

En cas de réservation réussie, une communication est établie entre les deux nœuds sur cet intervalle de temps. Pour transmettre, l'émetteur doit attendre une invitation à émettre incluse dans les RTR (poll). La figure 8 montre un exemple de phase de transmission impliquant un récepteur et trois émetteurs. Notons que les paquets de données sont acquittés par les RTR grâce à un numéro de séquence (sn).

A la fin d'un message, lorsque l'émetteur n'a plus de paquets à envoyer, la communication est coupée. Le récepteur conserve la main sur l'intervalle de temps si d'autres communications sont en cours. Dans le cas contraire, l'intervalle est à nouveau libre.

Un mécanisme d'équité oblige les récepteurs à libérer leur intervalle de temps lorsque la trame est trop longtemps pleine, c'est-à-dire qu'elle n'accepte plus de nouvelles communications avec de nouveaux récepteurs. A contrario, une communication a la possibilité d'être partagée entre plusieurs intervalles de temps si la trame est sous-utilisée (la transmission est alors dite *multi-intervalle*)
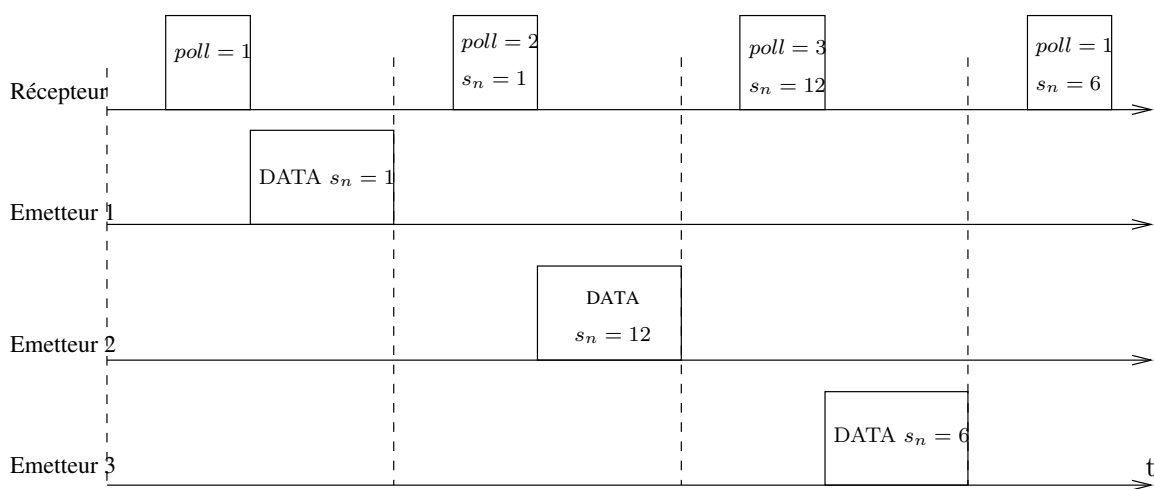
Figure 8: Invitations (poll) et accusés de réception (sn) durant la phase de transmission de CROMA.

CROMA assure qu'il n'y a pas de collision entre paquets de données. Il résout en outre les problèmes de terminaux cachés et exposés.

Analyse du protocole

Une étude de l'utilisation du canal basée sur les chaînes de Markov est présentée dans cette partie, d'abord dans le cas simple d'un seul intervalle de temps par trame puis dans le cas général. Il en résulte que dans un réseau complètement connecté, CROMA approche une utilisation des intervalles de temps de 100% lorsque la longueur moyenne des messages augmente.

Performances dans des réseaux multi-bonds

Cette partie évalue par simulations les performances de CROMA dans un réseau multi-bond classique souvent rencontré dans la littérature, la topologie *en carrés*, puis dans un réseau aléatoire, enfin dans un réseau mobile.

De ces simulations, nous pouvons tirer les conclusions suivantes : (i) CROMA permet d'atteindre des débits bien plus importants que IEEE 802.11 à forte charge. A faible charge, les délais sont plus faibles avec IEEE 802.11 qu'avec CROMA. (ii) Les communications multi-intervalles rendent le protocole presque indépendant de la longueur de trame choisie a priori (cf. figure 9 pour la topologie en carrés et différentes longueurs de trames). (iii) Une grande équité locale, c'est-à-dire ne

tenant pas compte des conditions de bout en bout, est assurée par CROMA. (iv) La mobilité dégrade les performances de CROMA, qui reste néanmoins plus robuste que IEEE 802.11.
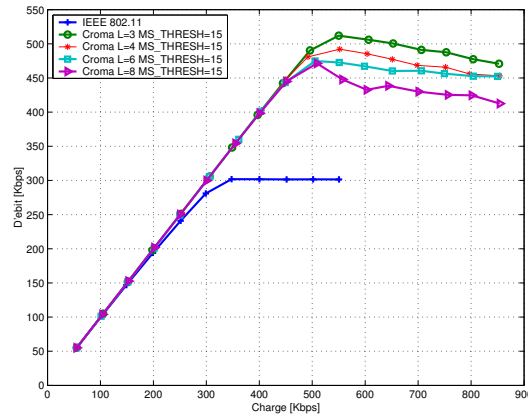

Figure 9: Débit en fonction de la charge, topologie en carrés.

Chapitre 4 - Mécanismes additionnels pour l'augmentation de la capacité

Il n'est pas possible d'accroître encore la capacité des réseaux ad hoc sans fil sans se heurter aux limites fixées par la théorie de l'information. Gupta et Kumar [104] en ont donné les bornes pour les réseaux fixes. Le but de ce chapitre est d'examiner les hypothèses du modèle proposé par [104] et d'essayer d'en relâcher certaines afin d'obtenir des gains de capacité. Cette démarche est illustrée à travers trois techniques qui utilisent les interactions entre couches protocolaires. Il s'agit enfin d'ouvrir de nouvelles pistes de recherche pour l'amélioration des performances de CROMA, et en particulier de sa phase de réservation fondée sur le protocole ALOHA.

La capacité des réseaux ad hoc sans fil

Dans [104], Gupta et Kumar ont abouti à des conclusions pessimistes concernant la capacité des réseaux ad hoc. Même si la topologie, le trafic et les rayons de transmission sont choisis de manière optimale, cette capacité décroît en $1/\sqrt{n}$ bits-mètres/s

par paire source-destination, où n est la densité des nœuds. Si maintenant, nous considérons des réseaux aléatoires, le débit agrégé décroît comme $1/\sqrt{n \log n}$ bits/s.

De leur étude nous pouvons déduire que : (i) Les réseaux ad hoc ne passent pas à l'échelle. (ii) Leur capacité est limitée d'une part par les interférences, d'autre part par la quantité de trafic relayé.

La mobilité comme source de diversité

Tse et Grossglauser [102] ont écarté l'une des hypothèses de [104] qui ne considère que des réseaux fixes. Ils ont montré que si les nœuds sont mobiles et que les paquets sont contraints à ne faire que deux bonds, l'un entre la source et un relais aléatoire, l'autre entre ce relais et la destination, le débit par paire source-destination est maintenu constant quand n augmente.

La contribution de cette partie est de simuler la politique d'ordonnancement proposée par [102] grâce à une version simplifiée de CROMA. Les résultats obtenus montrent le gain apporté par la mobilité des nœuds. Ils mettent également en évidence l'importance du modèle de trafic choisi. Enfin, le rayon de transmission optimal est tiré d'une analyse théorique.

Amélioration du protocole ALOHA synchronisé grâce à la diversité multi-utilisateur

Gupta et Kumar ont considéré que la capacité du canal physique était fixe et constante. Or la littérature a montré qu'elle pouvait être accrue dans un système multi-utilisateur en tirant partie des variations du rapport signal à bruit (RSB).

Dans cette partie, nous considérons n terminaux désirant transmettre des données à une station centrale en utilisant une version modifiée du protocole ALOHA synchronisé. De la même manière qu'avec le protocole traditionnel, les sources émettent sur un intervalle de temps avec une probabilité p. Ici, en revanche, la transmission est assujettie à de bonnes conditions de canal radio, c'est-à-dire qu'elle n'est possible que lorsque le RSB dépasse un certain seuil (RSB_0). Si tous les terminaux ont la même distribution et la même moyenne de RSB, p est directement reliée à RSB_0 par l'inverse de la fonction de répartition du RSB.

Dans cette partie, nous donnons les formules de capacité d'un tel système lorsque

l'émetteur choisit de manière optimale sa puissance et son débit. Ces résultats théoriques sont obtenus pour les canaux de Rayleigh, avec ou sans techniques de diversité, et pour les canaux de Nakagami-m. Les applications numériques montrent que la probabilité de transmission optimale diffère de $1/n$. Elles confirment également la caractéristique principale de la diversité multi-utilisateur : la capacité augmente avec le nombre de terminaux (cf. figure 10).
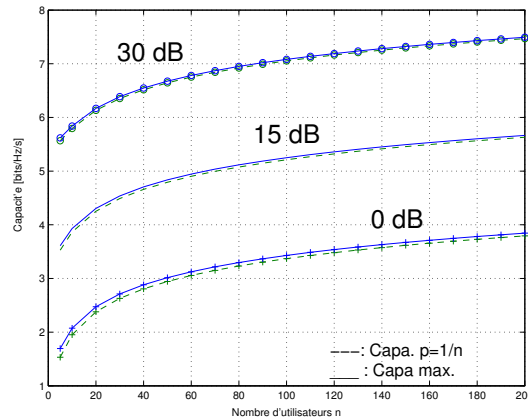

Figure 10: Capacité maximale et à $p = 1/n$ du protocole ALOHA synchronisé en fonction du RSB moyen.

ALOHA synchronisé avec réception multi-paquet

Dans cette partie, nous adoptons un modèle d'interférence différent de celui proposé dans [104]. En effet, nous considérons la possibilité pour un nœud de décoder plusieurs paquets reçus simultanément et nous montrons que les performances du protocole ALOHA dans un réseau multi-bond peuvent être ainsi améliorées.

Une analyse théorique donne les formules de débit local et de bout en bout pour un réseau multi-bond utilisant le protocole ALOHA synchronisé avec réception multi-paquet. Ces débits sont obtenus en fonctions des probabilités élémentaires $r_{n,k}$ pour un récepteur de décoder k paquets sachant que n ont été reçus.

Ensuite, trois modèles de réceptions sont considérés pour des transmissions avec spectre étalé : un modèle simple souvent utilisé dans la littérature, un banc de filtres adaptés et un détecteur multi-utilisateur. Pour chacun de ces modèles, les valeurs $r_{n,k}$ sont obtenues soit par calcul, soit par simulation. Les résultats numériques montrent les gains obtenus grâce à la réception multi-paquet. Ils confirment aussi

la supériorité de la détection multi-utilisateur sur la méthode de filtres adaptés (figure 11).
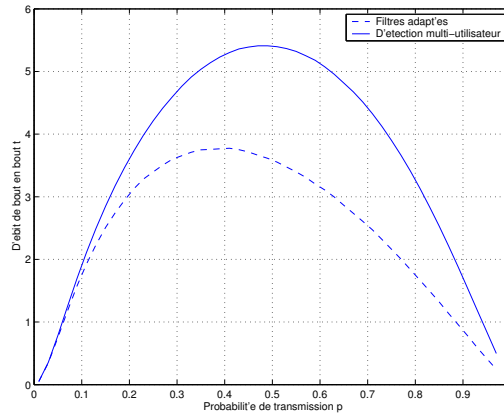

Figure 11: Débit de bout en bout du protocole ALOHA synchronisé avec réception multi-paquet.

Plusieurs conclusions peuvent être tirées de ce chapitre : (i) Contrairement à une idée reçue, la mobilité n'est pas forcément une source de dégradation des performances et au-delà du cas théorique présenté ici, elle devrait être exploitée de manière plus méthodique. (ii) Les interactions entre couches protocolaires sont un domaine actif de recherche qui peut permettre de dépasser certaines limites de capacité. (iii) Plusieurs résultats théoriques ont été obtenus dans ce chapitre : le rayon de transmission optimal pour l'algorithme utilisant la mobilité, la capacité du protocole ALOHA synchronisé avec prise en compte des conditions de canal radio et le débit de bout en bout d'un réseau multi-bond avec protocole ALOHA synchronisé et réception multi-paquet.

Conclusion

Le but de cette thèse était d'étudier les mécanismes de contrôle d'accès au canal pour les réseaux ad hoc sans fil. Nous énumérons à nouveau succinctement nos contributions :

(i) Nous avons proposé une classification des protocoles MAC distribués pour réseaux sans fil. Un accent a été particulièrement mis sur les mécanismes sans conflit AMRT.

(ii) Nous avons étudié le protocole IEEE 802.11 DCF et sa capacité avec plusieurs

types de trafic : TCP (navigation WWW et transfert de fichier) et UDP (voix sur IP). La notion d'éblouissement a été illustrée et plusieurs solutions ont été proposées pour en atténuer les effets. Un exemple de dimensionnement de réseau a été montré avec la notion d'Internet rapide pour zones peu denses. La possibilité d'une extension de couverture d'un PA grâce à des communications multi-bonds a été étudiée. Cette extension n'est possible qu'aux dépens du débit et de l'équité.

(iii) Nous avons proposé un nouveau protocole ad hoc sans conflit appelé CROMA, alternative à IEEE 802.11 dans les réseaux multi-bonds fortement chargés. Ce protocole offre de plus grands débits et une plus grande équité à forte charge.

(iv) Nous avons étudié plusieurs mécanismes d'interactions inter-protocolaires afin d'augmenter la capacité des réseaux ad hoc sans fil. Le premier utilise la mobilité pour réduire la quantité de trafic relayé. Le deuxième met en oeuvre la notion de diversité multi-utilisateur. Le troisième utilise la réception multi-utilisateur. Plusieurs résultats théoriques et numériques ont été obtenus.

Ces contributions ne représentent qu'un élément du vaste domaine des réseaux ad hoc. Ils ouvrent la voie à de nouvelles recherches.

Par exemple, l'étude des performances de CROMA nécessite maintenant des modèles de canal radio plus complexes afin d'en évaluer l'impact sur le protocole, les mécanismes de retransmission et d'accusé de réception, et la réutilisation spatiale.

L'introduction et l'étude de mécanismes de qualité de service semblent nécessaires pour une version ultérieure de CROMA.

Les études d'interactions inter-protocolaires ne sont que balbutiantes dans la littérature. Il serait intéressant d'intégrer les mécanismes présentés dans le chapitre 4 dans CROMA afin d'en améliorer les performances.

Dans l'étude de IEEE 802.11, cette thèse s'est essentiellement attachée à en évaluer les performances par simulation. Un autre champ de recherche concerne les analyses théoriques du protocole. Dans un réseau complètement connecté, elles sont déjà disponibles dans la littérature. Il serait intéressant d'étendre ces analyses au débit de bout en bout d'un réseau multi-bond. Pour CROMA, cette analyse est également souhaitable. Enfin, l'étude de CROMA dans un réseau avec PA se révèle nécessaire pour un manufacturier d'équipements de télécommunication.

Abstract

This thesis contributes to the domain of medium access control for wireless ad hoc networks. These networks are by definition created *for the occasion*, or *for a specific purpose*, and usually have to operate without any existing fixed infrastructure.

In the first chapter of this thesis, a synthesis of the MAC protocols and algorithms is proposed. The contention-based and the conflict-free protocols are two well known families.

IEEE 802.11 DCF is the heir of the first family. It is able to address both access point centric and ad hoc deployments, as well as single and multi-hop communications. The foremost objective of this dissertation is to find alternative schemes for improved MAC performance in highly loaded networks.

Chapter 2 studies the capacity of IEEE 802.11b in case of AP centric networks with TCP and UDP traffic. A specific kind of near-far effect has been detected. The consequent performance degradation is highlighted and some solutions are proposed. A real world deployment is presented for outdoor provisioning of high speed Internet to low density areas. Advantage of using the multi-hop concept to extend the coverage range of an access point and the corresponding issues of degradation in throughput and fairness at high input loads have been analyzed in detail.

In chapter 3, we propose a new slotted protocol, called CROMA, to overcome the weaknesses of IEEE 802.11 in highly loaded multi-hop ad hoc networks. We provide an analytical study of CROMA in a fully connected network. Extensive simulations considering a challenging network topology, a random network, and a mobile network show that CROMA clearly outperforms IEEE 802.11 in the targeted environments in terms of channel utilization and fairness.

Chapter 4 explores three examples of cross-layer mechanisms estimated to be useful for further improvement in capacity of ad hoc networks and CROMA evolutions. Impressive performance improvements have been demonstrated through theoretical analysis and simulations in the following three cases: (i) A simplified version of CROMA and an associated scheduling policy can take advantage of node mobility for throughput improvement. (ii) Multi-user diversity considerably improves the CROMA reservation scheme based on slotted ALOHA. (iii) Multi-user detection can offer additional improvement for reservation too.

Contents

Introduction

During the last fifteen years, personal wireless communications have seen a spectacular growth. Although radio technologies are more a century old, only ten years were needed to switch from the analog first generation of mobile telephony to the digital second generation (2G) in the 1990's. Ten years after the introduction of the Global System for Mobile communications (GSM), the number of mobile subscribers exceeds that of fixed phones even in many developing countries. Thus, even if a lot of effort has still to be spent to reduce the digital divide and to build an information society, wireless communications are unquestionably a great technological success.

Today, the 2.5G, the third generation (3G), that is about to be launched, and Wireless Local Area Networks (WLAN) provide the basis for multimedia contents. With these technologies, it may be conjectured that mobile terminals will become the principal means for Internet access.

In this context, what could be the fourth generation (4G) of mobile networks? Probably not a unique system able fulfilling all the requirements. Rather a kind of tree whose branches could be the evolutions of the 3G, next generations WLAN (e.g. IEEE 802.11n), Personal Area Networks, Body Area Networks, broadband wireless access WiMAX, or cellular Internet Protocol (IP) networks. The leaves could be the latest technological advances in physical and link layers mechanisms, e.g. multi-carrier modulations, spread spectrum, multiple antennas, beam-forming, or multi-user detection. One of the first tasks of the fourth generation will probably be to seamlessly integrate this diversity for maximizing advantages for the subscriber.

Ad hoc networking could be one of the leaves of 4G. Wireless ad hoc networks are by definition created *for the occasion*, or *for a specific purpose*, and have to operate without any existing fixed infrastructure. Initially designed for battlefield communications during the early 1970's, ad hoc networking has been of great interest to the research community during the past few years for applications in daily life. This is motivated by the advent of mobile computing, the interest in peer-to-peer

networks for information exchange, the widespread availability of IEEE 802.11 products, and the creation of the Mobile Ad hoc Networks (MANETs) working group at the Internet Engineering Task Force (IETF). Moreover, high data rates and low cost portable radios can be extremely useful in a lot of civilian applications: emergency and rescue operation networks, inter-vehicle communication, sensor networks, home equipments inter-connections, meshed networks, and coverage extension of future cellular networks.

Historically, research revolved around the two main problems of medium access control (MAC) and routing. This dissertation focuses on the former aspect. MAC design is indeed a critical point in ad hoc networks for a fair and efficient allocation of the shared radio medium. The technological challenges come mainly from the wireless channel and from the varying network topology. Additional problems are encountered due to the distributed nature of the network: All nodes play the same role and no one is supposed to organize the resource sharing as in cellular networks. Today, the straightforward but sub-optimal solution to address these issues is to use the existing MAC layer of IEEE 802.11 that presents major weaknesses - especially at high input loads. The foremost objective of this dissertation is to find alternative schemes for improved MAC performance in highly loaded networks.

In the first chapter of this thesis, the main characteristics of ad hoc networking are described. Literature search proves that capacity is presently the key research issue. The MAC layer appears to be one of the bottlenecks for better performance. In this chapter, a synthesis of the MAC protocols and algorithms is proposed. The contention-based and the conflict-free protocols are two well known families.

In the first category, IEEE 802.11 is the most famous one and also the MAC protocol used in most ad hoc networks implementations. Since it is a standard, it is also the first candidate for a commercial deployment. Almost all available WLAN cards implement the Distributed Coordination Function (DCF). DCF is able to address both networks with access point (AP) and ad hoc deployments for single and multi-hop communications.

Since inter-working between wireless access networks and the core network is a must in some application scenarios, chapter 2 deals with some associated problems. In particular, we underline the advantage of using ad hoc and multi-hop concepts in multiple types of deployment of AP centric networks. Firstly a clearly identified advantage of capacity improvement by using relay nodes to solve the problem usually encountered near-far effect is explained. Secondly, multi-hop communications are

shown to help coverage extension of an AP, e.g. to cover dead spots frequently encountered in indoor environments.

We also show that IEEE 802.11 DCF exhibits good performance in single-hop networks for Transport and Control Protocol (TCP) based traffic, and suffers from a high overhead for Voice over IP (VoIP). In multi-hop networks, IEEE 802.11 faces a degradation of the throughput with increasing number of hops and presents weaknesses for a fair allocation of resources.

To offer an interesting alternative to the existing solution, chapter 3 proposes a new protocol, called Collision-free Receiver-Oriented MAC (CROMA) that tries to take advantage of the advances in random schemes in a slotted environment. In CROMA, receiver nodes act as local and temporary base stations. Extensive simulations considering a challenging network topology, a random network, and a mobile network show that CROMA clearly outperforms IEEE 802.11 at high input loads in terms of channel utilization and fairness.

In order to further improve the capacity of ad hoc networks at the MAC layer and to prepare the evolutions of CROMA, chapter 4 explores the emerging research area of cross-layer interactions. A landmark paper attempts to show that fixed ad hoc networks are not scalable. We successively question the assumptions of this paper by considering mobile networks and by introducing new interference models and adequate receiver algorithms.

First of all, we exploit node mobility thanks to a simplified version of CROMA and a scheduling policy in order to improve the network throughput. This leads to the counter intuitive result that mobility may be beneficial to capacity. Then, we consider multi-user diversity with the study of the channel aware slotted ALOHA protocol: If users take into account their channel conditions before transmitting, the overall performance of the system is increased. Finally, we provide theoretical results for the throughput of the slotted ALOHA with multi-packet reception. We show the capacity improvement advantages of using multi-user detection.

	IEEE 802.11 DCF	CROMA and evolutions
Access point networks	Chapters 2 & 4	for further work
Ad hoc networks	Chapter 3	Chapters 3 & 4
Single-hop networks	Chapter 2	Chapters 3 & 4
Multi-hop networks	Chapter 2	Chapters 3 & 4

3

Chapter 1

Distributed MAC Protocols for Wireless Networks

1.1 Introduction

Ad hoc networking is a attractive field that challenges the imagination of researchers. It indeed mixes all the difficulties related to the study of varying radio channel, varying network topology, node mobility, the scarcity of resources (bandwidth, energy), and a distributed processing environment.

During the recent past, wireless ad hoc networking has become a major research area. Since the first papers published by e.g. F. Tobagi and L. Kleinrock in the early 1970's, very little has been reported in the literature. The present surge of interest is motivated by the advent of mobile computing, the interest in peer-to-peer networks for information exchange, the widespread availability of IEEE 802.11 products, and the creation of the IETF MANET working group.

No clear widely accepted definition of ad hoc networks is available. Section 1.2 presents some definitions and some applications of ad hoc networking. A list of selected areas of active research is also provided. Then, the focus is maintained on Medium Access Control protocols - basic underlying mechanisms and associated algorithms for performance enhancement. Section 1.3 provides a feature based classification of the MAC protocols.

Brief description of contention oriented schemes is given in section 1.4. With these solutions, nodes compete for the channel at each packet transmission. It is shown that IEEE 802.11 DCF is to a large extent the heir of this category of

protocols.

However, better performance and better channel utilization is expected with the conflict-free protocols, especially at high input loads. With these protocols, the channel is effectively reserved for a certain amount of time and then transmission is conflict-free. Conflict-free Time Division Multiple Access (TDMA) based schemes are summarized in section 1.5. It is shown that only adaptive slot allocations can address the issues of mobility, varying topology and traffic patterns.

1.2   Wireless Ad Hoc Networks

1.2.1   Definitions

Giving a precise definition of wireless ad hoc networks is not an easy task, and the literature on the subject doesn't provide an unified view of the concept. *ad hoc* essentially means *for the occasion*. This implies that these wireless networks are created for a specific purpose and disappear with the condition of their creation.
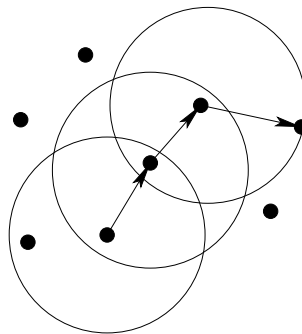
Figure 1.1: Example of multi-hop forwarding in a wireless ad hoc network.

From this common and generic definition, we can deduce the main characteristics of an ad hoc network [107, 136]:

- Because the network is created *on the fly*, it should not rely on a preexisting fixed and/or wired topology. In that sense, ad hoc networks can be clearly differentiated from cellular networks, where base-stations are needed. According to [76], the system may however have gateways to and interfaces with a fixed network.

- The preceding point often implies a distributed operation of the network, with no central control, all nodes having *a priori* the same role. This characteristic has a huge impact on the protocols design.

- Such a network has a priori no determined topology, so that a node may not be in direct radio connectivity with all other nodes in the network. In this case, nodes should be able to store, forward, and thus route the information from a source to a destination (see figure 1.1). With this capability, nodes are both terminals and routers and hence sometimes designated as *terminodes* [56]. The wireless ad hoc network is said to be multi-hop.

- An energy-constrained operation may be expected because some nodes may rely on batteries for their energy.

- Literature does not provide a definite answer to know whether or not an ad hoc network is mobile. [76] considers that dynamic topologies are a key characteristic of ad hoc networks because nodes are assumed to be mobile.

1.2.2 Some Applications

Ad hoc networks present several advantages. One of the main is to allow a quick deployment, where the telecommunication infrastructure is not available. This can be the case for reasons of cost, safety, security.

The first application of wireless ad hoc networking can be found in the **defense** environment. In 1972, the american (United States) Defense Advanced Research Projects Agency initiated a research program called Packet Radio Networks [125]. Its aim was to build a packet switched network on the battlefield for mobile users with relaying capabilities. Radio devices were likely to be carried by numerous kinds of supports (aircraft, vehicles, soldiers, ships...) creating dynamic situations. The network had also to be robust to the arrival or the departure of nodes and to link breaks.

After this precursor project, the technology found applications in civilian life in the 1980's, e.g. the **amateur packet radio networks** [128]. Today, ad hoc networks are proposed for **emergency and rescue** operations. After a catastrophe, an earthquake, or a blast, firemen and rescuers need to communicate with each other and coordinate operations sometimes in total absence of communication infrastructure.

For few years, applications are now anticipated for commercial deployments. We provide here a short list of known projects with examples of applications:

Inter-vehicle communication: Emergency and congestion warnings (figure 1.2), traffic information, lane change assistance, or cooperation at road intersections are provided by multi-hop communications between cars.
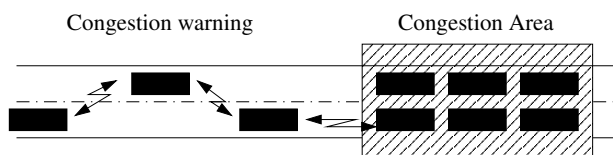
Figure 1.2: Example of inter-vehicle communication: congestion warning.

Sensor networks: Nodes are sensors in charge of collecting information on a field. As an example, scientific data can be collected by robots on a volcano and transmitted to a center located in a safe zone for analysis.

Home networks: Domestic and electronic equipments (TV, video recorder, fridge, heat controller...) are expected to be wireless enabled and able to communicate with each other in a near future via multi-hop networks.
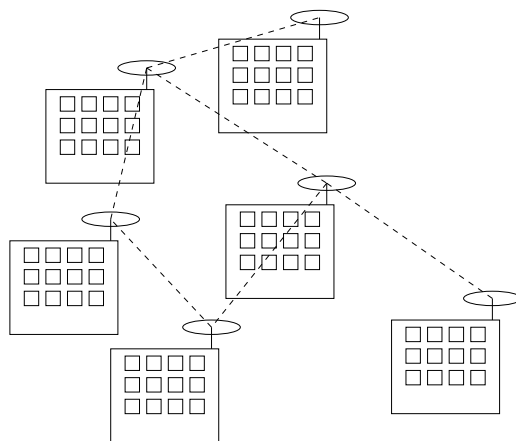
Figure 1.3: Example of meshed network: wireless metropolitan network.

Meshed networks: Backbone inter-connections for wireless networks can be of mesh topology. Such networks are used to inter-connect network equipments like access points (see figure 1.3). Such an inter-connection is foreseen in WiMAX. This can also be a good solution in indoor, where the wired network may not be easily accessible, e.g. in rail stations, hospitals, or harbors.
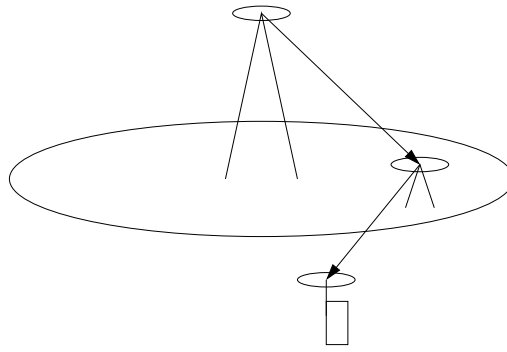
Figure 1.4: Example of coverage extension.

Coverage extension: The coverage of cellular networks, e.g. WLAN, can be extended thanks to multi-hop networks. The information of a terminal out of the range of an AP can be forwarded by intermediate terminals or specific relay nodes (figure 1.4).

1.2.3   Some Research Areas

Practically all protocol layers have to be specifically tuned to handle the dynamics of wireless transmission, distributed topology and distributed processing inherent in ad hoc networks. This paragraph lists some of the main research interests in the field.

Physical layer: Since ad hoc networks are wireless, research efforts on the transmission over radio channels is a requirement. In the field of multi-hop networks, it can be noted that several papers focus on the cooperation between nodes, and in particular on the performance of the relay channels (see e.g. [81, 109]). The design choices at the physical layer have also a great impact on higher layers. The MAC protocol may for example take into account the underlying modulation, e.g. Orthogonal Frequency Division Multiplexing (OFDM) [134] or Ultra Wide Band (UWB) [41]. Moreover, ad hoc networks are fundamentally bandwidth-constrained, this can lead to the consideration of high efficiency techniques like multi-packet reception [13].

Synchronization: This is a very critical issue for all distributed TDMA schemes. A possible solution in outdoor, now at low cost, consists in making use of the GPS (Global Positioning System) that provides a global synchronization for all nodes. Also the European satellite navigation system, GALILEO, will provide a very good

timing accuracy [85]. In this case, guard intervals have to be foreseen in MAC protocols. Another way of research is local synchronization, where nodes try to synchronize themselves by exchanging beacons with their neighborhood [79, 80, 147].

Routing layer: This is probably the most active research area in ad hoc networks, especially since the establishment of the IETF MANET working group, whose aim is to standardized IP unicast and multicast routing. More than thirty propositions of protocols have been submitted. Four of them are or are expected to become Experimental Request for Comments (RFC) documents: two reactive ones, Dynamic Source Routing (DSR) [122] and Ad hoc On Demand Distance Vector (AODV) [159], and two proactive ones, Optimized Link State Routing (OLSR) [74] and Topology Broadcast based on Reverse Path Forwarding (TBRPF) [156]. Reactive protocols look for routes on demand, while proactive ones continuously maintains routing tables.

Transport layer: Papers on the transport layer have the two fold objective to study TCP over ad hoc networks and to propose modifications or alternatives to it, see e.g. [98, 143, 184].

Energy efficiency: Minimum energy consumption in wireless communication devices is one of the major challenges for designing ad hoc networks. For the obvious sake of cost and portability, the battery-life of a wireless device has to be maximized, while maintaining network connectivity [168]. This issue has been addressed at the MAC layer (see e.g. [175]), and at the routing layer (see e.g. [176] or the power aware modification of Link State and AODV protocols [52])

Node cooperation: In multi-hop networks, a node may need to benefit from the cooperation of other nodes for packet forwarding. Some nodes could however deny this cooperation to save their battery power or even by bad behavior. Such a behavior impacts the network performance negatively. Several algorithm try to tackle this issue (see e.g. [58, 149, 150]).

Higher layers: Security, authentication, authorization, and accounting have specific issues in ad hoc networking [203]. Addressing is of particular interest in networks, where nodes have multiple radio interfaces, and/or can communicate with nodes on a wireline network via gateways (see e.g. [67]). Auto-configuration, service discovery, and address allocation are also active areas of research.

Information theory: In this field, a landmark paper is that of Gupta and Kumar [104] that tries to answer the question: How much information wireless networks can transport? In this paper, interference is treated as noise and nodes are

supposed to be fixed. In this case, it is proved that if n nodes capable of transmitting W bits/s are in a disk of area 1 m^2, the network can transport in the best case $\Theta(W\sqrt{n})$ bits-meters/s. In random networks, each node can obtain a throughput of $\Theta(W/\sqrt{n\log n})$ bits/s. [78, 102] have shown that these bounds can be overcome for networks with mobile nodes. These results will be detailed in chapter 4. In [105], Gupta and Kumar have extended their result to cooperative networks: Interference is not seen as noise but may be a source of information. Another recent work [188] presents a mathematical framework based on rate matrices for finding the capacity region of ad hoc networks.

Following [82], it can be noted that the use of information theory methods and concepts to communication networks is still widely unexplored.

MAC layer: This is historically one of the first issues of research in ad hoc networking [124]. The state of the art in this field is described in the following sections.

1.3 MAC Issues and Protocols Classification

In this section, we review the role of the MAC sub-layer and problems related to ad hoc networks. A classification of most commonly studied MAC protocols is provided.

1.3.1 MAC Issues

The basic function of the MAC protocol is to define the rules of the *radio resource sharing*. Beyond this first definition, a MAC protocol should have several desirable features in the context of wireless ad hoc networks. It should be *distributed* because nodes cannot rely on a fixed infrastructure and all nodes have initially the same role, *dynamic* in order to adapt to the changes in traffic patterns and eventually to the varying topology. Properties to ensure *reliable source to destination links* and *fairness* are also of importance. Even if fairness may be difficult to define (see e.g. [152]), we understand that nodes should not be unable to transmit. We have also seen above that radio devices should be *energy efficient*, the MAC sub-layer can implement mechanisms to address this issue. Finally, since the radio resource is scarce and ad hoc networks are particularly bandwidth constrained, it is expected that a MAC protocol provides an efficient utilization of the shared channel. MAC designers face however some specific problems [65] with respect to wireline and

cellular networks.

Half-duplex Operation

In radio devices, it is generally difficult to receive data when the transmitter is sending data. Even with two antennas, transmitted and received signals would interfere at the reception and both signal strengths can differ by orders of magnitude. With a single antenna, a circulator would be needed to isolate the reverse from the forward path (radars antennas adopt this solution). A circulator allows however in the best case a few tens of decibel attenuation at a high cost. Only signal processing at the output of the circulator could help [28].
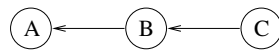
Figure 1.5: First order collision.

Hence, in the classical case, collision detection by the sender is not possible like in the Carrier Sense Multiple Access / Collision Detection (CSMA/CD) protocol for wireline LAN. Half duplex operation can lead to collisions of the first order: As shown in figure 1.5, node B cannot receive the data from node C because it is transmitting to A.

Varying Channel and Topology

The wireless channel is characterized by multi-path propagation, fast fading, and path loss. This implies much higher bit error rates (BER) than in wireline networks. Moreover, errors are correlated resulting in burst errors. Besides, changes in topology can lead to interruption in radio links.

Hence, several mechanisms are needed at the MAC and link layers to cope with these effects, e.g. smaller packets, forward error correction, frequent acknowledgments, or retransmission policies.

Hidden Terminal Problem

Carrier sensing, widely used in MAC protocols to prevent collisions, is based on the simple principle *listen before talk*. A node has to sense the channel before sending

any packet. Because of the signal power decay with distance, only nodes within a limited radius of the transmitter can however detect the carrier on the channel. This leads to the hidden and exposed terminal problems.
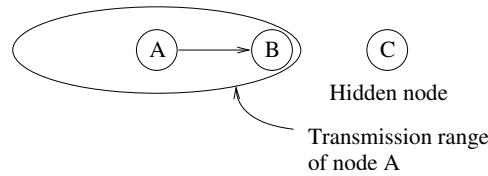
Figure 1.6: Hidden terminal problem.

An hidden terminal is a node within the transmission range of the intended destination node but out of range of the sender. In figure 1.6, node B is within the range of A, but C is not, C is hidden from A. Let us assume that A is transmitting data to B. By using carrier sensing, C would not be able to detect the signal from A and can thus decide to transmit also data to B. Such a transmission provokes a collision at B, both packets may be lost. "Two packets colliding at a receiver" is termed a second order collision.

If the received power level from node A is much higher than the received signal strength from C, B can decode A's packets. This can be easier if the packet from A arrives slightly before the packet from C, such that B can synchronize with A. This phenomenon is called the capture effect. A typical value for the needed signal strength difference for terminals operating with IEEE 802.11b is 10 dB.

The hidden terminal problem degrades the performance of the carrier sensing based protocols. We will see later in this chapter how MAC designers solve this issue.

Exposed Terminal Problem and Spatial Reuse

An exposed node is within the range of the sender but out of the range of the destination node. Let us consider the network of figure 1.7 with B transmitting data to A. C is within the range of B but not that of A, so it is exposed. By applying carrier sensing, C prevents itself from sending data to D, although its transmission wouldn't necessarily result in a collision.

Since there is no packet loss, the exposed terminal problem is less of an issue than the hidden terminal problem. However, the channel is under-utilized. Better

12
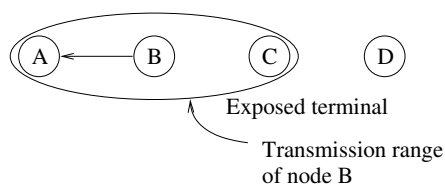
Figure 1.7: Exposed terminal problem.

performance can be expected by solving the exposed terminal problem. Also, MAC protocols for multi-hop networks have to take advantage of the spatial reuse of resources: On the same channel, two simultaneous transmissions can be successful if they are sufficiently far from each other. This is a source of capacity improvement [130].

1.3.2 MAC Protocols Classification

From the literature, two major families of protocols are:

(i) the contention-based protocols

(ii) the conflict-free protocols

In the former case, the channel is acquired by the nodes for each packet to be transmitted, whereas in the latter case, the channel is reserved for a certain amount of time and transmissions are conflict-free. In most protocols, the reservation is based on a random access. Among conflict-free protocols, we can further distinguish Time Division Multiple Access (TDMA), Code Division Multiple Access (CDMA), and Frequency Division Multiple Access (FDMA). In the next section, we provide a short overview of contention-based protocols. A bibliography of TDMA based protocols is given in section 1.5.

1.4 Contention-Based MAC Protocols

1.4.1 Basic Mechanisms and Classification

Contention-based protocols are widely discussed in literature. Some basic mechanisms have to be considered to study the functioning of such protocols. These are:

- Packet sensing: Following [95], we use this term to designate ALOHA-like random access. Contrary to carrier sensing, the channel is not sensed before transmission and a packet is recognized only when it is entirely received.

- Carrier sensing: A node performing carrier sensing is able to know whether or not the channel is busy, i.e., another node in its communication range is transmitting.

- Back-off algorithm: After a collision or possibly if the channel is sensed busy, the transmission of a packet is re-scheduled after a random delay. This delay is uniformly drawn in a time window, called the contention or back-off window. This prevents channel congestion.

- Handshaking: Exchange of short control packets precedes transmission of data packets. Rate of collisions of data packets and wasted time in case of short packet collision are thus reduced.

- Busy tone: The hidden and exposed terminal problems can be solved by the transmission of a busy tone by the receiver and/or the sender.

- Virtual carrier sensing: Control and/or data packets include the duration of current transmission or handshake, so that neighboring nodes are aware of the current handshake duration even if they do not sense any carrier. This method helps solving the hidden and exposed terminal problems.

- Collision resolution: The resolution of a collision is not deferred in the future as with the back-off algorithm. Instead, competition for the channel is solved as soon as it occurs. HiperLAN 1 is an example of protocol based on collision resolution (see section 1.4.2).

- Link level acknowledgment: The MAC protocol helps to improve link reliability by using acknowledgment control packets (ACK).

- Power control: The transmit power is controlled in order to reduce the energy consumption and to limit the amount of interference in the network.

This is of course not an exhaustive list. Moreover, due to rather interdependence between some of these mechanisms, a clear classification of contention-based protocols is difficult. Figure 1.8 is an attempt to classify the contention-based protocols according to some of the most important ideas described above.

Contention based Protocols

With handshake

Without handshake

Packet sensing

Carrier sensing

Carrier sensing

Packet sensing

MACA
MACAW
FAMA–NPS

Without
control channel

With
control channel

With
control channel

Without
control channel

ALOHA
S–ALOHA

Receiver
oriented

FAMA–NCS
MACA–BI
MACA/PR
DFWMAC

DBTMA
DBTMA/DA
PAMAS

BTMA
RI–BTMA

NP–CSMA
1 persistent CSMA
p persistent CSMA
EY–NPMA

Quality of
service

Energy
efficient

Collision
resolution

Figure 1.8: A classification of the contention-based MAC protocols for ad hoc networks.

Handshaking can be seen as a breakthrough for solving the hidden terminal problem (one of the main concerns of MAC designers). Moreover, we believe that carrier sensing, as opposed to packet sensing, plays a central role for todays protocols. Let us now examine in more details the contention-based propositions.

1.4.2 Description of Handshaking Free Protocols

ALOHA and slotted ALOHA (S-ALOHA) are historically among the first wireless MAC protocols [37]. Analysis for both infinite and finite populations in a fully connected network (all nodes are in the transmission range of each other, see figure 1.9) can be found in [169]. Originally designed for a star topology, such protocols can be used in an ad hoc network [185]. [153] shows that capture increases the throughput of ALOHA in multi-hop networks. Research on ALOHA and its variations is very active because it is the basis of many access schemes, in particular for channel reservation. One of the important recent publications [45] revisits the notion of spatial reuse and proposes protocol optimization mechanisms in a multi-hop environment.

As an improvement of ALOHA, the Carrier Sense Multiple Access (**CSMA**) protocol with collision detection (CSMA/CD) enjoys a huge success in wireline LAN. Collision detection is however difficult to implement in radio devices because of the aforementioned half-duplex operation. In non-persistent CSMA, a terminal with
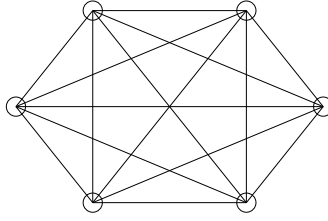
Figure 1.9: A fully connected network with six nodes.

a ready packet starts by sensing the channel. The packet is transmitted only in case the channel is found idle. Otherwise, the transmission is rescheduled to some random later time. At this instant, the channel is sensed again and the algorithm is repeated [131]. Although CSMA achieves much higher channel utilization than ALOHA, its performance considerably degrades in presence of hidden nodes [132].

In order to cope up with the problem, [132] proposes an alternative to CSMA called the Busy Tone Multiple Access (**BTMA**) protocol that divides the bandwidth in a data and a busy-tone channel: As long as a node senses carrier on the former, it transmits a busy-tone signal on the latter in order to prevent transmissions from hidden terminals. [200] further improves this solution in the Receiver-Initiated BTMA (**RI-BTMA**): Data is preceded by a preamble including the destination address, so that only the intended receiver sets its busy-tone signal.

Figure 1.10: EY-NPMA (HiperLAN 1) channel access scheme for a unicast packet.

CSMA is the basic mode of the HiperLAN 1 [190] MAC protocol: If a station, ready to send a packet, senses the channel idle for a certain time period, it is allowed to transmit. As the network load increases the Elimination Yield - Non-preemptive Priority Multiple Access (**EY-NPMA**) protocol is applied. EY-NPMA is based on a collision resolution mechanism that results in a single winning station with high probability (over 96% [39]). The access mechanism is divided into three phases (figure 1.10). After the P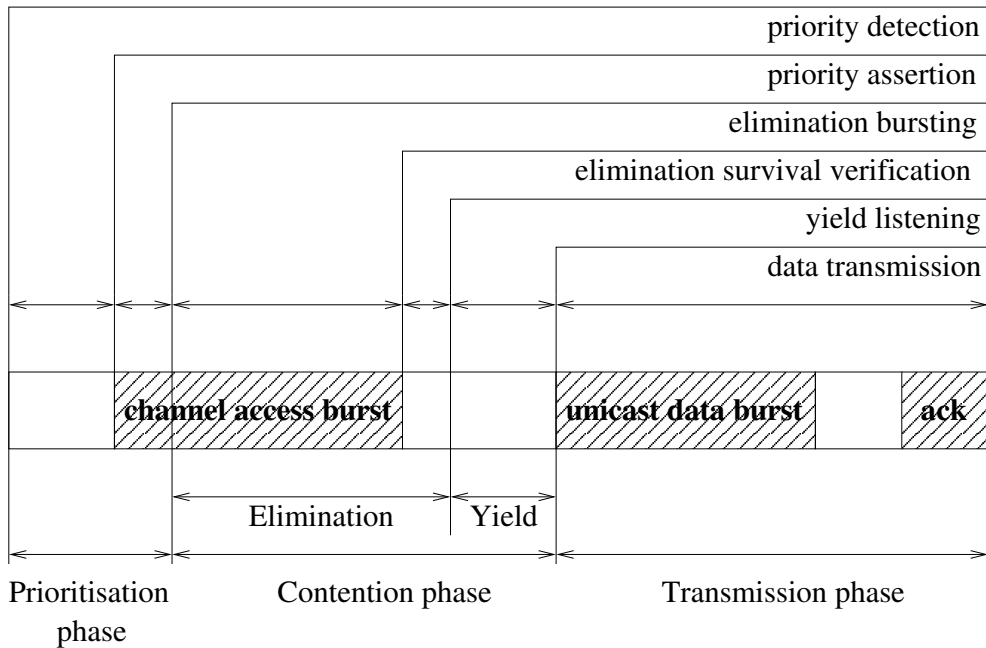riority Resolution phase, only competing stations with the highest priority survive. In the Elimination phase, surviving stations transmit a burst of random length and listen to the channel. Terminals that hear the carrier withdraw from the competition. During the Yield phase, a station senses the channel for a random time. If it is idle, it is allowed to transmit. EY-NPMA doesn't take into account the hidden terminal problem. The Elimination burst can however help detecting hidden nodes in some cases [164].

1.4.3 Description of Handshaking Based Protocols

Karn [127] brought with **MACA** a substantial advance by introducing the Request-to-Send / Clear-to-Send (RTS/CTS) handshaking for wireless networks (figure 1.11). The idea of carrier sensing was totally eliminated due to the poor performance of CSMA in multi-hop networks. Hence, RTS is sent by S according to the ALOHA protocol, stabilized by a binary exponential back-off (BEB). If D correctly receives the RTS, it sends back a CTS packet. RTS and CTS include information on transmit packet data length. Any station hearing the CTS prevents itself from transmitting, so that the hidden terminal problem is partly solved. A station hearing the RTS but not the CTS is allowed to transmit, thus the exposed terminal issue is also reduced. Moreover, only short control packets RTS or CTS are lost in case of collision: The time to solve the competition is reduced and the data packet doesn't need to be retransmitted.

BEB may result in an unfair sharing of resources [54]. An improved version called **MACAW** suggests three modifications of the back-off algorithm used by MACA.

- The first one is the Multiplicative Increase and Linear Decrease (MILD) mechanism: Upon a collision, the back-off window is multiplied by 3/2 and decreased by one unit at each success.

- Data packets include a field in their header which contains the current value of the back-off window. Whenever a station hears a packet, it copies this value

17
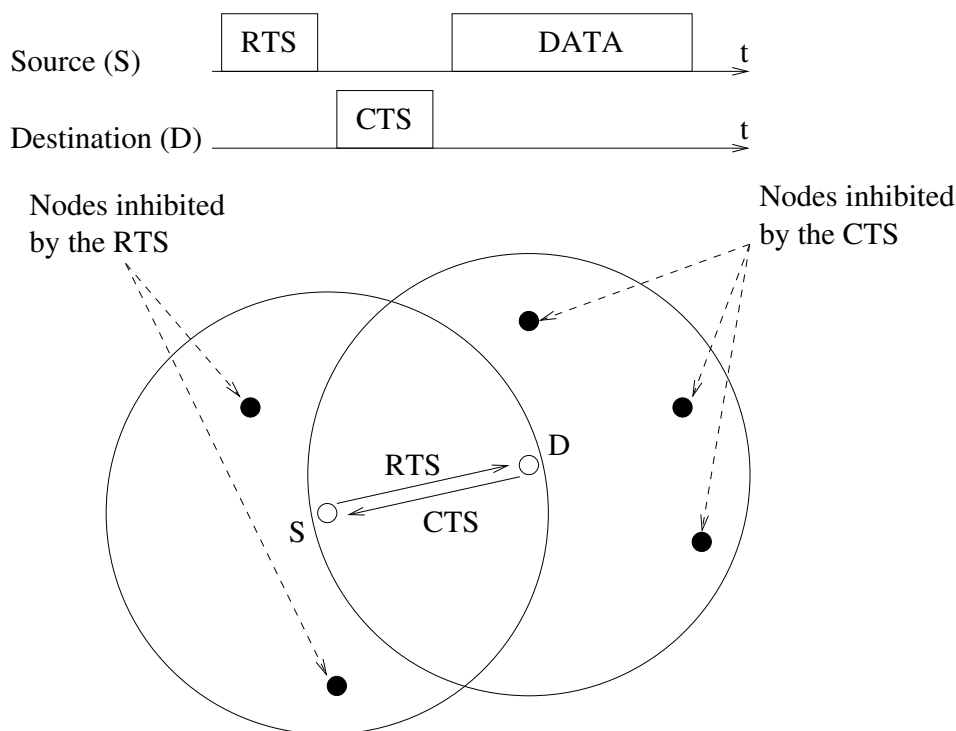
Figure 1.11: MACA data transmission with RTS/CTS handshake.

into its own window size.

- The back-off algorithm is run in each station independently for each stream.

Finally, the MACA handshake is modified in a RTS / CTS / DS / DATA / ACK sequence. Before sending data, a short Data Sending packet is transmitted by the sender to inform that the RTS/CTS dialogue was successful. A link level acknowledgment, optional with MACA, is considered here as mandatory.

Adapting the handshaking to the family of busy-tone protocols, the Dual-BTMA [106] (**DBTMA**) improves the short dialogue with two busy-tone channels, one for the sender, the other for the receiver. This has the main aim to combat both hidden and exposed terminal problems. An extension of DBTMA for terminals with directional antennas (**DBTMA/DA**) is proposed in [114]. A comparison of protocols using this technique can be found in [113].

Again in the family of out-of-band signaling protocols, the Power Aware Multi-Access Protocol with Signaling (**PAMAS**) considers reducing the energy consumption at each node. In order to avoid consuming any power for receiving packets with destination address of other nodes, [175] suggests the use of the RTS/CTS handshake on a separate channel. This helps nodes to know when and for how long

they can power themselves off. For a survey on energy efficient MAC protocols for ad hoc networks see [141].
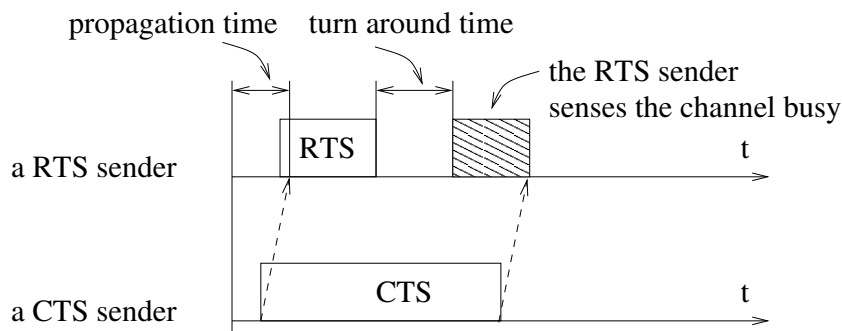
Figure 1.12: In FAMA, the dominating CTS plays the role of a busy tone.

Since using two separate channels is more cumbersome, modifications have been performed, e.g. [95] proposes to use the CTS as a kind of in-band busy-tone. In **FAMA** protocols, CTS and RTS lengths are calculated with due consideration of the radio propagation delay, the processing time, and the turn-around time, i.e., the time for radio device to switch between transmission and reception states. Hence, a station that transmits a RTS simultaneously with the CTS of a receiver would hear at least a portion of the CTS at the end of its transmission. This is shown in figure 1.12. FAMA-NCS is based on carrier sensing, while FAMA-NPS is based on packet sensing.
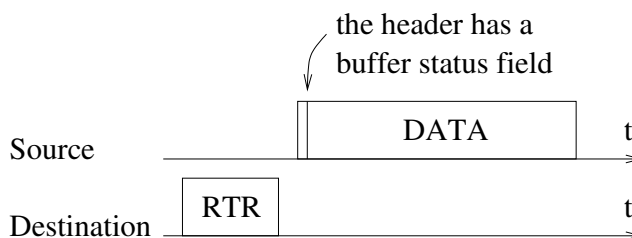
Figure 1.13: Data transmission with the receiver-oriented MACA-BI protocol.

Consideration of turn-around time imposes a heavy penalty on the use of re-sources. It is further observed in [186] that the relevant area of contention is the reception range of the receiver. In the MACA By Invitation (**MACA-BI**) protocol, the sender waits for an *invitation* by the receiver in the form of a Ready-to-Receive (RTR) packet (figure 1.13). This simple handshake improves the performance of the MAC layer with respect to CSMA, FAMA, and MACA by reducing the overheads.

19

An additional advantage is the reduced power consumption. However, MACA-BI relies on traffic prediction at the receiver. It must indeed be able to approximately know when the sender has a ready packet. This can be helped by the indication in data packets of the buffer status at the sender side. The receiver-oriented notion is exploited in chapter 3 by CROMA.

The notion of quality of service (QoS) in the MACA family of protocols is introduced in [140] with MACA with Piggyback Reservations (**MACA/PR**). Non real-time packets are transmitted according to the RTS / CTS / DATA / ACK sequence. The first packet of a real-time flow is sent using the RTS/CTS handshake and reserves time intervals for subsequent packets along the flow path in the network. Then, each real-time packet and associated ACK piggybacks a reservation for the next packets. MACA/PR is coupled with a QoS routing algorithm.

DFWMAC: the Distributed Coordination Function of IEEE 802.11

The Distributed Foundation Wireless MAC (**DFWMAC**) is the protocol adopted by the IEEE 802.11 standard [35] for its so called Distributed Coordination Function (DCF). This scheme is the direct heir of MACA/MACAW on the one hand, and of CSMA on the other. From MACA, it has taken the handshake, the virtual carrier sensing, the BEB ; from MACAW, the ACK control packet ; from CSMA, the physical carrier sensing.
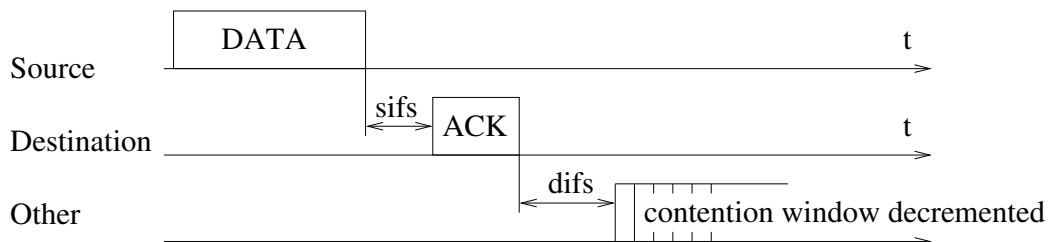
Figure 1.14: Basic mode access of IEEE 802.11 DCF.

DCF has two modes that can be dynamically chosen according to the transmit data packet length. In the basic mode (figure 1.14) rather used for short packets, a station ready to send picks a random number of slots (of length SlotTime) in its back-off window $[0; CW]$. The back-off timer is decremented at each time-slot only if the channel is sensed idle for more than a DCF inter-frame space (DIFS) interval. The station is allowed to transmit when the timer reaches zero.

20

The contention window parameter (CW) has an initial value CW_{min}. After each failure of the transmission, CW is doubled. It is however upper bounded by CW_{max}. CW is reset to the minimal value after each successful transmission.

Each correctly received frame is acknowledged by an ACK control packet. The interval between data and ACK is set to the Short inter-frame space (SIFS). Note that SIFS is shorter than DIFS, so that a receiver acknowledging a frame has priority over a new transmission.
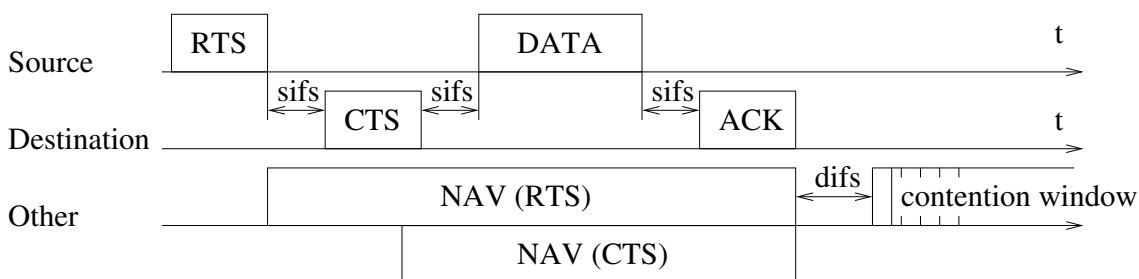
Figure 1.15: RTS/CTS mode access and NAV setting in IEEE 802.11 DCF.

With this handshake procedure, DFWMAC takes into account the hidden terminal problem. The basic access described above is applied to the RTS transmission (figure 1.15). The short dialogue is not recommended for data packet lengths smaller than the RTS because in this case a collision on the data packet is less time consuming.

In addition, a virtual carrier sensing is implemented. RTS, CTS and data packets include the information related to the sequence duration. Stations hearing this field set their Network Allocation Vector (NAV) to one, indicating that the channel is busy for the time of the transfer. Finally, the protocol allows the successive transmissions of several segments, each one individually acknowledged by the receiver.

Bianchi in [55] provides an accurate analysis of the saturation throughput of DCF in a fully connected network and for a finite number of terminals under ideal channel conditions. [60] analytically derives the average size of the contention window that maximizes the throughput and proposes to tune the back-off algorithm to increase the capacity. [68] studies the influence of hidden nodes on the performance of the two access modes of DCF.

Sometimes, IEEE 802.11 DCF has been criticized for lack in performance in multi-hop networks [202]. It is claimed that it has still the hidden terminal problem, it doesn't solve the exposed terminal problem, and the back-off algorithm is judged to

21

cause unfairness, especially with TCP. Besides, questions regarding the effectiveness of the RTS/CTS handshake have been raised in [201]. If the interference range of radios is much larger than the transmission range, the efficiency of virtual carrier sensing is reduced.

1.5 Time Division Conflict-free MAC Protocols

The contention-based protocols are particularly adapted to bursty traffic but can exhibit a high overhead due to frequent contention if the network load increases. On the other hand, deterministic scheduling may be preferred for networks with heavy load, carrying mixed traffic and realizing sophisticated functions. That is the reason why attention has to be paid to conflict-free MAC sub-layers, i.e., TDMA, FDMA, and CDMA based protocols. The present section focuses on the time division multiple access literature. Hereafter, we propose to distinguish papers related to the complexity issue, from those focusing on node and link allocation protocols and algorithms.

1.5.1 Complexity Issues

In the most basic TDMA scheme, time is divided in slots and each slot is preassigned to a single user (figure 1.16). The slot assignment or schedule is static and follows a periodical pattern, called *TDMA cycle* [169]. The cycle length in slots equals the number of nodes in the network in this basic case. Assuming that all users have always a packet ready to be sent, the pure TDMA scheme is optimal in terms of throughput for fully connected networks.
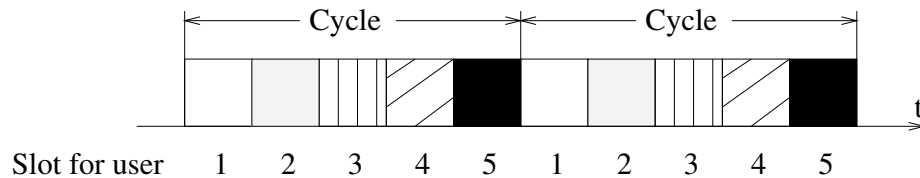

Figure 1.16: TDMA slot allocation for five users.

In multi-hop networks however adequately adapted TDMA schemes can generate significant performance improvement through spatial reuse of resources (see

section 1.3.1). This gives rise to the problem of optimizing the static slot allocation, e.g. in order to maximize the network throughput or to minimize the TDMA cycle length. It is very similar to the graph coloring problem or to the frequency assignment problem in cellular networks. Here, labels are time-slots.

In multi-hop networks, possible slot allocations are studied under the constraints of the broadcast nature of the radio channel and of the transceivers capabilities (e.g. the half-duplex operation, see section 1.3.1). Two categories of optimization problems frequently addressed in the literature are the broadcast scheduling problem and the link scheduling problem [73].

1. In the broadcast scheduling problem, a packet transmitted by a node has to be received by all its one-hop neighbors. In figure 1.17, neighbors of A and B are not allowed to transmit on the considered slot. On the contrary, nodes A and B, that are three-hops away are allowed to share the time-slot. In fact, we have here a conflict-free transmission, i.e., without collision, if two nodes one-hop or two-hops away are not assigned to the same slot. In the broadcast scheduling problem, slots are allocated to the nodes.

2. In the link scheduling problem, a packet transmitted by a node has to be received by a single particular one-hop neighbor. In figure 1.18, the packet sent by A is intended to B. Contrary to the previous problem, node C, a neighbor of A, is allowed to transmit on the same slot because its transmission towards D does not interfere. Thus, there is no conflict between two links if there is no collision at the intended receivers. In the link scheduling problem, slots are allocated to the links.
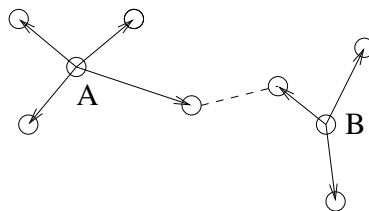

Figure 1.17: Broadcast scheduling.

Arikan [42] has shown that constructing an optimal schedule for the link scheduling problem to optimize throughput is NP-complete. According to Ephremides and Truong [83], scheduling broadcast transmissions for throughput optimization is also
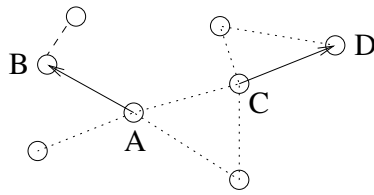
Figure 1.18: Link scheduling.

NP-complete. Finding a maximum broadcasting set or a broadcast frame of minimum length are also NP-complete problems [167]. It is possible to compute a minimum length TDMA cycle in a polynomial time only if spread spectrum communications are assumed [108].

Since most of the optimization problems are NP-complete, authors look for suboptimal algorithms and protocols while trying to maximize the resource utilization. Algorithms and protocols can be classified in two parts derived from the above two classical problems: the node allocation and the link allocation algorithms. A classification of the TDMA based conflict-free MAC algorithms and protocols is given in figure 1.19. Early papers focus on static networks with centralized algorithms and then deterministic reservations of the slots. The need to address topology changes and varying traffic patterns leads to dynamic reservation schemes based on random access.

1.5.2 Node Allocation

Node allocation algorithms and protocols try to solve the broadcast scheduling problem. Each node is assigned one or several time-slots in each TDMA cycle during which it can transmit to any of its neighbors. Thus, intended receivers are not taken into account, and two nodes that are one-hop or two-hops away cannot transmit simultaneously. Solutions can be classified in centralized algorithms, schemes with deterministic reservation, schemes with random reservation, and time-spread protocols.

Centralized Algorithms

These algorithms assumes fixed networks. Moreover, each node is assumed to know the entire topology and to execute the algorithm which produced the schedule.

24

Figure 1.19: A classification of the TDMA based conflict-free MAC algorithms and protocols.

Figure 1.20: Static multi-hop network with eight nodes.

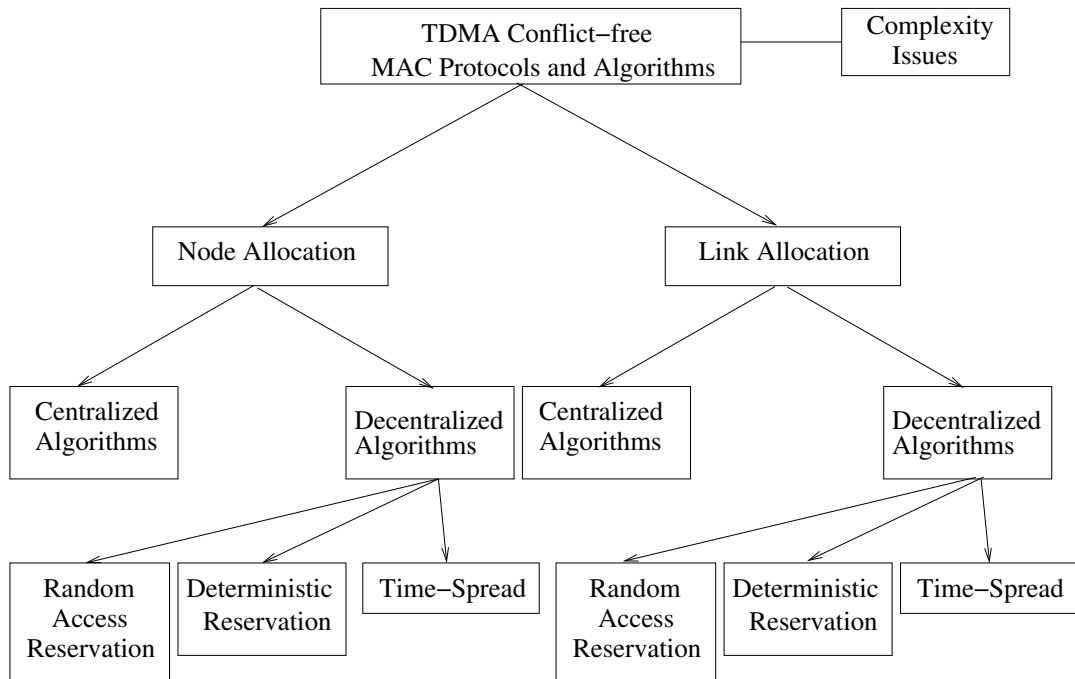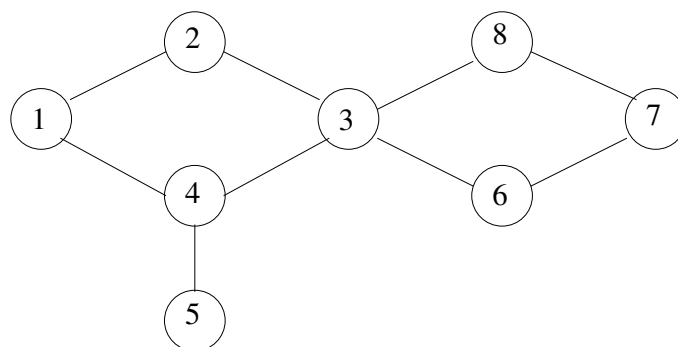**Ephremides and Truong** [83] propose a centralized algorithm which can be computed in a polynomial time. Let $A = (a_{i,j})$ be the connectivity matrix of the network graph, i.e., $a_{i,j} = 1$ if $i = j$ or if nodes i and j are in communication range, and $a_{i,j} = 0$ otherwise. An important result is that the zeros of A^2 provide the node pairs that are at least three-hops away. These pairs can share the same time-slot in the broadcast scheduling problem without interfering. For example, the connectivity matrix of the network showed in figure 1.20 is:

$$A = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

We can verify on the A^2 that for example nodes 6, 7 and 8 are more than three-hops away from node 1:

$$A^2 = \begin{bmatrix} 3 & 2 & 2 & 2 & 1 & 0 & 0 & 0 \\ 2 & 3 & 2 & 2 & 0 & 1 & 0 & 1 \\ 2 & 2 & 5 & 2 & 1 & 2 & 2 & 2 \\ 2 & 2 & 2 & 4 & 2 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 & 2 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 3 & 2 & 2 \\ 0 & 0 & 2 & 0 & 0 & 2 & 3 & 2 \\ 0 & 1 & 2 & 1 & 0 & 2 & 2 & 3 \end{bmatrix}.$$

Based on this property, the proposed algorithm iteratively chooses nodes according to their distance from each other. A single slot is allocated to the obtained maximum set of nodes.

Ramaswami and Parhi present in [167] a two phased heuristic approach. In the first phase, each node is assigned to a single slot. Nodes are processed in a certain order. Each node is assigned the smallest slot number not already assigned to a one-hop or a two-hop neighbor. After this first allocation phase, the cycle length is fixed and the second phase is run. It aims at reusing the allocated slots. Each node is processed in turn again and scheduled in as many slots as possible. The assignment is based on the slot availability given the first phase allocation.

Table 1.1 shows the slot allocation for the network of figure 1.20 after the first phase (T) and the second phase (T*).

Table 1.1: Slot allocation of the network of figure 1.20 according to the algorithm of Ramaswami and Parhi.

Slots	Nodes							
	1	2	3	4	5	6	7	8
1	T					T		
2		T			T		T	
3			T					
4				T			T*	
5	T*							T

In the two preceding algorithms, the node processing order is arbitrary. Moreover, if a node can be attributed several slots in the first phase, the choice is arbitrary either. According to **Hung and Yum** [115], the first phase should leave the greatest freedom for the second phase. To this end, the processing order and slot choices take into account the number of one-hop and two-hop neighbors as well as the degree of freedom left for the second phase. It is claimed that the two preceding propositions are outperformed in terms scheduling delay, TDMA cycle length, and fairness.

[166] generalized the previous idea by providing a unified algorithm for the channel assignment, called Unified Assignment for Medium Access (**UxDMA**). The algorithm consists of two phases, a labeling phase and a coloring phase. In the labeling phase, each node is assigned a unique label, as previously. In the coloring phase, nodes are considered in decreasing order of labels. The considered node is colored in a greedy fashion, i.e., a color violating none of the selected constraints is chosen. The paper gives three examples of labeling: the Random (RAND), the Minimum neighbors first, and the Progressive minimum neighbors first (PMNF) methods. PMNF provides the better performance.

Heuristic algorithms using graph theory are seen to be the early time approximations for solving the NP-complete scheduling problems. Another group of publications using different approaches like gradual neural networks [90, 91], mean field annealing [198], or genetic algorithms [64] are of considerable interest too.

All these algorithms are centralized and designed for static nodes, whereas ad hoc networks are by nature without any infrastructure, or centralized unit, and can be mobile. Distributed algorithms and protocols can be classified in three classes:

- protocols with deterministic reservation (most of the time they are derived from the centralized algorithms described above and do not address topology changes)

- protocols with random access based reservation (with a handshake protocol)

- Time Spread Multiple Access (TSMA) based protocols

Deterministic Reservation

Ephremides and Truong give in [83] a decentralized version of their algorithm. It starts initially with a skeleton schedule for which the ith slot is reserved for the ith node. The schedule information is gathered in a matrix, each row of which corresponds to a time-slot and each column to a node. Using their reserved TDMA slot, all nodes broadcast the status of their own column, and in the next TDMA cycle, the status of the columns of their neighbors. Thus, each node can locally know at which slot it is allowed to transmit, using informations of the schedule matrix and a priority rule. A periodic rotation of the priority rule provides a fairer schedule. A similar proposition is the Dynamically Reserved Slot Management (**DRSM**) [157].

Cidon and Sidi propose in [73] a protocol for both node and link assignment. The channel is split into a control channel and a transmission channel. For each slot of the information channel, there are two corresponding segments in the control channel called the request and the confirmation segments. Conflicts are solved using priorities among the nodes. The resulting assignment is maximal, i.e., no additional ready node can transmit during a given slot without interfering with the transmission of any other assigned node. The paper proposes also two methods to improve fairness.

[48] presents a Node-Activation Medium Access (**NAMA**) protocol based on Neighborhood-aware Contention Resolution (NCR). NCR is a method that allows each node to elect deterministically one or multiple winners for channel access at a given time-slot. It is assumed that every entity knows the identity set of its contenders. For each of them a given node computes a unique priority number thanks to a pseudo-random number generator, whose seed is related to the node identity and the slot number. At each slot, this method provides unique priority numbers to nodes known by their neighbors and that change at each time-slot. The node with the highest priority accesses the channel.

Ramaswami and Parhi in [167] propose a token based algorithm derived from

their centralized heuristic approach presented above. It assumes that each node knows the identity of its neighbors and maintains a schedule table for each of them. A token is generated at a source node and routed to all the nodes in the network. The path taken by the token is a Depth First Search (DFS) of the graph. A node picks its slot only after each of its children in the DFS tree has picked its slot. Two major disadvantages of the algorithm are that it may not be robust to topology changes and that no data transfer can take place during the execution of the algorithm.

Random Access Reservation

The necessity to address the problem of mobility, topology changes, slot management, and scalability (in the previous algorithms the TDMA cycle depends on the network size) gives rise to a new family of protocols where the reservation of the slots is done via a random access, most of the time a handshaking.

In this area, the most important and probably one of the first work that proposed this method is the paper of Zhu and Corson in [204] that defines the five-phase reservation protocol (**FPRP**) for mobile ad hoc networks. FPRP employs a contention-based mechanism where nodes compete with each other to acquire TDMA slots. A *multi-hop ALOHA* policy has been also developed to support FPRP. This policy uses a multi-hop, pseudo-Baysian algorithm to calculate contention probabilities and enables faster convergence of the reservation procedure.

The TDMA frame structure is as follows. There is a reservation frame (RF) for the handshake and it is followed by a sequence of information frames (IF) for data transmission (see figure 1.21). There are N information slots (IS) in an IF. If a node wants to reserve an IS, it contends in the corresponding RS. A RS is composed of M reservation cycles (RC). A reservation cycle has five phases:

1. the reservation request phase (RR), where nodes make their requests for reservations

2. the collision report phase (CR), where nodes report collisions that just occurred in phase 1

3. the reservation confirmation phase (RC), where nodes make confirmations of their requests

4. the reservation acknowledgment phase (RA), where nodes that heard a RC acknowledge with a RA packet. Hence, the nodes two-hops away are informed

of the reservation.

5. the packing and elimination phase (P/E), that enables efficient spatial reuse of the same slot, and eliminates some possible deadlocks that may exist between adjacent nodes
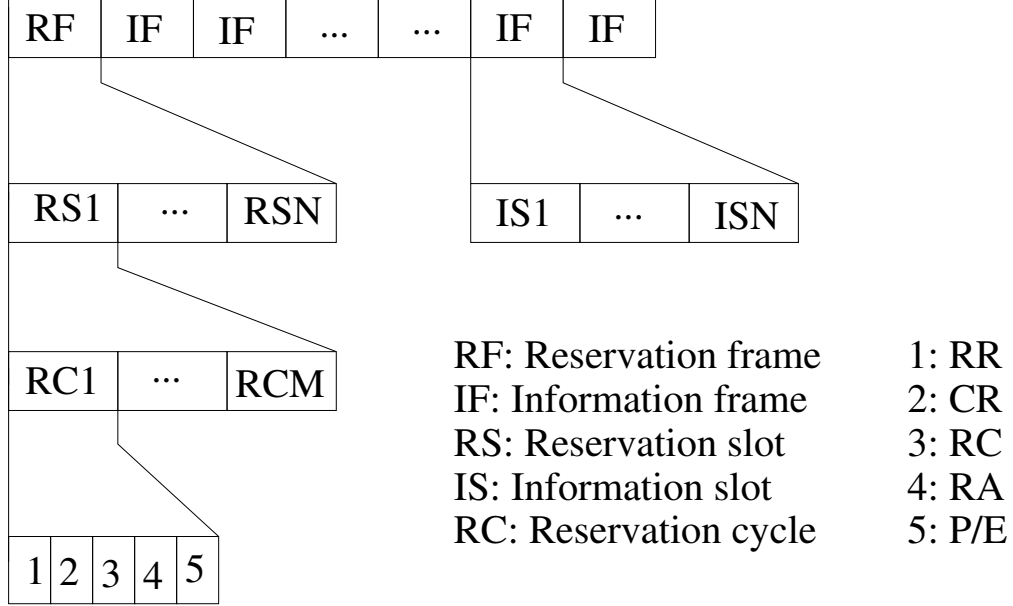
Figure 1.21: FPRP frame structure.

In terms of graph coloring, FPRP and RAND [166] exhibit comparable performances, i.e., the number of colors required for different network topologies are similar for the two algorithms, and it is not far from the degree-based lower bound.

Time Spread Protocols

Chlamtac and Fargó in [69] opened a new area of research by defining topology-independent protocols, making transmission schedules immune to topology changes. The original TSMA protocol is a node allocation version.

Each node is assigned a distinct polynomial P over $GF(q) = \{0, ..., q-1\}$ (q must be of the form $q = p^m$, where p is a prime number and m is a positive integer, here authors take $m = 1$) with maximum degree k. They define the graph of P as the set of pairs $(\beta, P(\beta)) \in GF(q) \times GF(q)$. Frames are divided in q sub-frames of q time-slots (figure 1.22). In the sub-frame β, a node transmits its packet in the slot $P(\beta)$, i.e., there is a one-to-one mapping of the polynomial graphs on the set of time-slots (figure 1.22).

Then, they consider a specific sender S and one of its neighbors D. D is guaranteed to receive at least once the packet sent by S within a frame if $q \geq kD + 1$, where D is an upper bound on the number of neighbors in the network. Indeed, since the degree of the polynomials is less than k, two graphs cannot have more than k common points. Thus, at D, the packet from S cannot collide more than kD times in the frame. As this packet is transmitted q times, it is received without any collision at least once.

This protocol is called Galois Radio Network Design (**GRAND**). GRAND guarantees conflict-free operation. However, the frame length depends on a maximum bound on the number of one-hop and two-hop neighbors for a node. This implies that frame lengths can be very long: GRAND is not scalable. It can be noted that the traffic pattern is not taken into account and that a unique code has to be allocated to each node. [123] proposes an optimization of GRAND. Figure 1.23 proposes a classification of the node allocation algorithms and protocols.
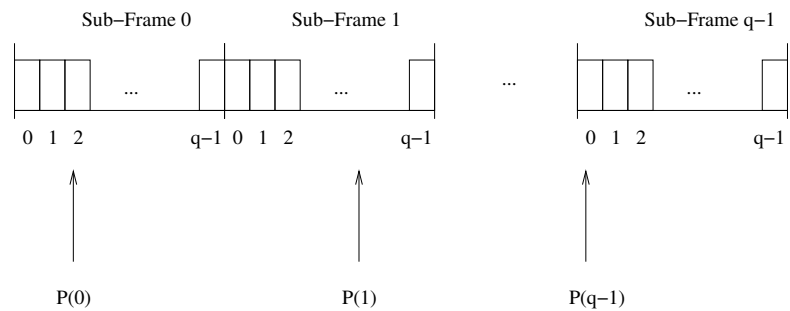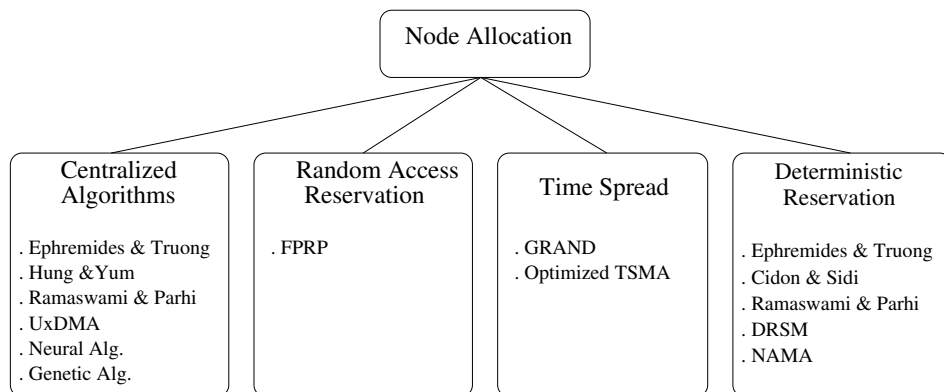

Figure 1.22: TSMA frame structure.

Figure 1.23: Classification of the node allocation algorithms and protocols.

Protocols and algorithms for the broadcast scheduling problem have several advantages [182], e.g. implied TDMA cycles are shorter and some higher layer multi-

destination protocols can take advantage of the broadcast nature of the radio channel. However, if a node has n neighbors, n cycles are required to transmit distinct single destination packets. Moreover, two neighbors cannot transmit simultaneously even if their packets will not collide at their intended destination. Link allocation schemes try to address these problems.

1.5.3 Link Allocation

Link allocation algorithms and protocols try to solve the link scheduling problem. They allocate unique time-slots in the TDMA cycle to a node for each directed link to its neighbors. Thus, the link method allows two neighboring nodes to transmit simultaneously whenever the destination nodes are not neighbors of both transmitting nodes. A better spatial reuse is expected from link allocation schemes. In addition, every node can send a single packet to each of its neighbors during every cycle.

In a sense, the link allocation scheme attempts to emulate a wireline network so that all higher layer protocols may be used without modification. However, specific advantages of node allocation techniques like short TDMA cycles and fast delivery of multi-cast traffic are lost.

In this section, we use the same classification as in the previous one and distinguish four categories of schemes: centralized, distributed with deterministic reservation, random access, and time-spread schemes.

Centralized Algorithms

One of the first approaches to find a sub-optimal algorithm for link scheduling is the protocol proposed by Nelson and Kleinrock in [154] and called **spatial TDMA**. Node locations are assumed to be fixed and known. Authors translate the scheduling problem in the maximal clique problem. A clique is a set of links allowing all its elements to transmit simultaneously successfully. A single slot can thus be assigned to a clique. A maximal clique is one in which no additional links can be added without creating a conflict. The schedule is a set of maximal cliques which contains all the links of the network topology.

As for node allocation, a neural approach can be used to solve the link scheduling problem [49]. Finally, the aforementioned unified algorithm in [166] can be applied to link scheduling.

Deterministic Reservation

Cidon and Sidi propose in [73] an extension of their protocol for link allocation. Again, the channel is split into control and transmission channels. The control channel consists of a request segment and a confirmation segment (see figure 1.24).
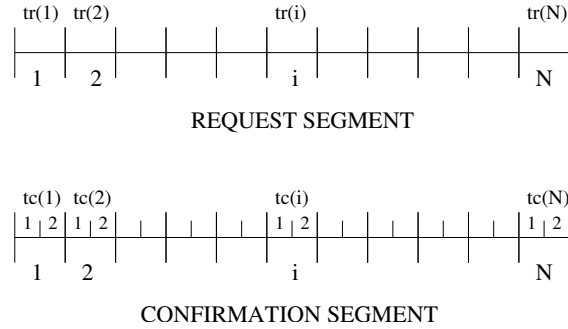

Figure 1.24: Control channel of the Cidon & Sidi algorithm.

Each segment is divided in N mini-slots where N is the number of nodes. Mini-slots in the request segment are used to specify the intended receiver. A confirmation mini-slot is divided in two parts. The first part is used to transmit a deletion signal. If this signal is sent by some of its neighbors, a node cannot transmit on this time-slot. The second part is used by a node to transmit a confirmation signal and to specify to its neighbors that it will send data on the time-slot.

[161] proposes also a protocol where the channel is divided into a data segment and a control segment and called Dynamic distributed Time-Slot Assignment Protocol (**DTSAP**). But the originality of this paper is that it addresses the problem of the topological dynamics. Here, three underlying problems arise: the network synthesis problem (establishing correct assignment tables at the genesis of the network and inclusion of new nodes), the new neighbors problem (two nodes become neighbors), and the connectivity loss (two neighbors lose their connectivity). Here, nodes exchange connectivity tables.

[48] uses the aforementioned NCR method to construct a link allocation protocol (**LAMA**) and a pairwise link allocation protocol (**PAMA**). But, in these cases, spread spectrum is assumed. In LAMA codes are attributed to receivers, whereas in PAMA, codes are attributed to sender-receiver pairs. The main drawback of this approach is that the one-hop and two-hop away neighbor identities are assumed to be known at each node.

Random Access Reservation

Tang and Garcia-Luna-Aceves suggest [187] that all previous algorithms and protocols are designed either for broadcasting or unicasting, but not for both. In addition, TSMA protocols suffer from two limitations: The sender is unable to know which neighbor can correctly receive the packet it sends in a particular slot, and these protocols are not scalable because the frame length must be larger than the number of nodes in a two-hop neighborhood.

Thus, they propose in [187] the Collision Avoidance Time Allocation protocol (**CATA**). CATA allows nodes to contend for and reserve time slots by means of a distributed reservation and handshake mechanisms. CATA ensures that no collisions occur in successfully reserved time-slots, and reservations support unicasting, multicasting, and broadcasting.

In CATA, a frame is divided in L slots. In each slot, four control mini-slots (CMS) are followed by a data segment (see figure 1.25).
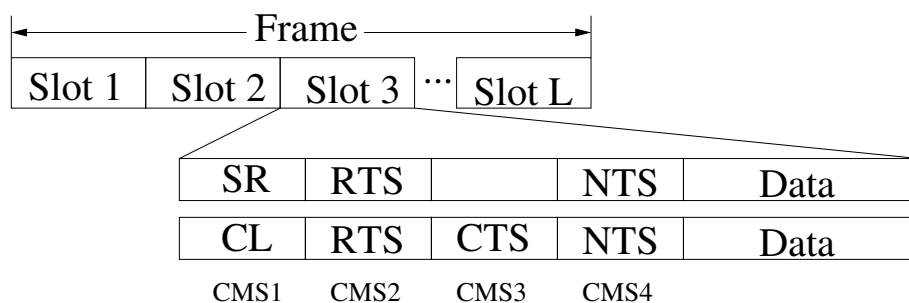
Figure 1.25: CATA frame structure.

Every node receiving data in a given slot transmits a slot reservation (SR) packet in CMS1, this is a busy tone to senders attempting to establish transmissions. Every node that sends data transmits a RTS during CMS2 to jam any possible RTS addressed to its neighbors. Both sender and receiver send a not-to-send packet (NTS) during CMS4 in order to jam any broadcast or multicast reservation. The reservation of a unicast transfer is made as follows. The sender sends an RTS during CMS2 if the channel is clear during CMS1. It detects a successful reservation with the reception of a CTS during CMS3. The sender of a broadcast or multicast RTS detects the failure of its request when it either receives an NTS or noise during CMS4. If the channel is clear during CMS4, the reservation is considered successful.

Figure 1.26: DPRMA frame structure.

The idea, taken from MACA, to use RTS/CTS handshaking in TDMA based protocols is also exploited in [121] for mixed voice and data traffic. The frame structure of the Distributed Packet Reservation Multiple Access (**DPRMA**) protocol is shown in figure 1.26.

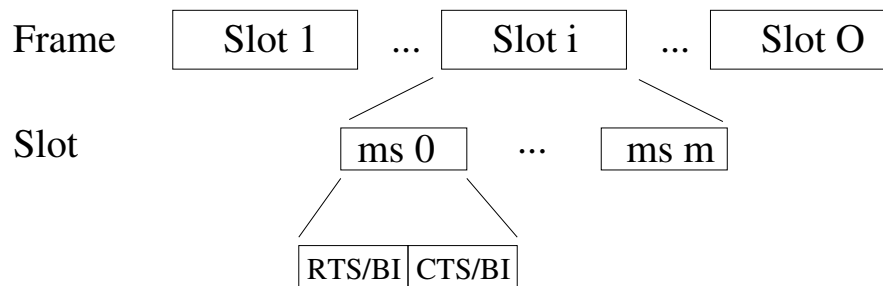As in MACA, nodes contend on the mini-slot 0 using the RTS/CTS handshake. In case of failure, the contention continue on the m remaining mini-slots. Otherwise, data transmission takes place. Voice terminals start this contention process on mini-slot 0 with probability 1, while data terminals start with a probability p smaller than 1. Moreover, winning voice terminals reserve the slot for several frames. On the contrary, data terminals can use only one slot.

A part of the RTS and CTS time intervals are used for carrier sensing. A terminal having won the slot transmits, during this part, a busy tone signal to prevent any contention on this slot.

DPRMA introduces the notion of quality of service by giving a higher priority to voice calls. It also solves the hidden terminal problem but leaves aside the exposed terminal issue. Another protocol for speech communication is proposed in [151].

Finally, to complete the picture, Space-time division multiple access (**SDMA**) [59] is discussed. Here, the nodes are assumed to know their location and time-slot are associated to locations in the network area. Inside a so called space slot, reservation is made in a random fashion.

Time Spread Protocols

Initially designed for node allocation in TDMA networks, TSMA can be used for the design of link allocation protocols [70]. [71] defines the concept of threaded TSMA. The proposed protocol is link oriented: Every neighbor of a sender receives

the transmitted packet, but it is destined to a single user. In the original TSMA protocol, called GRAND, the frame length is dependent on the maximal degree in the network. In paper [71], the authors try to overcome this limitation. The basic idea is the interleaving of several different TSMA protocols on a time-sharing basis to obtain a threaded TSMA (**T-TSMA**).
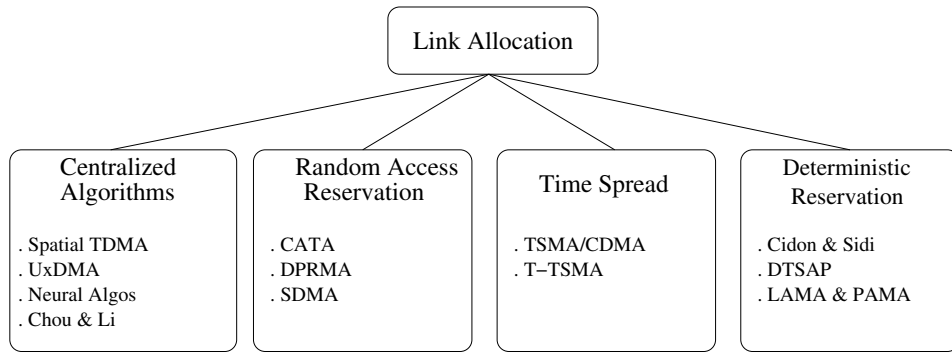
Figure 1.27: A classification of the link allocation algorithms and protocols.

1.6 Conclusion

From this chapter and the bibliographic study, some main conclusions can be drawn:

> Two families of MAC protocols are: the contention-based and the conflict-free protocols. In the former case, the channel is acquired by nodes for each packet whereas in the latter, the channel is reserved for a certain amount of data.

> In the contention-based family, IEEE 802.11 appears to have borrowed from numerous protocols, e.g. short packet handshake, physical and virtual carrier sensing, or back-off algorithm. At high input loads, TDMA based conflict-free schemes are known to be more suitable for a better channel utilization.

> However, an optimal slot allocation is not reachable since most of scheduling problems are NP-complete. Only sub-optimal solutions are proposed. Centralized algorithms don't take into account the fundamentally distributed nature of ad hoc networks, while decentralized solutions with deterministic reservation are not robust to traffic and topology changes.

> Time-spread protocols are mobility independent but can exhibit very long frame length in certain cases. Thus, they are not scalable. Moreover, they do

not address the problem of varying traffic patterns and faces the problem of unique code distribution among the nodes.

> Only adaptive slot allocation scheme seem to address the issue of mobility and traffic, while guarantying conflict-free transmissions at high input loads. Moreover, in order to favor spatial reuse, link allocation is preferred to node allocation.

Chapter 2

IEEE 802.11 DCF Performance Evaluation

2.1 Introduction

The Wireless Local Area Network (WLAN) technology specified in IEEE 802.11b is being widely deployed for high data rates networks. The areas of application include: residential networks, enterprise networks and hot spot coverage such as conference centers, railway stations, airports, hotels. In a vast majority of the cases, off-the-shelf equipment is used and the Distribution Coordination Function (DCF) of IEEE 802.11 is the regular feature implemented in such equipment.

DCF being a fully distributed protocol can be seen as a MAC sub-layer for ad hoc networking. It is probably at the origin of the renewed interest for ad hoc networks and most of the test beds that demonstrate ad hoc networking rely on IEEE 802.11 DCF. Finding an alternative to IEEE 802.11 should thus take into account the capability of the protocol to address both AP centric and ad hoc networks.

In this chapter, DCF performance in single-hop and multi-hop networks with AP is studied. Such networks present a special case whereby the AP forms a point of traffic concentration. The problems related to maximizing the number of simultaneous users, dimensioning and fine tuning network parameters are of paramount interest here.

We present a summary study on the IEEE 802.11b capacity for TCP based traffic models. In section 2.2, it is shown that a lot of simultaneous users can navigate on the web or download files with a high quality of service.

Then, we turn to the voice over Internet Protocol (VoIP) capacity (section 2.3). In enterprise and hot spots, VoIP and voice over WLAN (VoWLAN) are indeed becoming attractive technologies. The main goal is here to reduce the communication costs by merging data and voice networks. In this case, we show that IEEE 802.11b suffers from a high overhead that considerably limits the number of simultaneous calls.

The performance evaluation of single-hop networks is illustrated by an original real world case study presented in section 2.4 that aims at providing high speed Internet access to rural areas thanks to a WLAN coupled to a satellite connection.

A specific attention is paid in section 2.5 to the influence of link adaptation on throughput, namely the near-far effect. This chapter proposes some solutions to mitigate this effect. One of those is based on the utilization of a relay node able to route packets destined to low data rate users. This shows that ad hoc concepts can be efficiently adapted to infrastructure networks.

Finally, section 2.6 presents an application of multi-hop networking for coverage extension. Routing capable mobile devices or operator controlled fixed relays are shown to extend the range of fixed APs. Nevertheless, the extended covered area is gained at the expense of throughput which decreases with an increasing number of hops. In fact, we show that IEEE 802.11 exhibits weaknesses in a multi-hop environment in terms of capacity and fairness.

Most of the results presented in this chapter have been published by the author in [1, 2, 10, 11]

2.2 Capacity Results for TCP Based Traffic

Carrying TCP based traffic is the first and still the main purpose of WLAN deployment. This is an advantageous alternative to Ethernet wireline data networks.

In this section, we study two type of data transfers: the file transfer via the File Transfer Protocol (FTP) and the World Wide Web (WWW) traffic. We provide some capacity results on the standard IEEE 802.11b DCF based on the Ready To Receive / Clear To Send (RTS/CTS) mechanism. The capacity is evaluated either in terms of maximum achievable throughput or in number of simultaneously active users.

2.2.1   Maximum Achievable TCP Throughput

The throughput achieved by a simplex communication between two nodes gives an upper bound on the capacity. The impact on throughput due to the distributed mechanism DCF is estimated and compared with the results obtained with the Point Coordination Function (PCF) centralized MAC protocol. A simple computation is seen to provide results that are also very close to the simulations results.

The channel is assumed to be error-free and the other parameters in this study are header length (long or short), the basic rate (1 or 2 Mbps), and the packet size at the application layer. Without contention, the mean back-off timer is $\lfloor CW_{min}/2 \rfloor$. With TCP and a communication between two nodes, this value is optimistic because the acknowledgment packets of the transport layer (denoted ACKTCP hereafter and that should be differentiated from the acknowledgment control packet of IEEE 802.11, denoted ACK) contend for the channel with data packets. This contention implies a higher back-off window in average for both nodes. Simulations show that a mean back-off timer of $\lfloor CW_{min}/2 \rfloor + 1$ can be chosen in this situation. Also, every TCP packet is assumed to be acknowledged by a single ACKTCP. Moreover, TCP packets are sent using the RTS/CTS handshake, while ACKTCP are not.

Now, by dividing the packet length by the duration of the frame sequence RTS + CTS + DATA + ACK + ACKTCP + ACK, including back-off and guard intervals, we obtain the throughputs of table 2.1 for two typical packet lengths. This is compared to simulation results (figures in brackets) obtained with the Network Simulator v.2 (ns2) [155]. It can be noticed that analytical results are close to the simulation results, even if sometimes the calculation provides an overestimate of throughput.

From these results, we can draw two important conclusions.

(i) The overhead introduced by the protocol stack and by the IEEE 802.11 DCF protocol is very high in terms of channel resources utilization: Even for the best case presented here (1024 byte packets, a short preamble, a physical mode at 11 Mbps and a basic rate of 2 Mbps), the user throughput is only 30% of the available physical data rate. This is mainly due to the exchange of control packets at a low basic rate (up to 11 times less than the data physical mode).

(ii) Achievable throughputs are very sensitive to the physical layer parameters. For example, 1.40 Mbps are achieved with 512 byte packets, a long preamble, a data rate of 11 Mbps, and a basic rate of 1 Mbps. If 1024 bytes are considered

Table 2.1: Maximum achievable TCP throughput of IEEE 802.11b in DCF mode.

Packet size/preamble/basic rate	Physical modes			
	1 Mbps	2 Mbps	5.5 Mbps	11 Mbps
512 bytes/short/2 Mbps	0.59 (0.57)	0.96 (0.93)	1.59 (1.53)	1.95 (1.87)
512 bytes/long/1 Mbps	0.53 (0.49)	0.80 (0.73)	1.20 (1.06)	1.40 (1.22)
1024 bytes/short/2 Mbps	0.74 (0.73)	1.30 (1.27)	2.46 (2.39)	3.31 (3.20)
1024 bytes/long/1 Mbps	0.69 (0.66)	1.15 (1.07)	1.97 (1.78)	2.48 (2.20)

with a short preamble and 2 Mbps of basic rate, the throughput reaches 2.48 Mbps. In practice, deployment engineers need to find the right trade-off between a wider covered area and a better channel utilization.

The distributed nature of DCF is a factor of capacity reduction. This can be seen if we look at the performance of IEEE 802.11 in PCF mode (table 2.2). The assumptions and the analytical model are detailed in appendix C. With the already cited example (1024 byte packets, short preamble, a physical mode of 11 Mbps, and a basic rate of 2 Mbps), the user throughput is now 55% of the channel capacity. Hence, in a fully centralized solution, PCF may be preferred for reasons of throughput.

Table 2.2: Maximum achievable TCP throughput of IEEE 802.11b in PCF mode.

Packet size/preamble/basic rate	Physical modes			
	1 Mbps	2 Mbps	5.5 Mbps	11 Mbps
512 bytes/short/2 Mbps	0.68 (0.67)	1.25 (1.24)	2.77 (2.75)	4.32 (4.31)
512 bytes/long/1 Mbps	0.63 (0.63)	1.12 (1.11)	2.22 (2.21)	3.15 (3.14)
1024 bytes/short/2 Mbps	0.80 (0.80)	1.53 (1.53)	3.63 (3.61)	6.09 (6.06)
1024 bytes/long/1 Mbps	0.76 (0.77)	1.41 (1.42)	3.10 (3.08)	4.79 (4.78)

The maximum achievable throughput presented in this section assume a perfect error-free channel. A more accurate model is described in the next section.

2.2.2 Channel Model and Link Adaptation Strategy

Four physical modes have been defined by the standard IEEE 802.11b, namely 1, 2, 5.5, and 11 Mbps. 1 and 2 Mbps belongs to the basic rate set, i.e., all terminals must be able to receive and transmit at these data rates. In our implementation, the RTS / CTS handshake is used for each packet. RTS, CTS, and ACK control packet are transmitted at the basic data rate of 1 Mbps. The physical mode of data packet is chosen according to the implemented link adaptation strategy. This mechanism is based on the channel quality. So, let us describe the considered channel model.

The path loss is given by the following formula:

$$L = 32.4 + 20 \log(f) + 10n \log(d) , \qquad (2.1)$$

where $f = 2.4$ GHz is the frequency in GHz, $n = 4$, and d is the distance in meters between the sender and the receiver. This model corresponds to an indoor propagation scenario. The transmit power is 15 dBm. The received power is computed from link budget calculations. A 4 dB standard deviation of log-normal shadowing is taken into account.

Bit errors are assumed to be independent, so that the Packet Error Rate PER can be easily deduced from the Bit Error Rate (BER) ($PER = 1 - (1 - BER)^N$, where N is the number of bits). Then, the BER is computed by the analytical formulas of DBPSK[1], DQPSK, and the MBOK (CCK is considered as a variation of MBOK) modulations over Additive White Gaussian Channels (AWGN) [163].

Link adaptation in our simulations is based on PER metric and carrier to noise ratio (C/N) switching thresholds. So, the decision for the physical mode takes into account both the received power C (C have to be above the sensitivity threshold of the mode) and C/N measurements (with PER constraint 0.1). More details on the link adaptation strategy can be found in appendix A. The TCP throughput from the AP to a single terminal in the cell is presented in figure 2.1 for 1024 byte packets and a window size of 64 segments of 1024 bytes. The computed range for an indoor environment is in accordance with experimental ranges observed in an office building (see [26] and [27]). We now focus on two specific TCP based applications: the FTP download of files and the WWW traffic.

---

[1]Differential-Binary Phase-Shift Keying (DBPSK), Differential-Quadrature Phase-Shift Keying (DQPSK), M-ary Bi-Orthogonal Keying (MBOK), Complementary Code Keying (CCK).
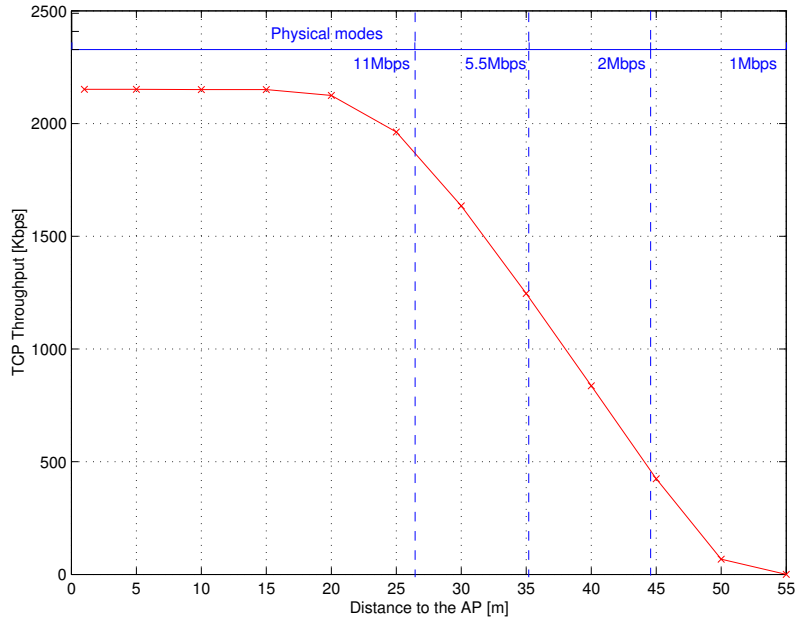
Figure 2.1: TCP throughput vs. distance to the AP, packet length = 1024 bytes.

2.2.3 FTP Download

We would like to know how many simultaneous FTP downloads are possible in a single cell and with which quality of service. The interesting metrics for this type of traffic are the throughput and the download time. These metrics are dependent on the distance of the mobile station to the AP. So, let us consider the FTP download of a file of 10 Mbytes (a relatively big file for the TCP case).
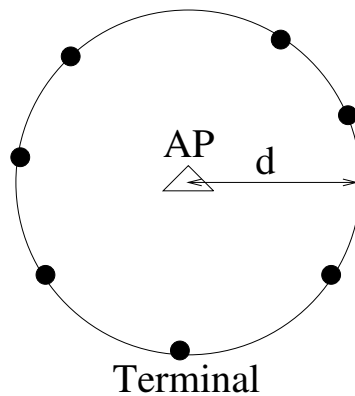

Figure 2.2: Network topology with seven terminals.

The topology for our simulations is shown in figure 2.2: A growing number of terminals are located at a fixed distance of the AP, i.e., the average received power

is the same. The simulated TCP version is the traditional TCP Tahoe with a window size of 64 segments of 1024 bytes. We have also assumed a network delay of 50 ms. The link adaptation algorithm is used (see appendix A) and we assume long preambles and a basic rate of 1 Mbps.
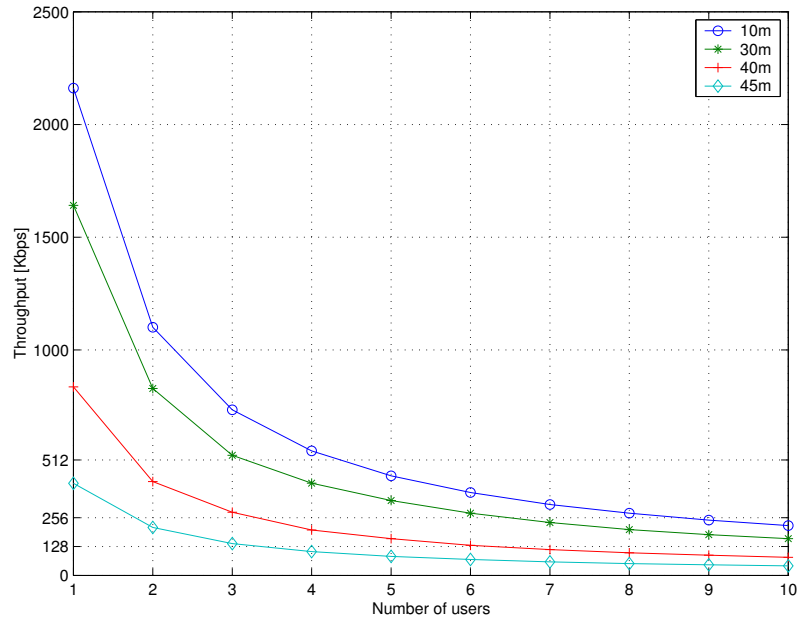

Figure 2.3: TCP throughput vs. number of users and distance to the AP, download of a 10 Mbyte file.

In figure 2.3, the throughput is given as a function of the number of users for several distances. Time transfers of the file are provided in figure 2.4. If a single user is downloading at 10 m from the AP, its throughput exceeds 2100 Kbps and only 37 s are needed for a 10 Mbyte file. If a throughput of 512 Kbps is required, the capacity of the cell is 4 users at 10 m, but only 1 user at 40 m can be allowed. Such a quality of service is not achievable at 45 m. If only 256 Kbps are required, the capacity becomes 8 users at 10 m, 3 users at 40 m, and 1 user at 45 m.

2.2.4 WWW Traffic

We look now for the capacity in terms of number of WWW users. Let us first describe the ETSI model of WWW browsing implemented in ns2 [173, 191].
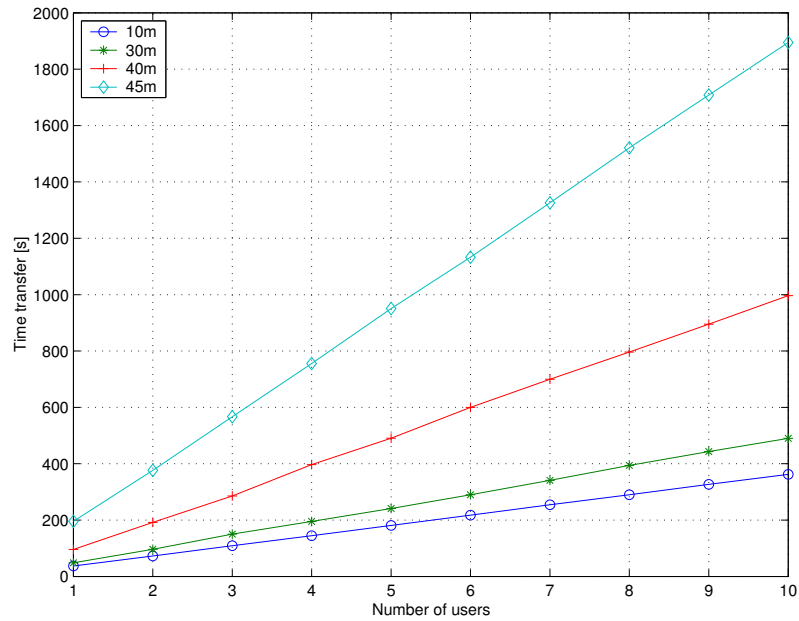
Figure 2.4: Time transfer vs. number of users and distance to the AP, download of a 10 Mbyte file.
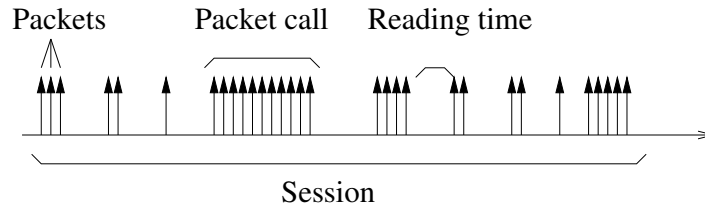

Figure 2.5: ETSI model for WWW traffic.

ETSI Model

This model is applied from the network to the terminal. Note that the uplink traffic that consists mainly in requests is neglected. The ETSI model, proposed by 3GPP, is characterized by three levels of communication: session, packet call, and packet transmission (see figure 2.5).

A session is a period of activity of the user, during which several WWW documents are downloaded. The transfer of a WWW document is modeled by a packet call, which in turn consists in a bursty sequence of packets. After the document is entirely arrived to the terminal, the user needs some time, the reading time, to study it.

The session arrival times is modeled by a Poisson process. The number of packet

calls per session is assumed to be geometrically distributed. The reading time follows an exponential distribution. The number of packets in a packet call is a geometric random variable. The packet size is modeled by a random variable with Pareto distribution and cut-off value (see [191]). However, at the transport layer, packets are concatenated in 1024 byte frames [173]. The default values used in our simulations are given in table 2.3. Note that an extra delay of 50 ms is assumed in the wired network.

Table 2.3: Default values for the ETSI model.

Parameter	Value
Mean session time	1800 s
Mean number of packet calls per session	5
Mean reading time	90 s
Mean number of packets per packet call	25
Mean packet size	900 bytes
Mean page size	22.5 Kbytes

WWW Capacity

In this section, we are interested in the number of possible simultaneous WWW users. Thus, we do not take into account the session level: All users are assumed to be active. In figure 2.6, the WWW page throughput has been depicted. This measure is defined as the throughput observed between the first and the last packet of a packet call. We chose this metric in order to evaluate the user perceived throughput. Figure 2.7 shows the time transfer of a WWW page. This is also a good metric for the perceived quality of service.

Note that we didn't let the number of WWW users grow indefinitely to know the limit of the WWW capacity. High numbers of users is indeed unrealistic for the considered cell areas. This is particularly true at 10 and 20 m, where page throughputs above 380 Kbps are reached even with 50 users. In these cases, the page download time is less than 1 s. If we require a throughput of 256 Kbps, 30 users can be aggregated at 40 m. Even with 50 users at 45 m, a page is retrieved in average in less than 3 s. Only at 50 m, the quality of service becomes unacceptable.
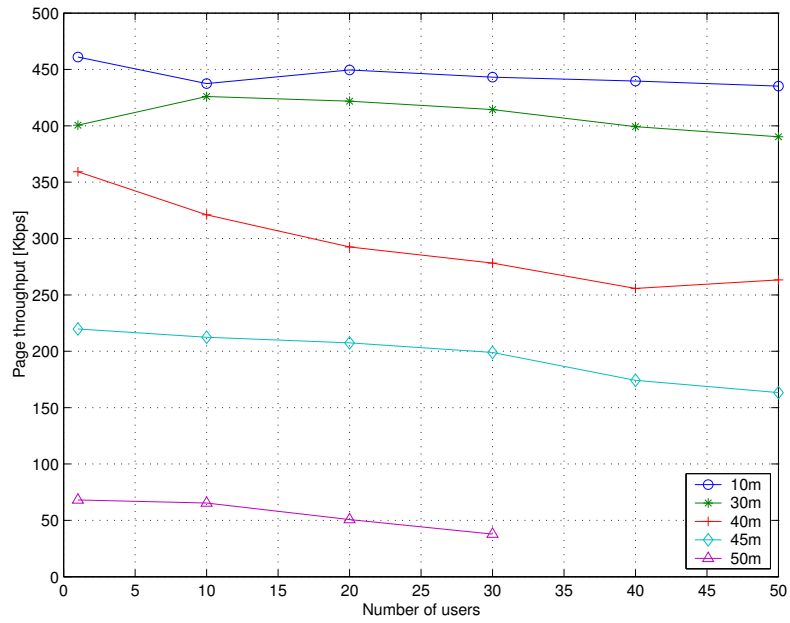
46

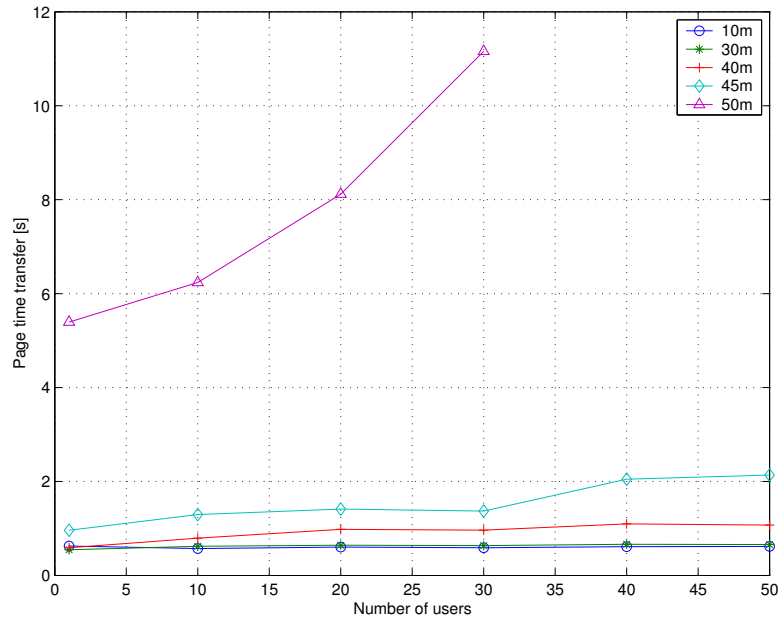Figure 2.6: WWW page throughput vs. number of users and distance to the AP.

Figure 2.7: Page time transfer vs. number of users and distance to the AP.

2.3 Voice over WLAN Capacity

This section provides the VoIP capacity of IEEE 802.11b, i.e., the maximum number of voice calls that can simultaneously take place in a WLAN cell. Providing such a real-time application over a CSMA/CA based network, originally designed for best-effort traffic is a challenging issue. Numerous papers try to address this problem. [194] evaluates the capacity with the Point Coordination Function of the standard. This feature is however optional and most of the card manufacturers do not implement it in commercial products. Instead, the Distributed Coordination Function is always used in practice. Some papers, e.g. [178] or [142], propose adaptations of DCF to allow voice traffic. [162] compares DCF, PCF, priority queuing and blackburst mechanisms.

Several experimental results based on IEEE 802.11b DCF are also available in the literature, e.g. [43] [40] or [97]. [96] provides an analysis of the number of VoIP calls for different codecs and data rates. However, none of these studies bases its conclusions on an efficient model for voice quality. All cited papers use quality metrics like packet loss and the average packet delay.

In this section, conclusions are based on the E-model which is an efficient tool to predict the voice quality (section 2.3.2). The system description detailed in section 2.3.3 includes codecs, the traffic model, the MAC protocol, the channel model, and the link adaptation strategy. Section 2.3.4 provides simulation results on capacity as a function of the chosen codec and the distance of the users to the AP. The influence of network delay and of the dejittering buffer are also studied.

Since most of the real-time applications are built over the User Datagram Protocol (UDP), it is of interest to evaluate achievable throughput for this protocol.

2.3.1 Maximum Achievable UDP Throughput

As for TCP, a simplex communication between two nodes provides an upper bound on the achievable throughput. Contrary to TCP, with UDP, the back-off timer is assumed to be equal to $\lfloor CW_{min}/2 \rfloor$. This is true in average since the whole traffic is coming from a single node and thus does not contend with backward traffic. Throughput results of table 2.4 are obtained by dividing packet length (two cases) by the duration of the frame exchange sequence RTS + CTS + DATA + ACK. Analytical values can be compared to simulation results evaluated with ns2 [155] (in

brackets).

Table 2.4: Maximum achievable UDP throughput of IEEE 802.11b in DCF mode.

	Physical modes			
Packet size/preamble/basic rate	1 Mbps	2 Mbps	5.5 Mbps	11 Mbps
512B/short/2 Mbps	0.74 (0.74)	1.26 (1.26)	2.27 (2.29)	2.96 (2.73)
512B/long/1 Mbps	0.67 (0.67)	1.07 (1.07)	1.72 (1.73)	2.09 (2.05)
1024B/short/2 Mbps	0.85 (0.85)	1.54 (1.54)	3.22 (3.23)	4.67 (4.69)
1024B/long/1 Mbps	0.80 (0.80)	1.39 (1.39)	2.62 (2.63)	3.51 (3.52)

In this ideal case of error-free channel, the proportion of protocol overheads is smaller than with TCP. With the same example as before (1024 byte packets, a short preamble, a physical mode of 11 Mbps, and a basic rate of 1 Mbps), the user throughput represents now 42% of the channel data rate. With PCF, this value is increased to 66% (see table 2.5 and appendix C for more details).

Table 2.5: Maximum achievable UDP throughput of IEEE 802.11b in PCF mode.

	Physical modes			
Packet size/preamble/basic rate	1 Mbps	2 Mbps	5.5 Mbps	11 Mbps
512 bytes/short/2 Mbps	0.80 (0.81)	1.51 (1.52)	3.47 (3.50)	5.60 (5.65)
512 bytes/long/1 Mbps	0.76 (0.77)	1.37 (1.39)	2.85 (2.90)	4.25 (4.28)
1024 bytes/short/2 Mbps	0.89 (0.89)	1.70 (1.72)	4.19 (4.24)	7.27 (7.34)
1024 bytes/long/1 Mbps	0.85 (0.87)	1.61 (1.63)	3.68 (3.74)	5.96 (6.04)

2.3.2   E-model for Speech Quality Evaluation

We propose the use of the E-model recommended by the ITU. This is considered to be an adequate model for evaluating voice over IP networks.

Description

The E-model is a tool to predict how an "average user" would rate the voice quality of a phone call [197]. This model has been standardized by the ITU [92][93] and provides a R-scale. The rating factor R is composed of several additive terms, each one representing a specific source of voice quality degradation: $R = R_0 - I_s - I_d - I_e + A$.

R_0 is usually set to 94.3 and represents the basic signal-to-noise ratio. I_s represents impairments simultaneously occurring with the voice signal (e.g. quantization). I_d represents impairments due to transmission delays. I_e represents impairments caused by the use of a specific equipment, e.g., I_e depends on the selected codec and on packet loss. A is the expectation factor, it represents the degradation that a user is likely to accept because he is aware that the technology is wireless and mobile.

The range of R is from the worst quality, 0, to the best one, 100. The quality classes are shown in table 2.6. Note that the Public Switched Telephone Network quality falls in the range $70 - 100$, so that $R = 70$ will be our cut-off value for the capacity evaluation.

Table 2.6: Quality classes according to the E-model.

R range	90-100	80-90	70-80	60-70	0-60
Quality	best	high	medium	low	poor

The fine-tuning of the parameters of the E-model is important to get accurate results. For our simulations, we chose the values given in table 2.7. These are the default values given by the ITU standards. Note that in VoIP, no wireline type echo has to be considered.

Mouth-to-ear Delay Budget

One of the main source of quality degradation is the mouth-to-ear delay. In this section, details of this delay are given. The main sources of delays are the following:

- The packetization time T_{pack}, i.e., the time needed to collect all voice samples that form a packet. In our simulations, each voice frame is packetized in a single IP packet, so that $T_{pack} = T_F$, where T_F is the voice frame duration.

Table 2.7: E-model parameters values.

SLR	8 dB
RLR	2 dB
LSTR	18 dB
STMR	15 dB
Ds	3 dB
Dr	3 dB
TELR	65 dB
WEPL	110 dB
qdu	1
Nc	-70 dBm0p
Nfor	-64 dBmp
Pr	3 dB(A)
Ps	35 dB(A)
A	5
T_a	T
T_r	$2T$

- The voice encoding and decoding process T_{DSP}, i.e., the time needed to encode the analog voice source or to decode the voice samples to an analog signal. According to [100], $T_{DSP} = 12,5$ ms for the codec G729. We will assume that this value is the same for other considered codecs.

- The look ahead delay T_{LA} if any. Some codecs need indeed to collect a few samples before producing a voice frame [120].

- The network delay T_{nw}, i.e., the delay caused by the transmissions in the wired network and by the different buffers in IP routers. In our simulations, two fixed values have been chosen: $T_{nw} = 20$ ms and $T_{nw} = 100$ ms. These values are in accordance with those given in [116] and [119], and are expected to represent two extreme cases.

- The access delay T_{WLAN}, i.e., the queuing delay in the AP and the delay added by the MAC layer of IEEE 802.11. This delay is evaluated through ns2 simulations.

- The dejittering delay T_{jitt}, i.e., the delay introduced by the dejittering buffer at the receiver side. The computation of T_{jitt} is detailed in the next section.

Hence, the overall mouth-to-ear delay is given by:

$$T = T_{pack} + T_{DSP} + T_{LA} + T_{nw} + T_{WLAN} + T_{jitt} \; . \tag{2.2}$$

Dejittering Mechanisms

In principle, the receiver of a voice call could play out the first packet as soon as it arrives in its reception buffer. Then, following packets have to be played out at regular intervals in order to reproduce the streamed information. In practice, packet based networks introduce transport delay variations (or jitter), so the too delayed packets could be lost for the application, even though they have been correctly received by the receiver.

That is the reason why the receiver has a dejittering buffer that retains fast packets until they have to be played out. The buffer "absorbs" the delay variations.

Let $s_n = nT_s$ be the sending instant of the n-th packet, where T_s is the sending period of the voice frames. Let d_n be its delay and $a_n = s_n + d_n$ its arrival instant at the receiver. We assume that the packets arrive in correct order. Network routes are considered stable. Moreover, on the radio link, the "stop-and-wait" acknowledgment policy ensures the correct order.

The dejittering buffer retains the first packet for a time D. Then the buffer is read at regular intervals, i.e., $a_0 + D + nT_s$. If the n-th packet is present in the buffer at this time, it is played out. Otherwise, i.e., if it is too late, the packet is lost. This occurs if:

$$
\begin{aligned}
a_n &> a_0 + D + nT_s & (2.3)\\
nT_s + d_n &> s_0 + d_0 + D + nT_s & (2.4)\\
d_n &> d_0 + D & (2.5)\\
D &< d_n - d_0 & (2.6)
\end{aligned}
$$

To avoid any packet loss, the dejittering delay has to be chosen as follows:

$$T_{jitt} = d_{max} - d_0 \; , \tag{2.7}$$

where d_{max} is the maximum delay. However, the voice traffic can tolerate some packet loss without a big degradation of the quality. If P_{loss} is tolerated in the

dejittering buffer, T_{jitt} is now:

$$T_{jitt} = d_q(P_{loss}) - d_0 \ , \tag{2.8}$$

where $d_q(P_{loss})$ is the $(1 - P_{loss})$-quantile of the delay. At this point, a trade-off has to be found because increasing P_{loss} reduces T_{jitt} and so the mouth-to-ear delay.

In practice, the receiver doesn't have the probability density function (pdf) of the delay. Several adaptive algorithms are presented in [193]. In this study, a perfect mechanism is assumed to match equation 2.8.

Packet Losses

There are two main sources of packet loss. The first one is due to the MAC layer. If a packet is not correctly received because of the channel conditions or because of a collision, the MAC layer retransmits the lost packet. After four unsuccessful retransmissions of a RTS or after seven unsuccessful retransmissions of a data packet, the packet is definitely dropped. The proportion of such packets is P_{loss}^{MAC}.

The second reason is due to the dejittering mechanism implemented at the receiver. The proportion of such packets is P_{loss}, as discussed in the previous section.

A third reason of packet loss could be the congestion in one of the nodes of the wired network, including the AP. However, it is not taken into account in this study.

2.3.3 System Description

The proposed system characteristics include the network topology, the codecs, the traffic model, the channel model, and the link adaptation strategy.

Since the IEEE 802.11b includes a rate (or link) adaptation mechanism able to switch between the physical modes 1, 2, 5.5, and 11 Mbps, the capacity of the cell depends on the spatial distribution of users. All users are assumed to be at an equal distance to the AP, as shown in figure 2.2 (seven terminals at a distance d from the AP).

Three codecs are considered: GSM-Enhanced Full Rate (EFR), G711, and G723.1. Their main characteristics are summarized in table 2.8. Note that the value I_e assumes no packet loss transmission. The dependence of I_e with packet loss is provided by [93].

Table 2.8: Codecs main characteristics.

	GSM-EFR	G711	G723.1
Bit rate [Kbps]	12.2	64	6.3
Packet size [bits]	244	640	189
Frame duration [ms]	20	10	30
Look ahead [ms]	0	0	7.5
I_e	5	0	15

Packet Loss Concealment (PLC) mechanism of G711 is useful to improve robustness against packet loss. The study includes the two cases of using PLC or not with bursty packet loss.

Voice frames are sent in a Real Time Protocol (RTP)/UDP/IP packet. For these protocols, there is an overhead of $20 + 8 + 12 = 40$ bytes. Note that the physical header of IEEE 802.11b (with long preamble) adds 24 bytes, and the MAC header adds 34 bytes.

The voice traffic is modeled by a ON/OFF source in each direction. The mean ON period duration is 1.0 s, and the mean OFF period duration 1.35 s. Both follow an exponential distribution. Thus, the voice activity is 42.6%. These values are in accordance with [158].
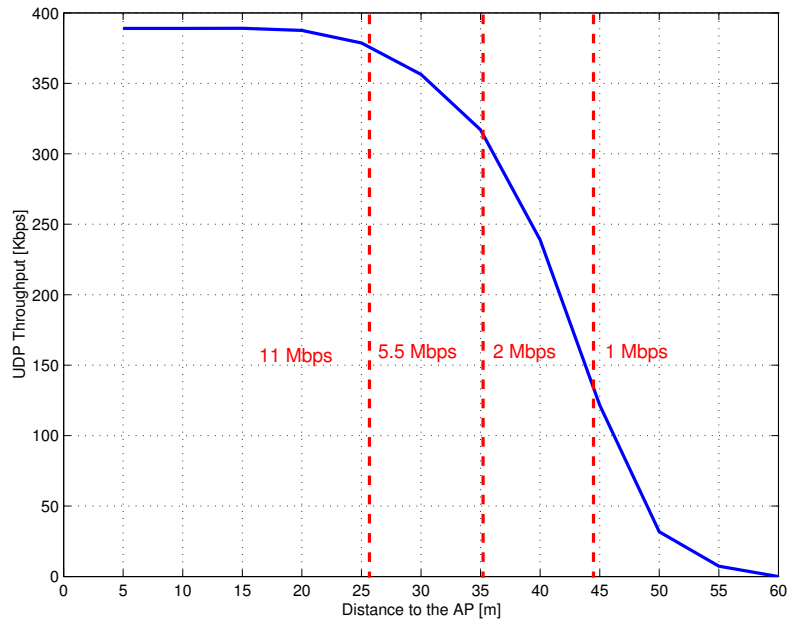

Figure 2.8: UDP throughput vs. distance to the AP, packet load = 80 bytes.

The channel model and the link adaptation strategy have been described in section 2.2.2. With these assumptions, the UDP throughput from the AP to a single terminal in the cell is presented in figure 2.8 for 80 byte packets. The ranges associated to the physical modes based on sensitivity thresholds are also shown (see appendix A for further details).

2.3.4 Simulation Results

In this section, the aforementioned system is simulated using the Network Simulator v.2 ns2 [155]. Capacity values are deduced from the E-model for different codecs, distances to the AP, and P_{loss} values. The simulated time is 200 s.

Influence of Distance

As the distance from the AP to the terminals increases, the link adaptation mechanism selects slower rate physical modes. As a consequence, the available throughput above the MAC layer is reduced and less simultaneous voice calls are possible. In this section, $T_{nw} = 100$ ms is assumed.
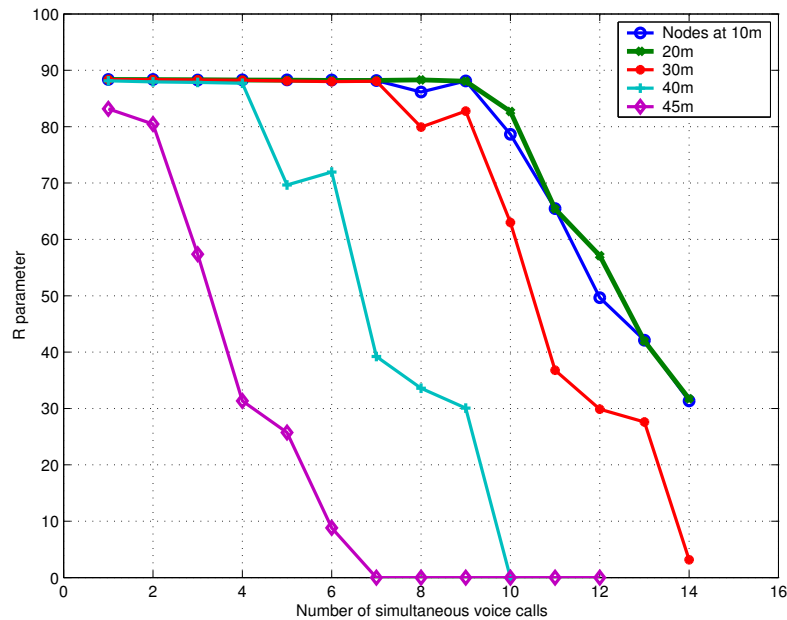

Figure 2.9: R parameter vs. number of simultaneous voice calls, influence of the distance to the AP, GSM-EFR.
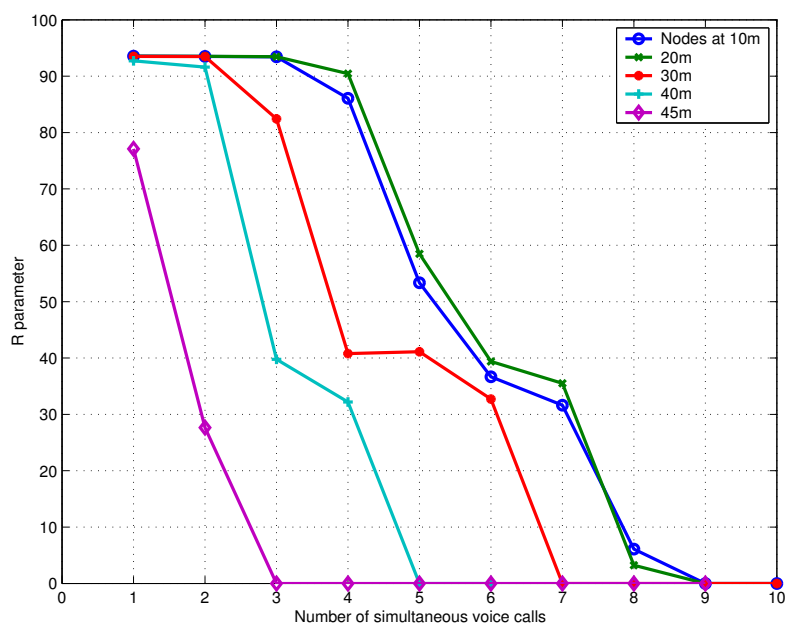
Figure 2.10: R parameter vs. number of simultaneous voice calls, influence of the distance to the AP, G711.

The performance of the codecs GSM-EFR, G711, G711 with PLC, and G723.1 are shown respectively on figures 2.9, 2.10, 2.11, and 2.12 as a function of the distance to the AP. Note that in this case $P_{loss} = 0\%$ (see equation 2.8). This explains the small difference between G711 and G711 PLC.

For the three codecs, there is a significant degradation of capacity with distance. With GSM-EFR, increasing the AP/terminal distance from 10 m to 45 m reduces the capacity from 10 voice calls to 2. With G711, while 4 simultaneous calls are possible at 10 m, only one call can be made at 45 m. The biggest degradation can be seen with G723.1, from 17 calls at 10 m down to 4 at 45 m.

These values can be surprising with respect to the available physical data rate in the cell, especially at 10 m, where the physical mode is 11 Mbps. In fact, IEEE 802.11b suffers from a huge overhead, due to the RTS/CTS handshake, the acknowledgment, the MAC header, the back-off window, and the basic rate of 1 Mbps used to transmit the control packets and the physical header. Moreover, for each voice frame, a RTP/UDP/IP header has to be added. The proportion of this overhead is particularly high for small data packets.

The overhead budget for the transmission of a small packet of payload 80 bytes at 11 Mbps is given in table 2.9. Note that the data part as well as the RTP/UDP/IP headers are sent at 11 Mbps. The back-off has been set to 15 SlotTime.
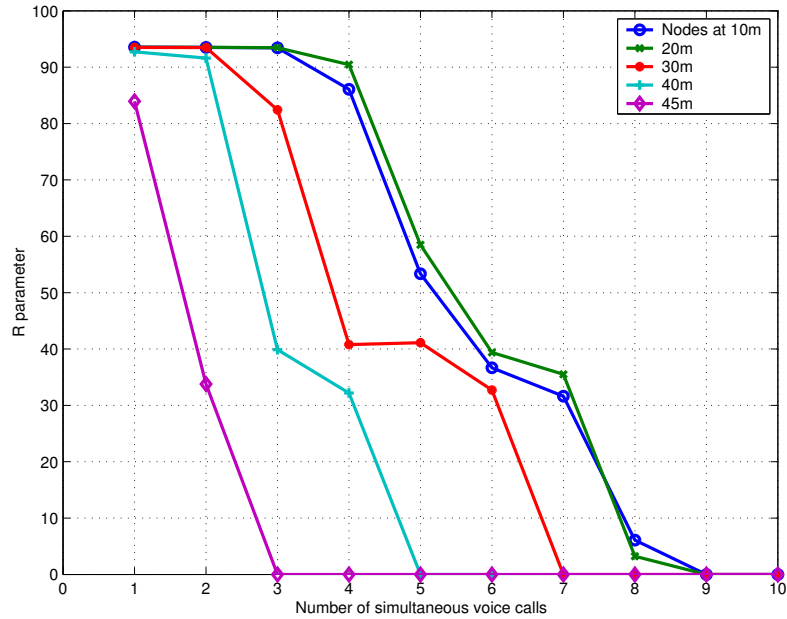
Figure 2.11: R parameter vs. number of simultaneous voice calls, influence of the distance to the AP, G711 with PLC.

Since 1644.0 μs are needed for the transmission of 80 bytes, the user throughput is approximately 389 Kbps, which is in accordance with figure 2.8. Now, a G711 call needs a bandwidth of 54.5 Kbps if we take into account both directions and voice activity. Thus, an upper bound for the capacity is $\lfloor 389/54.5 \rfloor = 7$. However, transmitting voice over IP is not only a question of bandwidth. Simulation results show that voice quality requirement of $R = 70$ reduces to the capacity to 4 calls.

Influence of Packet Loss in the Dejittering Buffer

The influence of P_{loss} is now studied ($T_{nw} = 100$ ms). On the one hand, increasing P_{loss} increases also the I_e parameter in the computation of R, on the other hand the delay added by the dejittering buffer is reduced. Figures 2.13, 2.14, 2.15, and 2.16 illustrate this trade-off for respectively GSM-EFR, G711, G711 with PLC, and G723. Terminals are at 10 m of the AP. Similar results were obtained for other distances.

With GSM-EFR, one additional voice call is possible for 10 m, if $P_{loss} = 2\%$ is allowed on the dejittering buffer. Although G711 can take advantage of the PLC in terms of voice quality, yet the maximum number of calls is limited to 4. Without PLC, G711 is very sensitive to packet losses: With $P_{loss} = 1\%$, the voice quality is
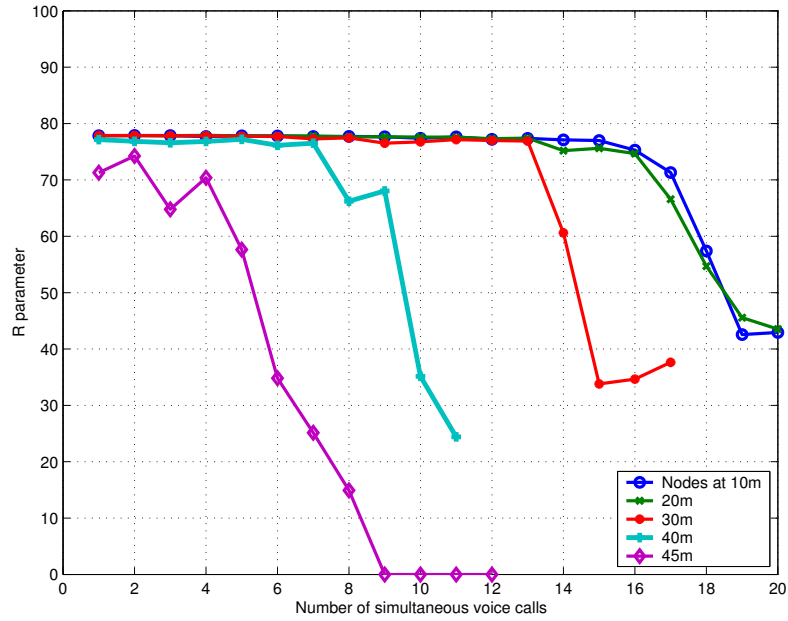
Figure 2.12: R parameter vs. number of simultaneous voice calls, influence of the distance to the AP, G723.1.

already below our requirement. Letting P_{loss} be 1% or 2% allows to have 18 G723.1 calls instead of 17.

Influence of the Network Delay

Now, the influence of the network delay is studied by comparing the achievable capacity for its two extreme values of $T_{nw} = 20$ ms and $T_{nw} = 100$ ms.

The R parameter is given in figure 2.17 for the two considered network delays and for terminals at 10 m from the AP. G711 with PLC performance is very similar to that of G711 for $P_{loss} = 0\%$. Reducing the network delay allows to obtain one additional voice call for GSM-EFR. There is, however, no gain for G711 and G723.

Figure 2.18 now shows the performance of the codecs with the best choice of P_{loss} among 0, 1, 2, 3, 4, and 5%. In this case, one more call can be obtained with G711 if PLC is used and $P_{loss} = 3\%$.

Thus, even huge reduction in network delay does not bring much increase in capacity. How can this phenomenon be explained? The explanation consists in the shape of the delay curves (figure 2.19). For all codecs, packet delays grow slowly with the number of voice calls. Even with $T_{nw} = 100$ ms, the mouth-to-ear delay is very low, so that the voice quality is not impacted a lot by the network delay. As

Table 2.9: Overhead budget for a packet with a payload of 80 bytes, a basic rate of 1 Mbps, and a data rate of 11 Mbps.

	Transfer time [μs]	Percentage
RTS (1 Mbps)	352	21.4
CTS (1 Mbps)	304	18.5
ACK (1 Mbps)	304	18.5
PLCP Header (1 Mbps)	192	11.7
3xSifs	30	1.8
Difs	50	3.0
Back-off	300	18.3
Data	58.2	3.5
MAC header	24.7	1.5
RTP/UDP/IP header	29.1	1.8
Total	1644.0	100

the number of voice calls reaches the capacity level, packet delays increase sharply: The voice quality is highly degraded. Hence, the network delay plays a role only around the capacity limit.

As a conclusion for this section, we recall the capacity results. Simulation results provide the following maximum achievable simultaneous voice calls: 5 for G711, 12 for GSM-EFR, and 18 for G723.1. Further work could include the fine-tuning of the MAC parameters. For example, the RTS/CTS handshake may not be needed on the downlink. Or the number of retransmissions could be reduced. Header compression could be also considered. Additionally, the concatenation of voice frames has to be investigated and could have a deciding role for reducing the MAC overhead.
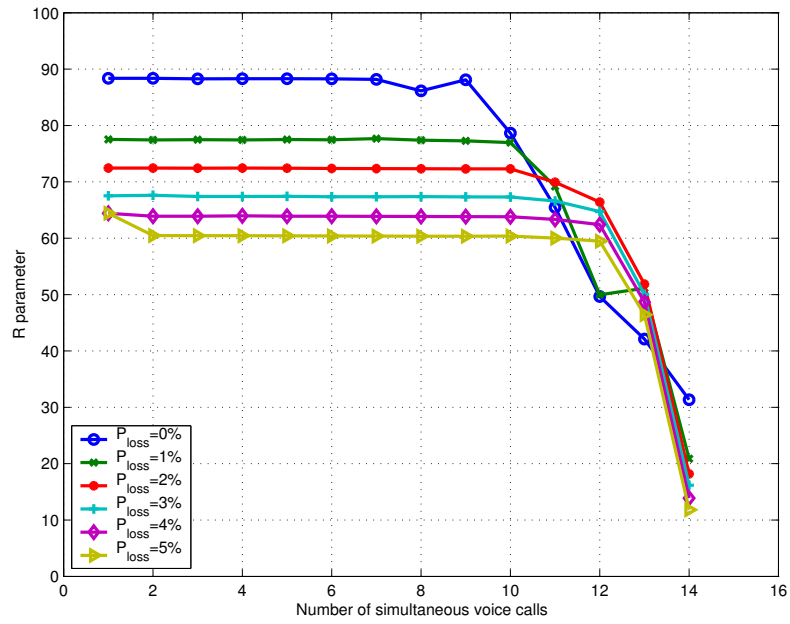
Figure 2.13: R parameter vs. number of simultaneous voice calls, influence of P_{loss}, GSM-EFR.
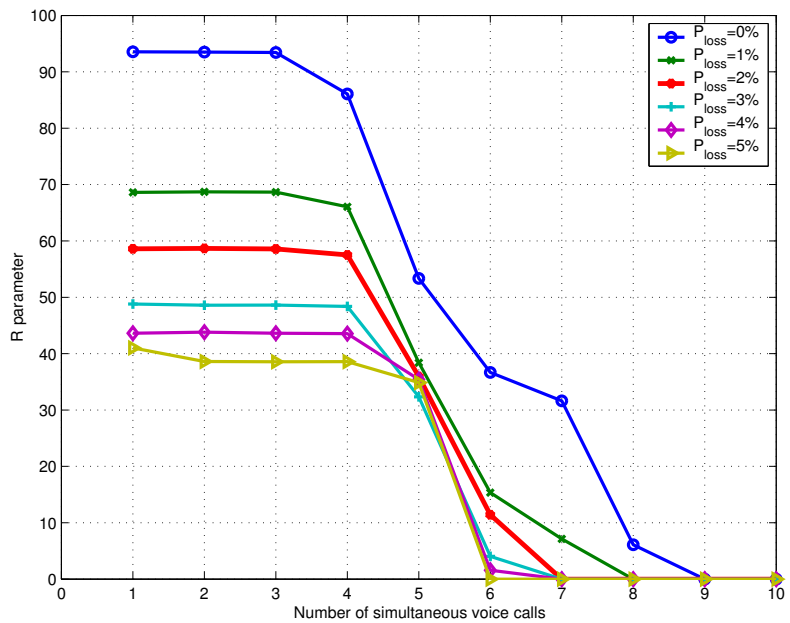

Figure 2.14: R parameter vs. number of simultaneous voice calls, influence of P_{loss}, G711.
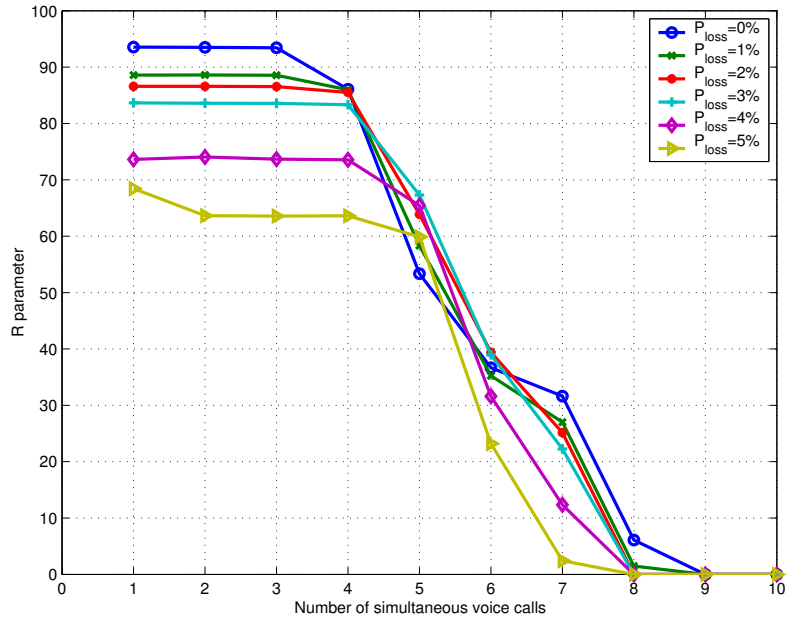
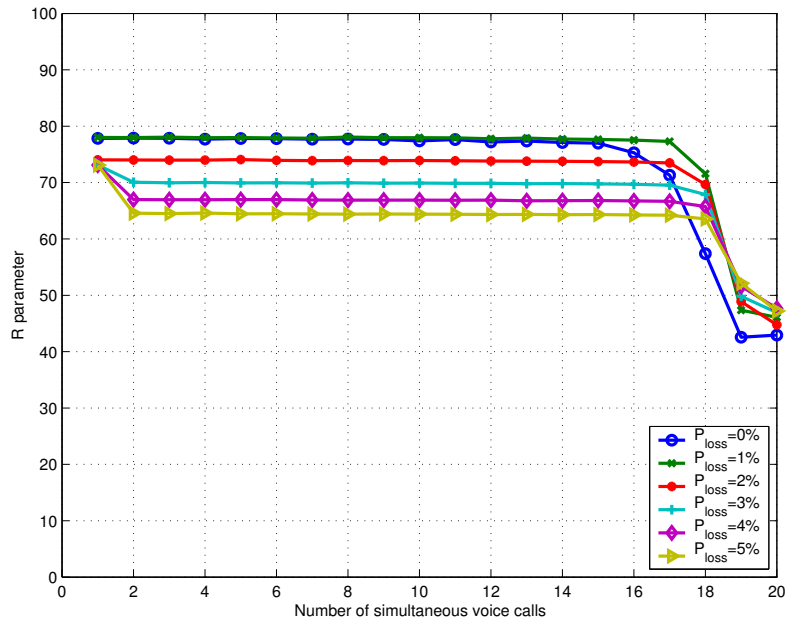Figure 2.15: R parameter vs. number of simultaneous voice calls, influence of P_{loss}, G711 with PLC.

Figure 2.16: R parameter vs. number of simultaneous voice calls, influence of P_{loss}, G723.1.
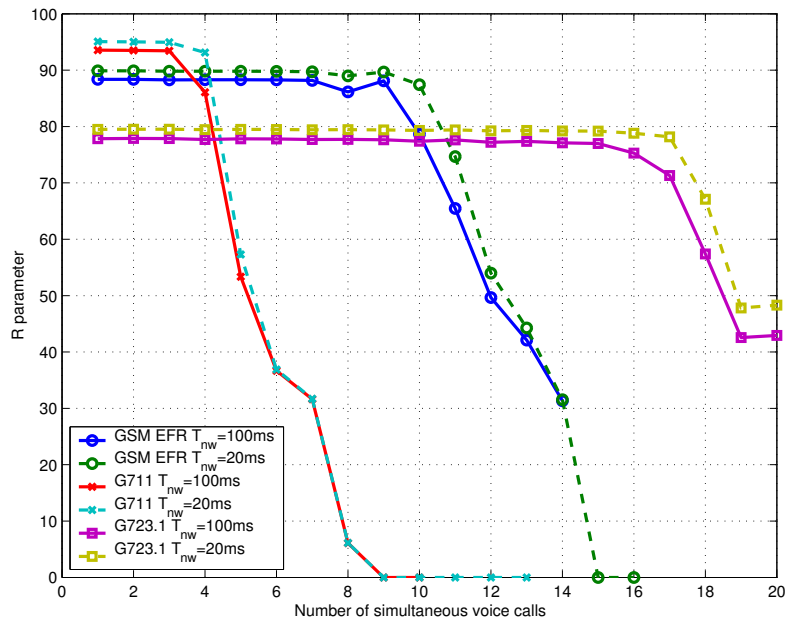
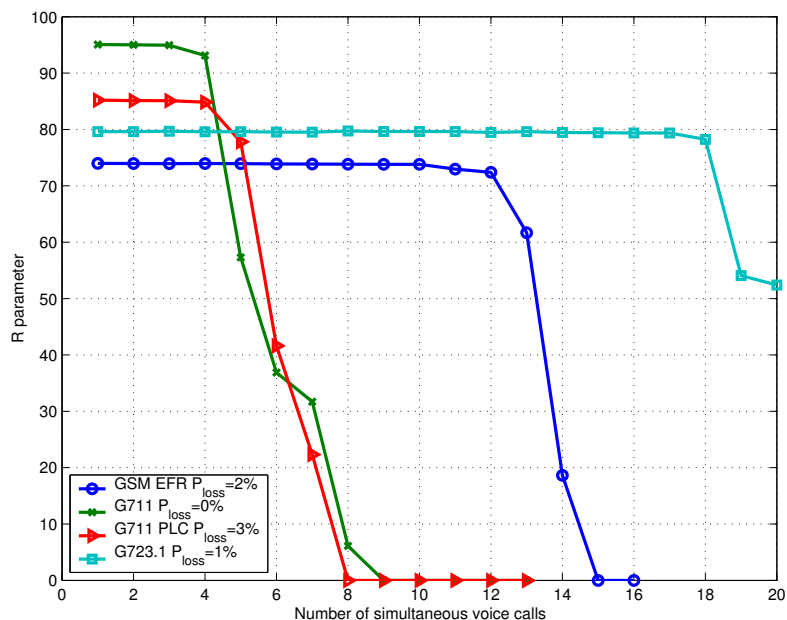Figure 2.17: R parameter vs. number of simultaneous voice calls at 10 m, influence of T_{nw}.

Figure 2.18: R parameter vs. number of simultaneous voice calls at 10 m, $T_{nw} = 20$ ms.
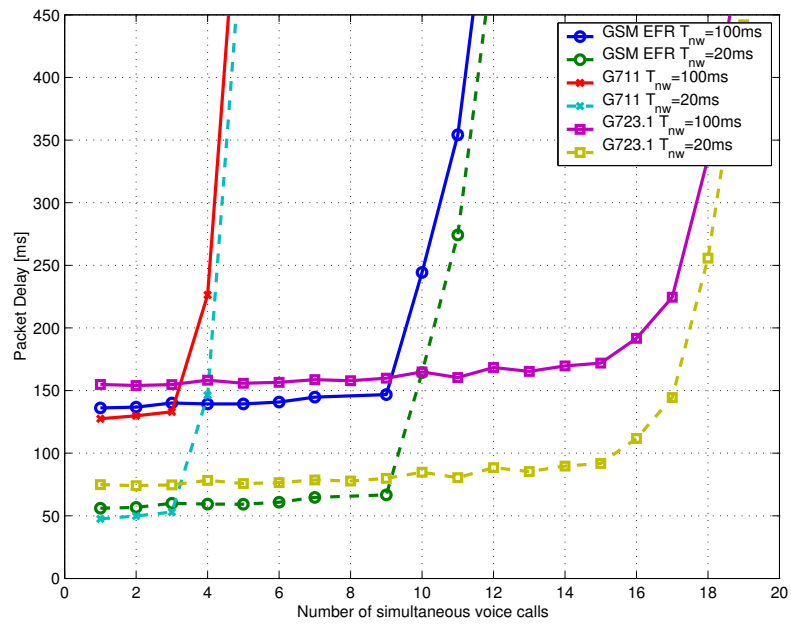
Figure 2.19: Packet delay vs. number of simultaneous voice calls at 10 m, influence of T_{nw}.

2.4 A Case Study: High Speed Internet for Rural Areas

In this section, we provide a unique commercial application example of IEEE 802.11 usage. The basic idea of the High Speed Internet (HSI) project is to provide high data rate Internet connection to the Internet to rural areas. In these areas, providing an Asymmetric Digital Subscriber Line (ADSL) is often too costly because of the small density of population or because of the long distance from a Digital Subscriber Line Access Multiplexer (DSLAM, a central office multiplexer). In this context, a telecommunication manufacturer can propose an end-to-end solution based on the coupling of a satellite connection to the Internet with a low cost WLAN network for end-user connectivity (figure 2.20). The aim of the presented study is to dimension
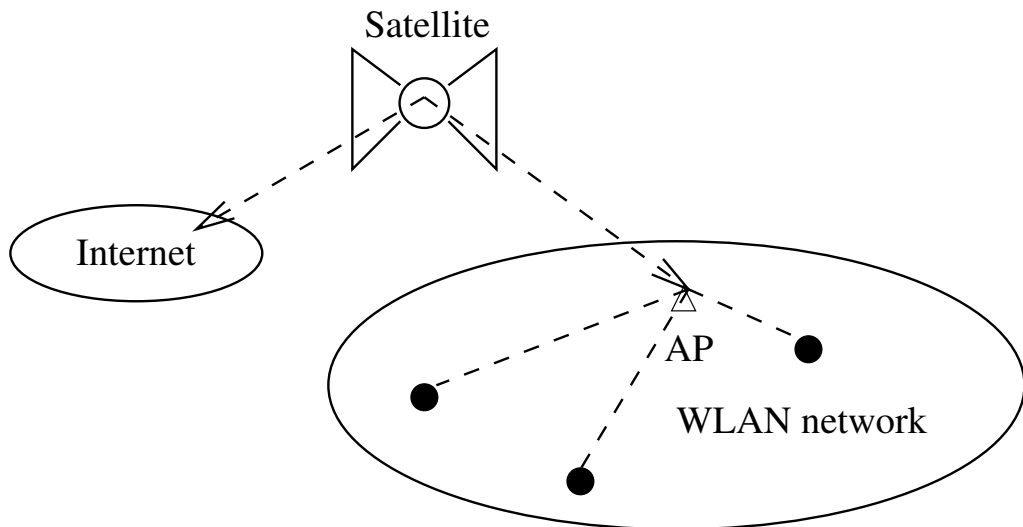

Figure 2.20: High Speed Internet for rural areas.

the satellite link according to the number of users. Specific requirements for radio propagation model, data throughput and backbone bandwidth need to be considered here.

2.4.1 Models

Satellite Link

The satellite reverse link (from the AP to the Internet via the satellite) and forward link (from the Internet to the AP) are modeled by error-free, fixed bandwidth and

fixed delay propagation media.

The considered forward and reverse data rates are the following in Kbps: $(128, 64)$, $(512, 128)$, $(1000, 256)$, $(1000, 384)$, $(2000, 512)$, $(2000, 2000)$, and $(48000, 384)$. Those cases are compared to the ideal case of zero network delay.

The packet queue in the forward direction is considered to be infinite, while the reverse link buffer is limited to 1000 packets of 1024 bytes. Similarly, the AP packet queue is restricted to 1000 packets of 1024 bytes.

The propagation delay from the network to the AP has been chosen to be 250 ms and includes the extra delay due to the satellite link. Obviously, the reverse link is also characterized by an extra delay of 250 ms.

Path Loss Model

The selected path loss model is the free space model with additional loss due to foliage [61]. With this model, the path loss L is given by $L = 32.4 + 20\log(f) + 20\log(d)$, where f is the frequency in MHz and d is the distance between the transmitter and the receiver in km. Note that the path loss exponent is $n = 2$. To this loss, an attenuation due to vegetation ($AttVeg$) is added. Random shadowing with standard deviation of 3 dB is considered. $AttVeg = 0.187 f^{0.284}(dx)^{0.588}$, where x is the vegetation density. In our simulations, $x = 1/500$. All antennas are assumed to be omni-directional with a gain in transmission $G_t = 4$ dBi and in reception $G_r = 9.8$ dBi. Finally, the transmit power of AP and end-user equipments is 16 dBm.

Traffic Models and Spatial Distribution of Users

Three types of traffic are considered: video streaming, ETSI WWW traffic, and e-mail traffic. Video streaming is a downlink traffic using UDP as transport layer. Application packets are concatenated in 1024 byte UDP packets[2]. The trace of a H.263 movie with average data rate 256 Kbps has been used.

The ETSI model has been already described in section 2.2.4 and is used for the WWW traffic. Parameters of the model are given in table 2.3. Contrary to section 2.2.4, the session level of the model is taken into account. WWW traffic is a downlink traffic based on TCP. During the simulations, some e-mails are also

---

[2]A more realistic model with variable UDP packets is for further studies

sent using FTP uplink transmissions. Small e-mails are made of 3 Kbyte files, while e-mails with attachments are 300 Kbytes long. The assumed version of TCP is TCP Tahoe with an advertized window of 64 segments of 1024 bytes.

End-user receivers are randomly distributed in a disk of radius 4 km. 50% of the users are in a disk of radius 2 km. The link adaptation mechanism described in appendix A provides 11 Mbps links up to approximately 2.2 km, 5.5 Mbps up to 3.6 km, and 2 Mbps between 3.6 km and 4 km. Three cases are studied: twenty, sixty, and hundred fixed users.

Three specific scenarios are considered. The number of subscribers per traffic type is given in table 2.10. Note that video streaming receivers are always located in the 11 Mbps area.

Table 2.10: Traffic scenarios for High Speed Internet.

	Twenty users	Sixty users	Hundred users
Video streaming	1	3	5
ETSI WWW traffic	14	42	70
E-mail (3 Kbps)	4	14	24
E-mail (300 Kbps)	1	1	1

2.4.2   Outdoor Coverage with IEEE 802.11

The standard IEEE 802.11 was designed for use of WLAN only in indoor environments. In this context, the propagation is typically less than 1 μs. For our study, distances up to 4 km have to be considered which lead to a value of delay over the air interface up to 13 μs.

The standard defines a parameter, $aAirPropagationTime$, that has an influence on the definition of $aSlotTime$. The value of $aSlotTime$ is set by the standard for the different physical layers, e.g., 20 μs for the Direct-Sequence Spread Spectrum (DSSS) physical layer of IEEE 802.11b. As a consequence $aAirPropagationTime$ is not adaptable.

Our simulations show that the values of the standard can be adopted for our study provided that two timers, $CTSTimeout$ and $ACKTimeout$, are adapted to the outdoor situation.

When a sender transmits a RTS to its intended receiver in order to initiate the four way handshake, it has to wait at most $CTSTimeout$ for a CTS packet from its counterpart. If no CTS has arrived before the timer expires, the sender understands that its transmission failed. A similar situation exists after the transmission of a data packet. In this case, the sender has to wait at most $ACKTimeout$ for an answer from the receiver. This is illustrated in figure 2.21: If the air propagation
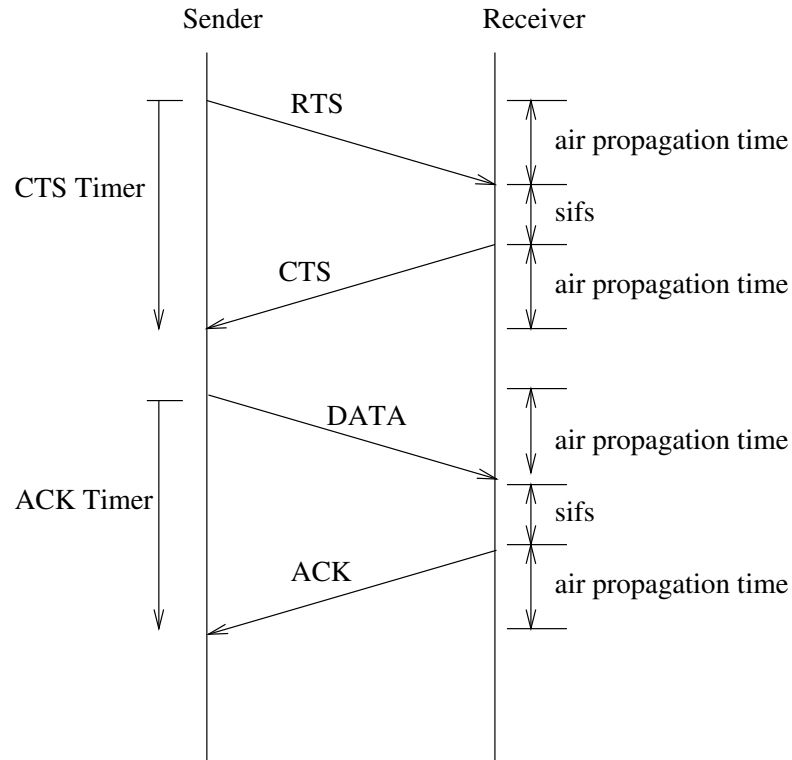

Figure 2.21: Four way handshake in IEEE 802.11 with air propagation time.

time is too long, the CTS timer can reach $CTSTimeout$ before the arrival of the CTS. In this case, the transmission is aborted and no data transmission is possible.

On the other hand, if $CTSTimeout$ and $ACKTimeout$ are long enough to compensate for higher air propagation times, the MAC protocol works correctly. However, the number of collisions slightly increases because vulnerability periods are longer and there are also bad evaluations of the NAV, this is shown in appendix B.

2.4.3 Capacity Study

Dimensioning the satellite link according to the number of users and other system requirements is based on the packet delivery ratio (PDR) of UDP packets and on

the WWW throughput experienced by subscribers. In a first approach, a PDR of 99% is chosen as cut-off value for an acceptable video streaming service [3]. Moreover, a minimum of 128 Kbps is required for WWW throughput. Simulation results for
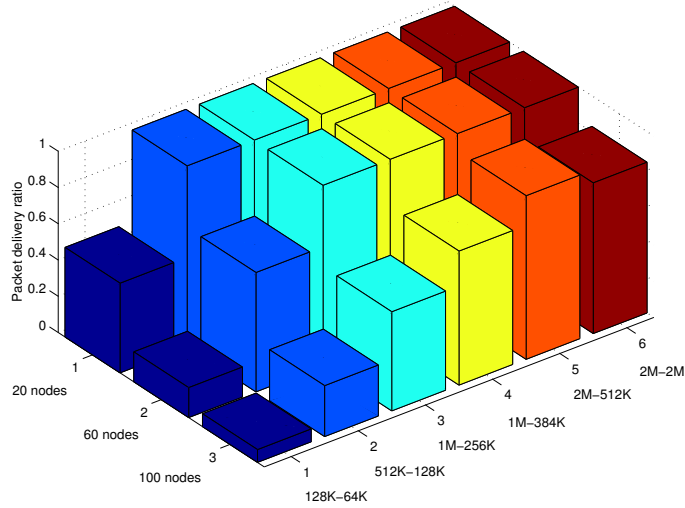

Figure 2.22: UDP packet delivery ratio for High Speed Internet scenarios.
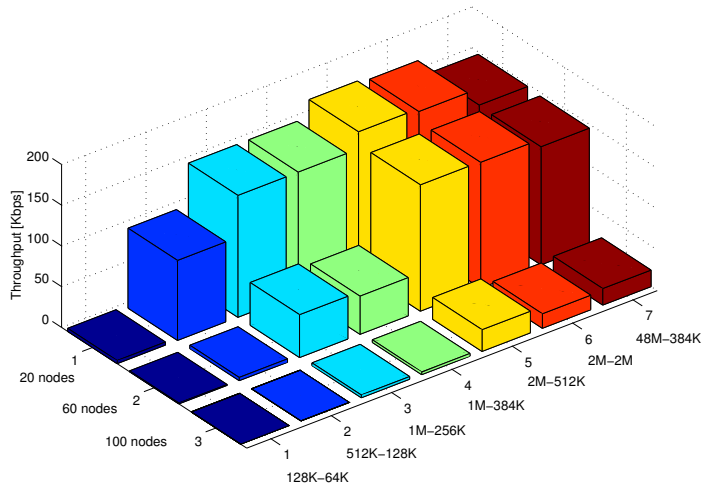

Figure 2.23: WWW throughput for High Speed Internet scenarios.

these two metrics are provided in figures 2.22 and 2.23. With twenty subscribers, the minimum configuration is (1000, 256) Kbps. With sixty users, (2000, 512) Kbps are needed. However, the system cannot support hundred users with the given scenario.
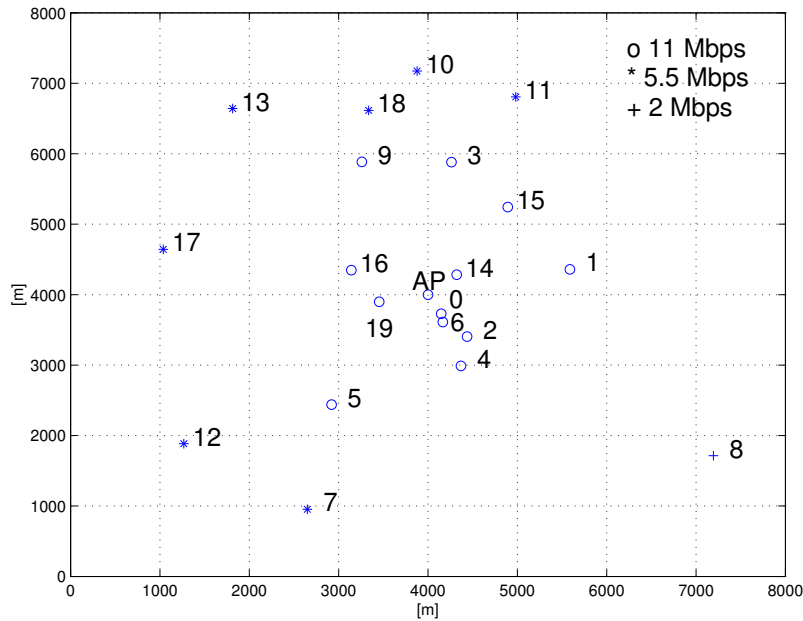
Figure 2.24: Twenty subscriber network.

Twenty Subscribers

Figure 2.24 shows the geographical distribution of twenty subscribers. Note that most of them are in the 11 and 5.5 Mbps areas. If we look at the packet delivery ratio of UDP packets, the $(128, 64)$ Kbps solution can be immediately excluded. In all other cases, UDP packet delays and jitters are acceptable (figures 2.25 and 2.26). Even in the worst case, the jitter is approximately 0.17 s. It means that 99% of the packets arrive in a interval of 0.34 s around the mean delay. Hence, the interval between two consecutive packets is not greater than 0.68 s in most cases. With an average throughput of 256 Kbps and 1024 byte packets, the buffer size of the subscriber has to be approximately 22 packets. This is a reasonable value. If the $(128, 64)$ Kbps solution is excluded, WWW users experience in all cases at least 100 Kbps, and more than 128 Kbps if the $(1000, 256)$ Kbps combination is chosen (figure 2.23).

The huge difference with the reference case without satellite is explained by the large round trip time introduced by the satellite link. As the slow start period begins, the transmission window of TCP is 1. Thus, 500 ms are needed for the server after the first TCP segment to send the second one. The same phenomenon is observed for e-mail traffic (figure 2.27). It is further worsened by the small volume

---

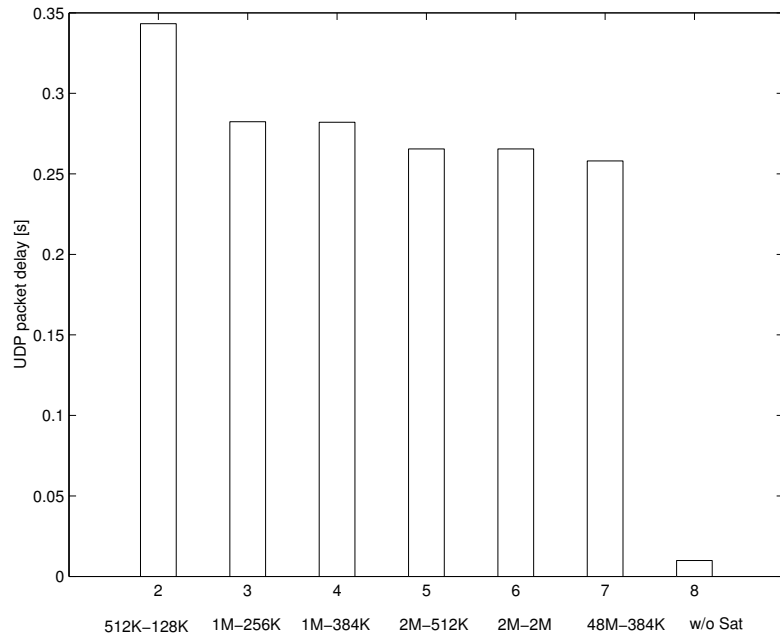[3]This criterion should be refined in further studies.

Figure 2.25: UDP packet delay, twenty subscriber network.

of the transmitted data: TCP remains in the slow start period during the e-mail transfer. This highlights the necessity to use a different version of TCP adapted to satellite links.

As a conclusion for twenty users, the $(512, 128)$ Kbps solution is the smallest possible configuration but $(1000, 256)$ Kbps is needed if a throughput of 128 Kbps has to be achieved.

Sixty Subscribers

The network of sixty subscribers is depicted in figure 2.28. The mean UDP packet delay figures confirm that a forward link of 1 Mbps is not sufficient for this scenario (figure 2.29). UDP packet jitters have the same order as for twenty users provided that the 1 Mbps forward link solutions are excluded (figure 2.30). With a 2 Mbps downlink, e-mail senders can expect 15 Kbps throughput (figure 2.31).

Hundred Subscribers

While in the first two cases, the limiting factor was the downlink bandwidth from the satellite to the AP, here (the network is shown in figure 2.32), the input load exceeds the WLAN capacity. The PDR of UDP packets (figure 2.33) doesn't match the
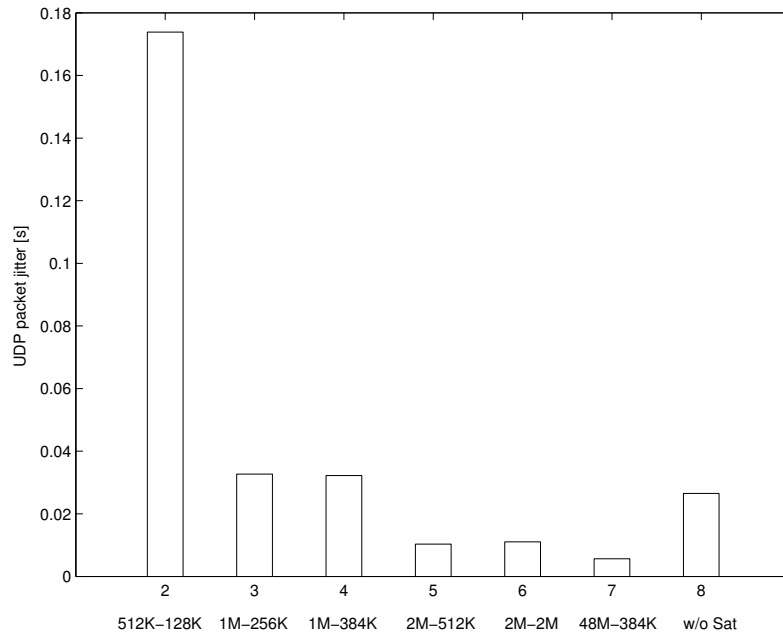
Figure 2.26: UDP packet jitter, twenty subscriber network.

requirement even without satellite link. As far as WWW throughput is concerned, no configuration provides an acceptable quality of service. Even if the network delay is zero, the throughput doesn't exceed 100 Kbps (figure 2.34).

One output of the preceding studies on TCP, UDP, and high speed Internet is that the cell capacity is highly dependent on the spatial distribution of the terminals. This is due to the link adaptation mechanism that reduce the transmission rate when the channel conditions are degraded. Some solutions are detailed in the next section.
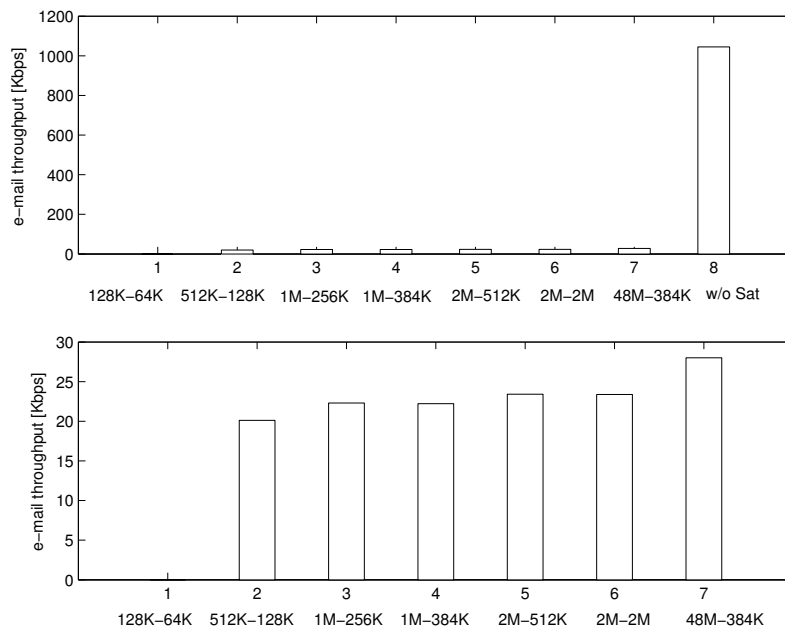
Figure 2.27: E-mail throughput, twenty subscriber network.

Figure 2.28: Sixty subscriber network.

Figure 2.29: UDP packet delay, sixty subscriber network.
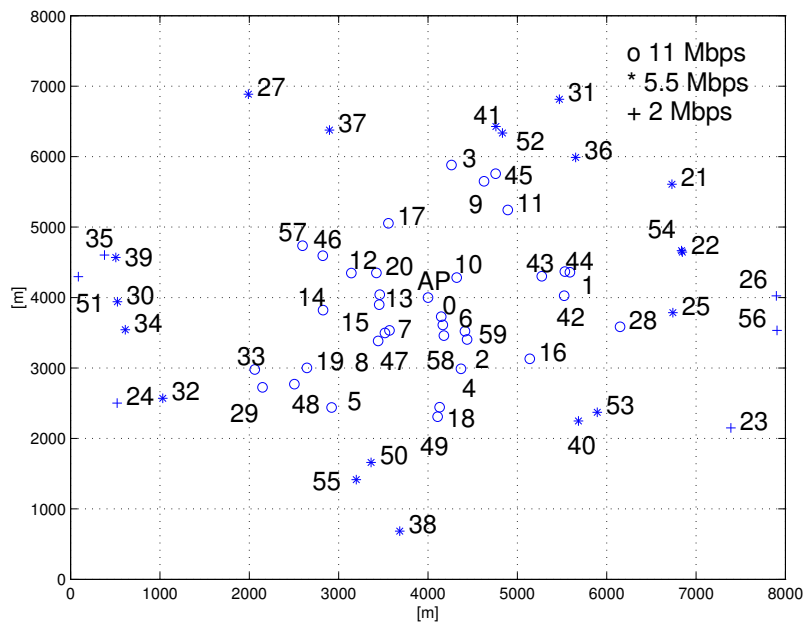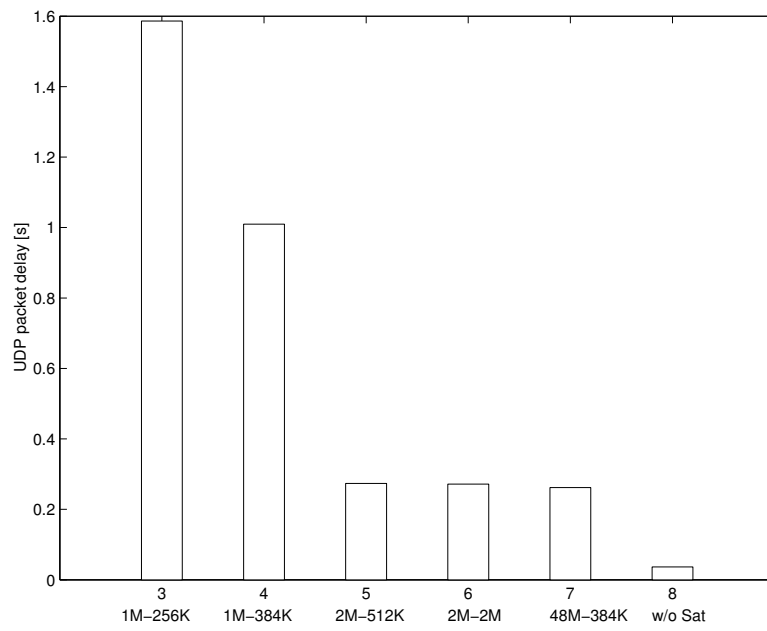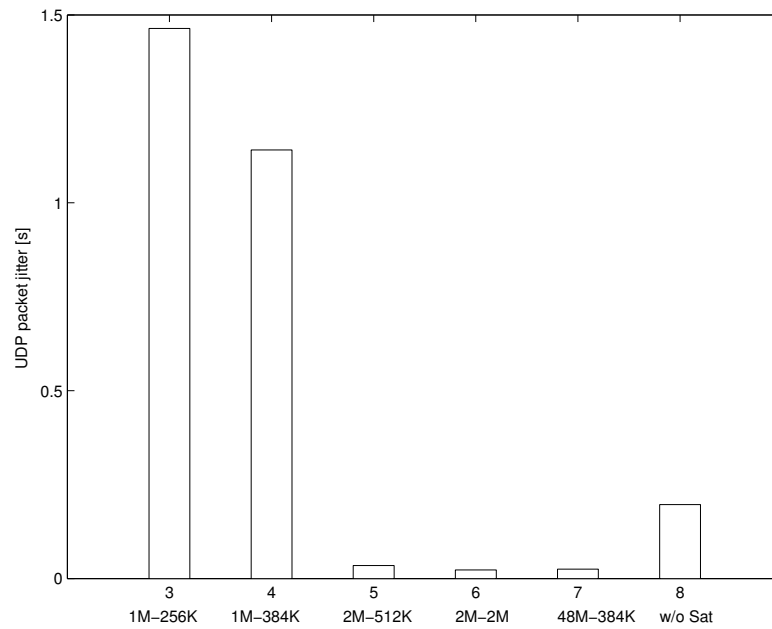

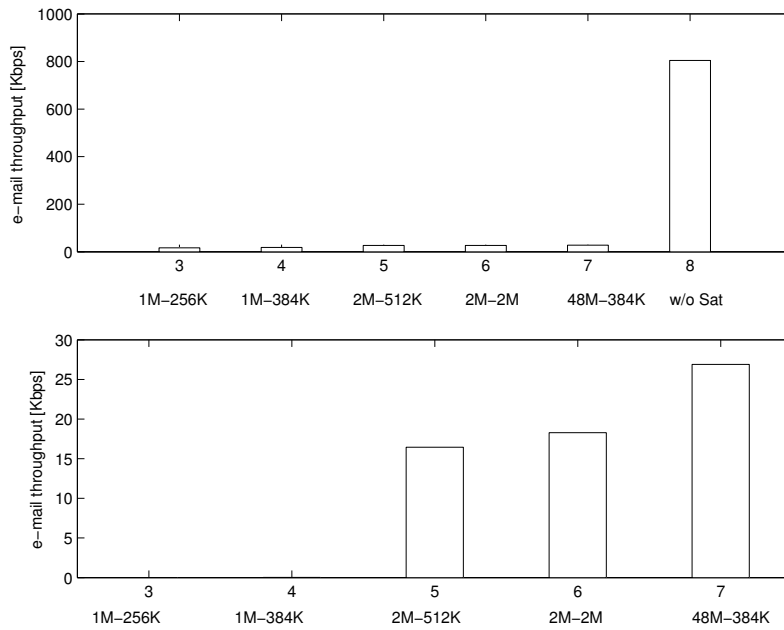Figure 2.30: UDP packet jitter, sixty subscriber network.

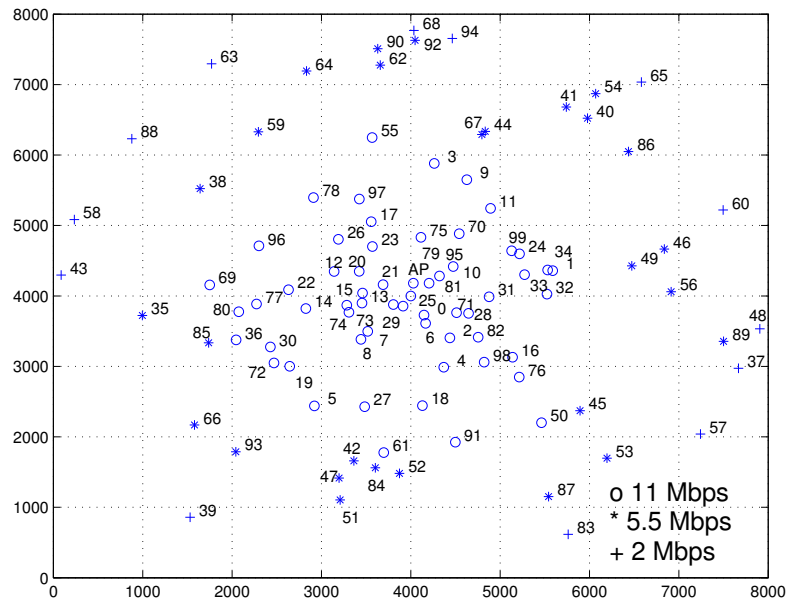Figure 2.31: E-mail throughput, sixty subscriber network.
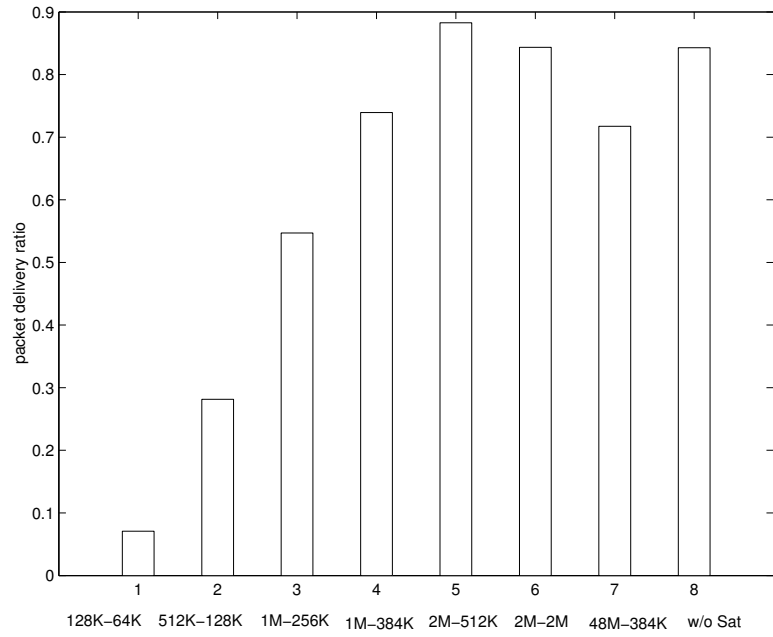

Figure 2.32: Hundred subscriber network.

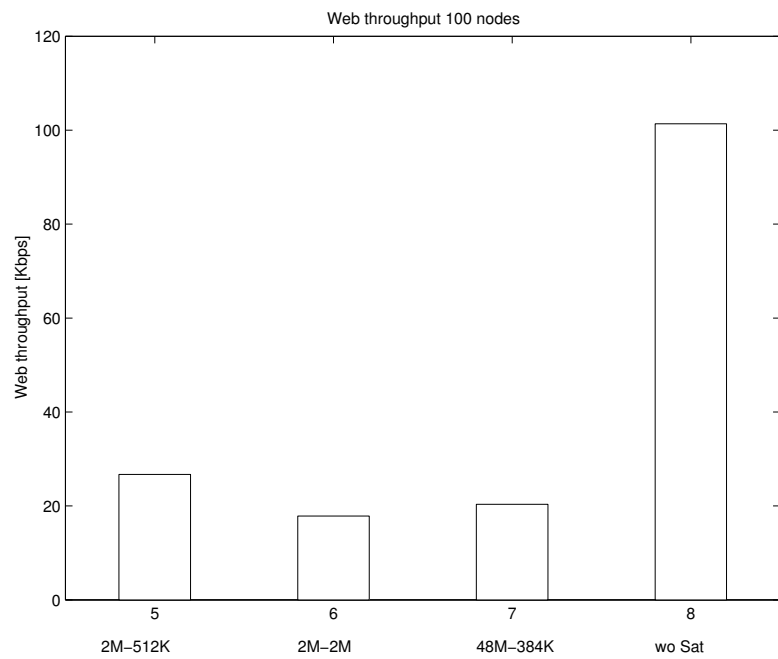Figure 2.33: UDP packet delivery ratio, hundred subscriber network.

Figure 2.34: WWW throughput, hundred subscriber network.

2.5   The Near-far Effect

The near-far effect has a determining impact on the overall capacity and cost of an AP centric IEEE 802.11 network. This problem has been identified in [110]. It has been observed that when some mobile terminal uses a lower bit rate than the others, the performance of all terminals is considerably degraded. Besides, the cell aggregate throughput is also significantly affected.

This problem has been illustrated for IEEE 802.11b. However, any multi-mode WLAN system based on DCF, i.e., 802.11a/b/g, is affected in the same manner. In this section, we provide an analysis of the problem through simulations for various transmission scenarios with IEEE 802.11b. Some solutions are suggested.

2.5.1   Illustration

The near-far effect is now illustrated with different traffic types and transport protocols. For improving the downlink, two solutions are proposed, namely the relay based and the window based solution. For the uplink, packet sizes of the terminals can be adapted to their physical mode.

Downlink Traffic

A simple scenario is considered with two terminals in the cell (see figure 2.35). *Terminal 1* is assumed to be close to the AP, i.e., it can transmit and receive packets most of the time at 11 Mbps. In our simulations, terminal 1 is at 5 m. After 5 s of simulation, *terminal 2* is introduced at d m (35 or 45 m) from the AP with the same type of traffic. 35 m is in the 2 Mbps area and 45 m is in the 1 Mbps area.

The AP is sending packets of 1024 bytes at a constant bit rate over UDP to the two terminals. The input is so high, that it has always something to send. The buffer of the AP is assumed to be infinite and serves the packets according to First In First Out policy.

Simulation results show a very fair behavior of the IEEE 802.11 DCF MAC protocol, since both terminals get the same throughput irrespective of their distance to the AP, i.e., irrespective of the physical mode they use for data transmission (figure 2.36).

But, this fairness leads to a very bad situation for terminal 1 that experiences a
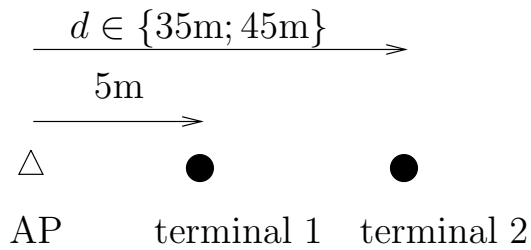
Figure 2.35: Scenario with two terminals to illustrate the near-far effect on the downlink.
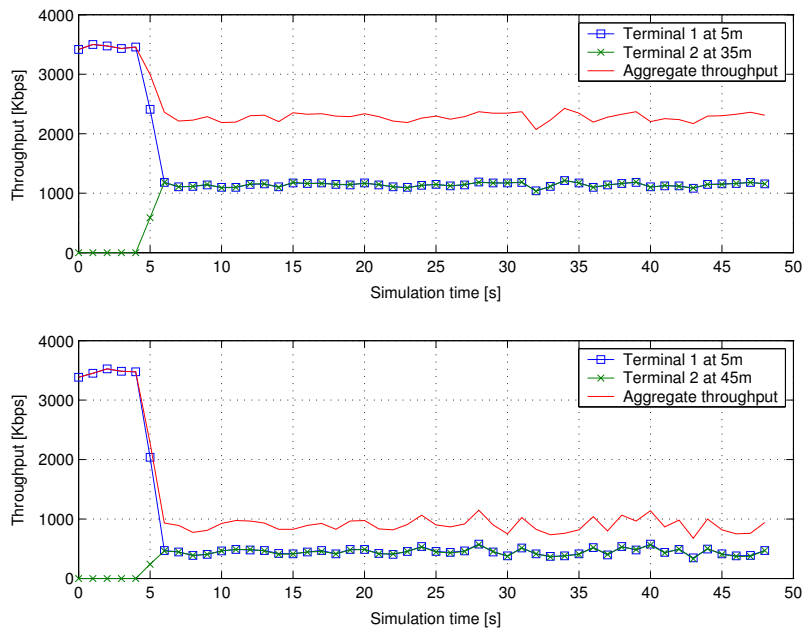
Figure 2.36: Terminals and aggregate throughput vs. simulation time, UDP downlink traffic, $d \in \{35\text{m}; 45\text{m}\}$.

significant performance degradation. When terminal 2 is at 35 or 45 m, terminal 1 sees a loss of respectively 57% and 86% of its throughput. The aggregate throughput drops too. This is not the case when terminal 2 is at 5 m or even at 25 m (figure 2.37)

Indeed, as explained in [110] and for equal size packets, a 1 Mbps terminal will occupy the channel approximately 11 times more than a 11 Mbps terminal to transmit a packet. Its data rate is smaller, but its channel occupancy is higher. This phenomenon leads to an equal throughput for both terminals. Besides, since most of the time the channel is used by the low bit rate terminal, the aggregate throughput is also reduced.
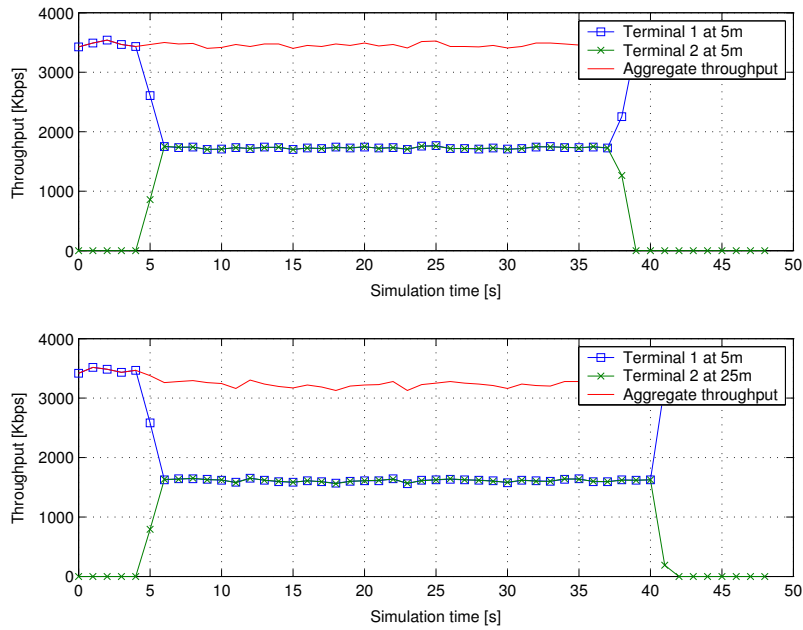
Figure 2.37: Terminals and aggregate throughput vs. simulation time, UDP down-link traffic, $d \in \{5\text{m}; 25\text{m}\}$.

A field experimentation in an office building illustrates also the near-far effect. As shown in figure 2.38, six measurements have been done for six positions of terminal 2 thanks to a radio sniffer. The UDP throughput and the physical mode distribution is given on figure 2.39 when terminal 2 is alone in the considered cell and 1500 byte packets are sent from the AP. We see a predominance of the 2 Mbps data rate. Figure 2.40 shows the UDP throughput of two terminals when terminal 1 is close to the AP and terminal 2 moves along the six measurement points. The throughput of terminal 1 alone in the cell is also given. Until 64 m, we have an illustration of the
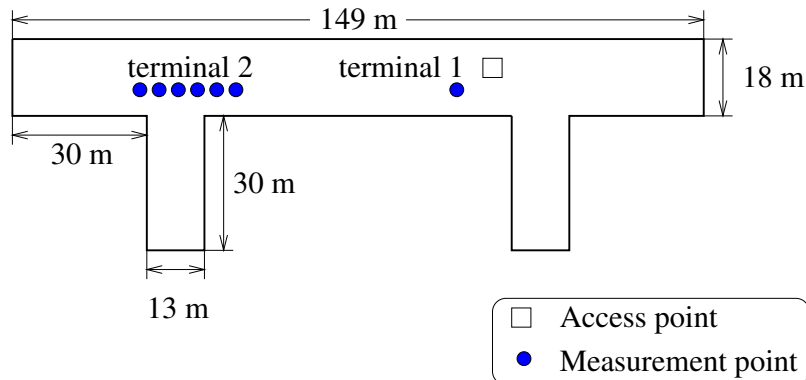
Figure 2.38: Field experimentation in an office building.

78
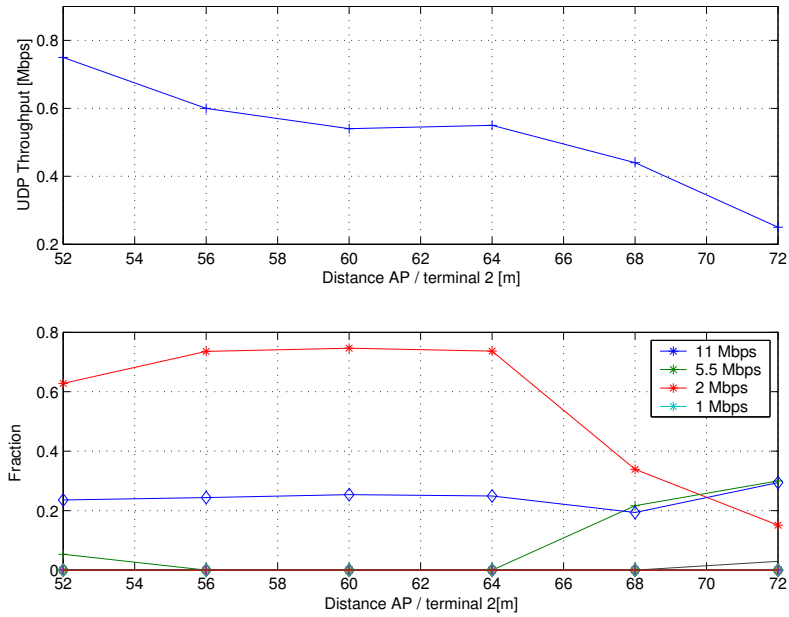
Figure 2.39: UDP throughput and physical mode distribution vs. the distance AP/terminal 2, field experimentation.

near-far effect: The aggregate throughput is reduced and for terminal 2 positions between 60 and 64 m, both the terminals have a similar throughput. After 64 m, terminal 2 experiences a lot of disconnections and terminal 1 takes advantage of this to increase its performance.

Note that in the downlink case, it is inefficient to adapt the packet size to the physical mode, e.g., allowing only long packets for terminal 1 and short ones for terminal 2. For the same input load for both users, the number of short packets would indeed exceed the number of long ones in the AP buffer and so on the channel (figure 2.41). Terminal 2 packets would in this case still occupy the channel most of the time.

It is however possible to adopt a scheduling policy at the AP allocating more time for the transmissions of terminal 1.

It can be expected that the impact of near-far effect will smooth out as the number of terminals increases. The degradation of performance depends on the proportion of terminals in each physical mode areas. Let us consider four terminals at 5 m, and a single terminal at d m (45 m and 5 m) as shown in figure 2.42. In the first simulation (figure 2.43), the significant degradation of the aggregate throughput is again observed (from 3460 Kbps to 1530 Kbps, i.e., -56%). This result has to be
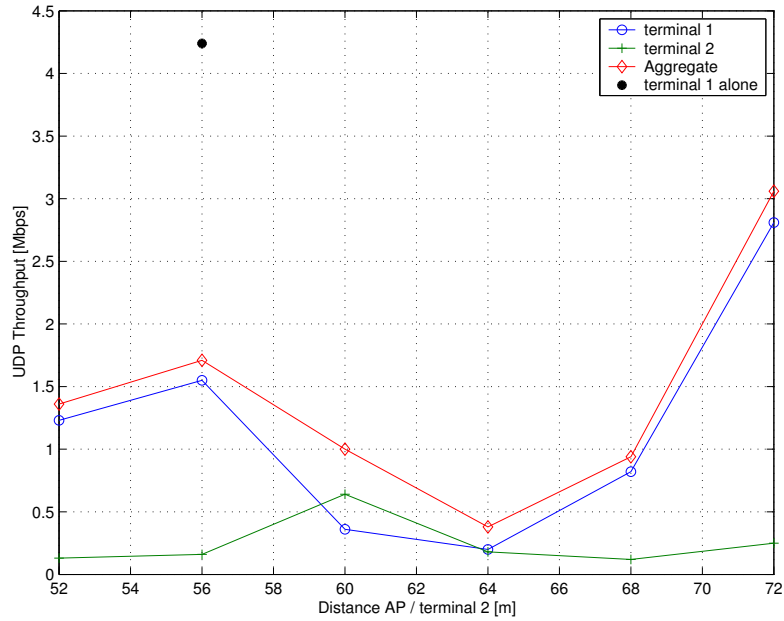
Figure 2.40: Terminals and aggregate UDP throughput vs. the distance AP/terminal 2, field experimentation.

compared with that of figure 2.36. In this case, the aggregate throughput doesn't exceed 1 Mbps. The difference is explained by the proportions of terminals in the two physical modes areas: 50% are in the 11 Mbps zone in figure 2.36, while they are 80% in figure 2.43. Terminals 1 to 4 see their throughput dropping from 864 Kbps to 306 Kbps. For $d = 5$ m it drops to 693 Kbps but the aggregate throughput is constant.

We now turn to a scenario with TCP and two terminals as in figure 2.35. Simulations are done with an advertized window of 64 segments of 1024 bytes and the near-far effect is still observable (figure 2.44).

It can be noted that at 45 m the throughput is less stable. This is due to a higher PER of terminal 2 in this particular scenario. If a packet is lost, the TCP congestion window for terminal 2 is reduced. Terminal 1 takes advantage to receive more data. Consequently, the aggregate throughput increases as well.

We now propose two solutions for the near-far effect on the downlink. The first one is to have a relay node in the direction of terminal 2. The second one is based on an adaptive advertized window of TCP.

80
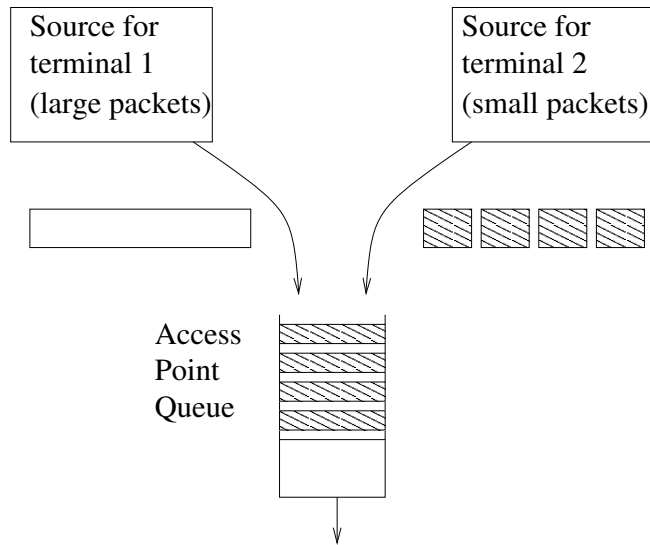
Figure 2.41: Variable packet sizes on the downlink in case of near-far effect.
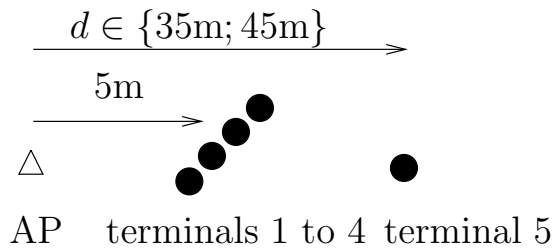
Figure 2.42: Scenario with five terminals to illustrate the near-far effect on the downlink.

2.5.2 Relay Based Solution

In this solution, a relay node is placed at a distance d_R from the AP (20 m and then 30 m, i.e., successively in the 11 Mbps and in the 5.5 Mbps area). Terminal 1 is still located at 5 m from the AP, while terminal 2 is at 35 m and then 45 m from the AP (see figure 2.45). A UDP downlink traffic is considered for both terminals and all packets destined to terminal 2 are routed through the relay node, and thus transmitted twice to reach the receiver.

Simulations results show that the near-far effect is partially mitigated thanks to the relay node when terminal 2 is at 45 m from the AP (figure 2.46): The aggregate throughput equals or exceeds 2 Mbps in all cases. This value has to be compared to the hardly achieved 1 Mbps in figure 2.36. The communication between the AP and terminal 2 is indeed made of two high data rate links. The relay-terminal 2 link is often a 11 Mbps link, while the AP-relay link is either at 5.5 Mbps or 11 Mbps
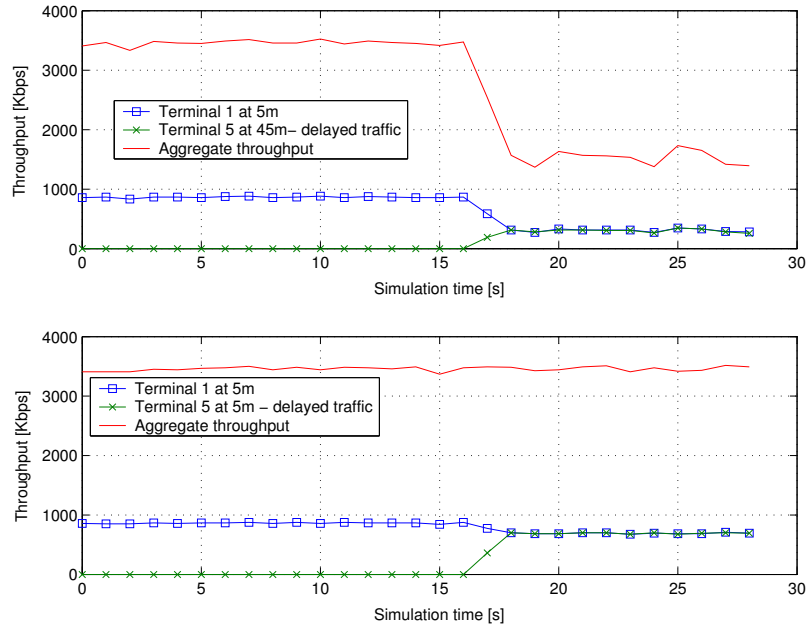
Figure 2.43: Terminals and aggregate throughput vs. simulation time, five terminals, UDP downlink traffic, $d \in \{5m; 45m\}$.

according to the relay position. Packets at this high data rate occupy much less the channel than in the previous section. And this, even if they have to be routed.

When terminal 2 is located at 35 m from the AP, there is however no mitigation of the near-far effect (see figure 2.47 to be compared with results in figure 2.36). In this case, the direct link is more efficient than two links with a higher data rate. Hence, a trade-off has to be found by the routing protocol according to the respective locations of the terminals and of the relay node.

The mitigation of the near-far effect has been shown in the downlink case, but the result holds for the uplink. It may not be realistic to deploy a specific node in various directions to solve the problem. However, the proposed solution can be seen as a interesting byproduct of coverage extension (see section 2.6).

2.5.3 Window Based Solution

The question is now to know if we can combat the near-far effect by using the advertized window of TCP, i.e., the maximum size for the transmission window. The answer is yes, as shown in the following simulation results. The considered scenario is again that of figure 2.35 with two TCP connections and $d = 35$ m. The TCP connection with terminal 2 starts after 5 s of simulation and stops after 20 s.

82
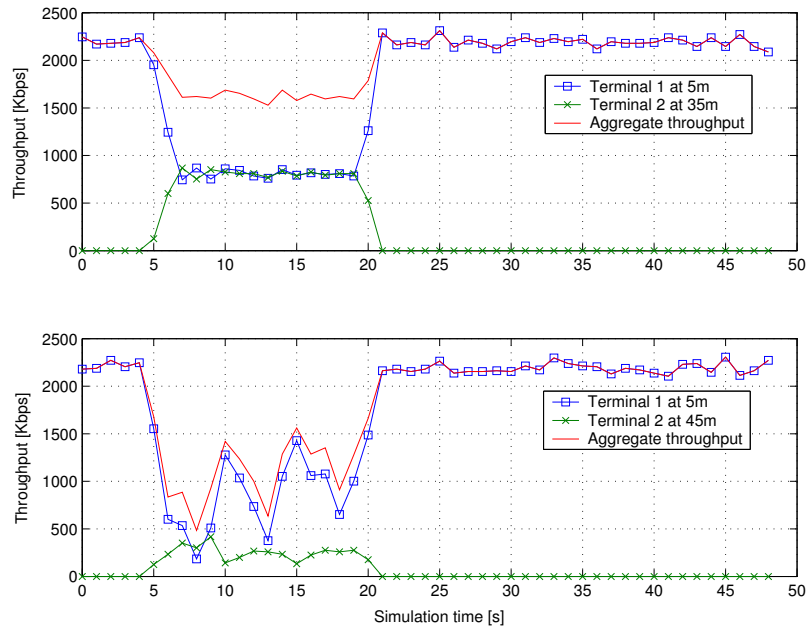
Figure 2.44: Terminals and aggregate throughput vs. simulation time, TCP down-link traffic, $d \in \{35\text{m}; 45\text{m}\}$.

The advertized window of terminal 1 has been set to 64 segments of 1024 bytes. The advertized window of terminal 2 varies from 2 to 256 segments of 1024 bytes[4].

Figures 2.48 and 2.49 show that the terminal throughput can be modulated using the advertized window of terminal 2. Besides, going from 64 to 256 doesn't bring anything to the throughput of the terminals. In this case, the maximum value of the transmission window of terminal 2 is indeed not reached.

[4]The maximum window size allowed by TCP is 64 Kbytes, extended versions allow 128 Kbytes, 256 Kbyte performance is provided here for reference only.
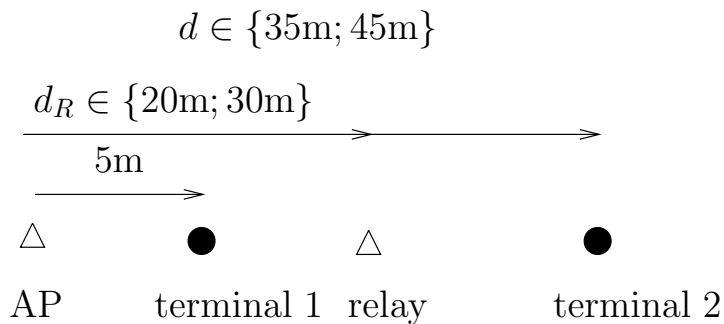
Figure 2.45: Scenario with two terminals and a relay node.

Figure 2.46: Terminals and aggregate throughput vs. simulation time, relay node solution, UDP downlink traffic, $d = 45$m, $d_R \in \{20m; 30m\}$.

Figure 2.50 highlights the fact that a trade-off has to be made by the system designer. The shorter is the advertized window of terminal 2, the higher is the aggregate throughput of the cell. In our simple scenario, an advertized window of 16 seems to be appropriate. We have indeed only a small degradation of the aggregate and terminal 1 throughputs. Besides, terminal 2 is satisfied in the sense that it gets approximately a similar throughput as if it were sharing the medium with another terminal in the 1 Mbps area.

These simulations suggest a simple algorithm that adapts the advertized window of the terminals according to their physical mode. This mechanism could be implemented without modifying the off-the-shelf products, which is a significant advantage for a company not involved in card design nor in MAC standardization (see [29]). On the uplink, the terminal imposes a maximal value for its transmission window. On the downlink, the terminal adapts its advertized window and sends back this value in its ACK packets.

2.5.4 Uplink Traffic and Packet Size Based Solution

Is the uplink also affected by the near-far effect? The problem is less pronounced, and adapting the packet length according to the physical mode can help. The reason
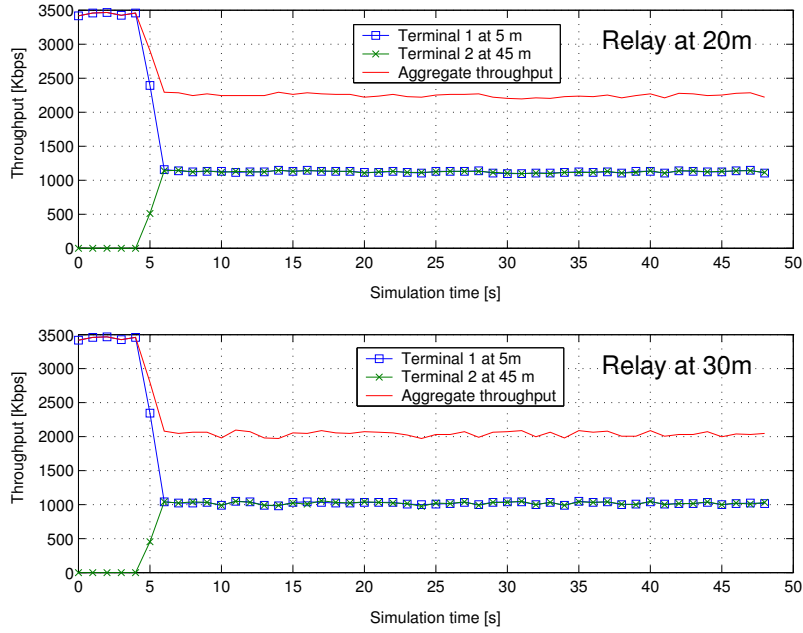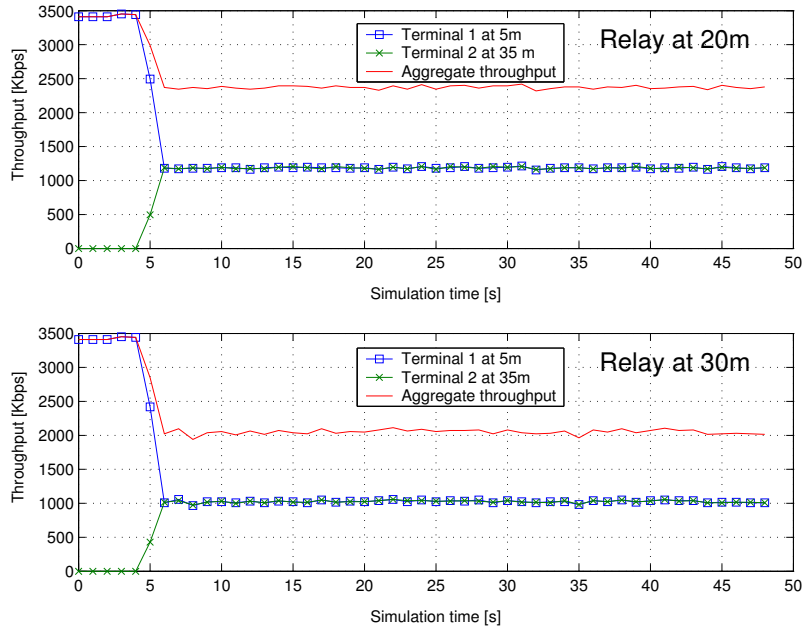
Figure 2.47: Terminals and aggregate throughput vs. simulation time, relay node solution, UDP downlink traffic, $d = 35$m, $d_R \in \{20$m$; 30$m$\}$.

is that short and long packets have the same probability of accessing the medium and thus alternate on the channel. Recall that it is not the case on the downlink. For the same input load towards both terminals, there are indeed much more small packets than long ones in the buffer of the AP.

Let us consider the scenario of figure 2.51 and a UDP uplink traffic. Terminal 1 sends only 1500 byte packets. On the upper part of figure 2.52, terminal 2 sends 1000 byte packets, while on the lower part, it transmits 1500 byte packets. In the first case, the near-far effect is reduced and terminal 2 doesn't see a very high degradation of its throughput.

This suggests that the aggregate and terminal throughput can be modulated by the packet length for the uplink case. This is shown in figure 2.53, where the terminal 2 packets vary from 50 bytes to 1500 bytes.

A question remains: Why communications in figure 2.52 have not the same throughput as in the downlink case (figure 2.36)? This can be explained again by the different PER experienced by the two terminals in this particular case. Each error is interpreted by the MAC layer as a collision and implies an increase of the back-off window. Terminal 2 experiences a higher PER and thus spends more time to back-off: Its throughput is hence reduced. This is confirmed by the back-off
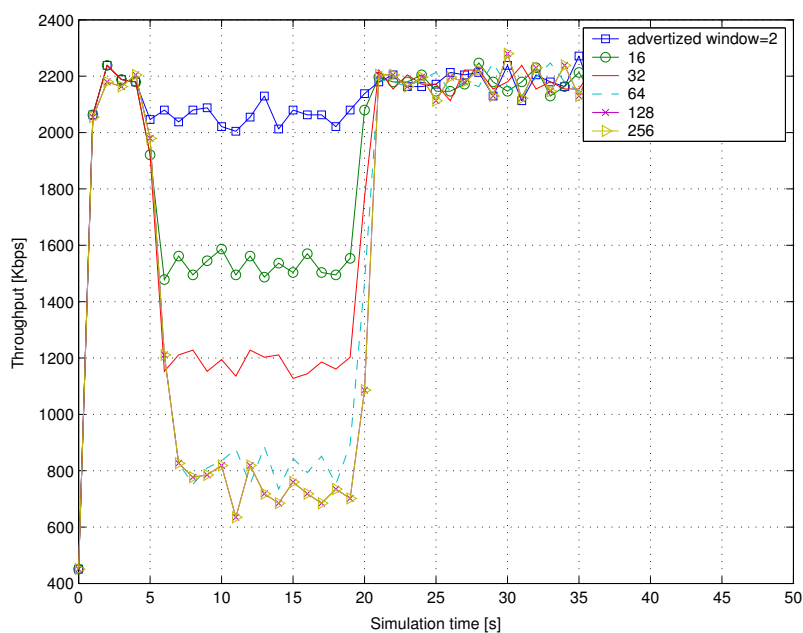
Figure 2.48: Terminal 1 throughput vs. simulation time, window based solution, TCP downlink traffic, $d = 35$m.

distributions of both terminals in figures 2.54 and 2.55.

2.5.5 Summary of Possible Solutions to the Near-far Effect

Table 2.11 gives a short summary of possible solutions to the near-far effect compatible with IEEE 802.11 DCF. Among them, we have seen an interesting application of multi-hop communications that is the relay based solution. In the next section, we further investigate the use of ad hoc concepts in AP centric networks.
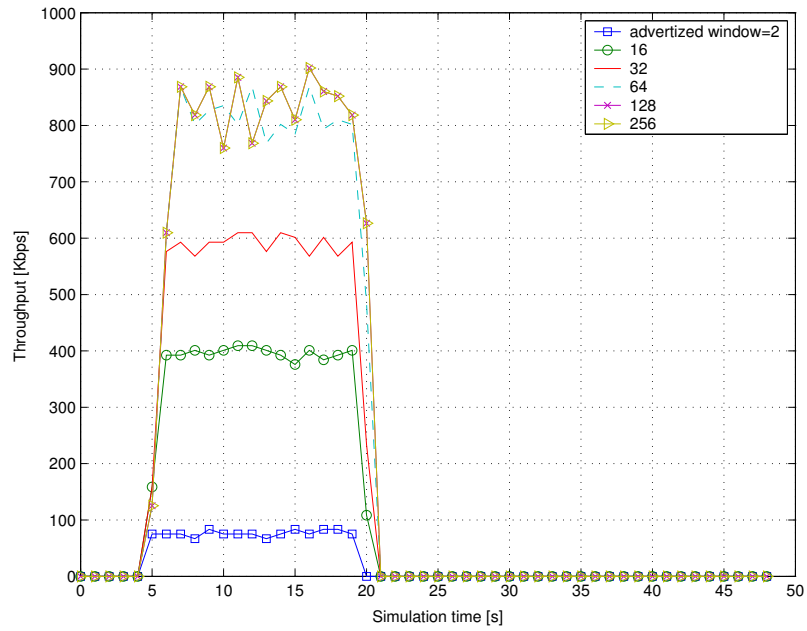
Figure 2.49: Terminal 2 throughput vs. simulation time, window based solution, TCP downlink traffic, $d = 35$m.
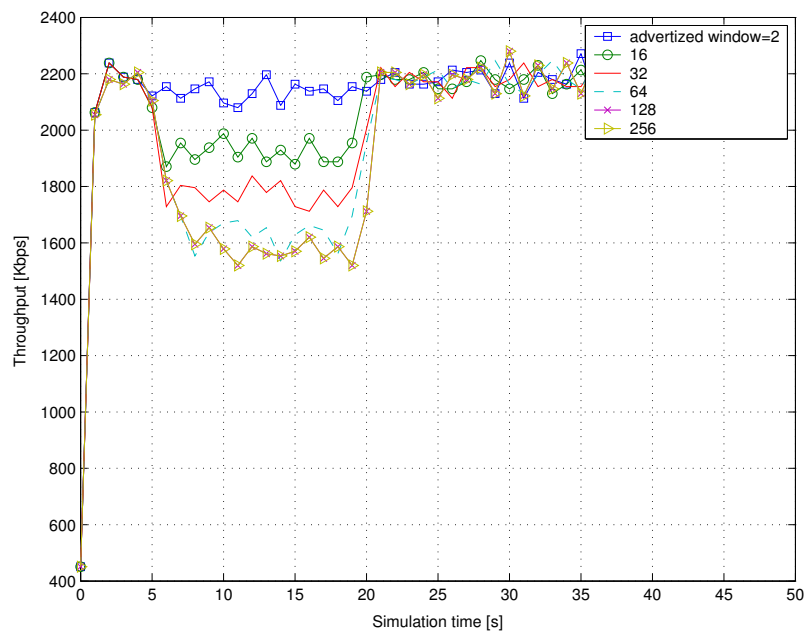
Figure 2.50: Aggregate throughput vs. simulation time, window based solution, TCP downlink traffic, $d = 35$m.
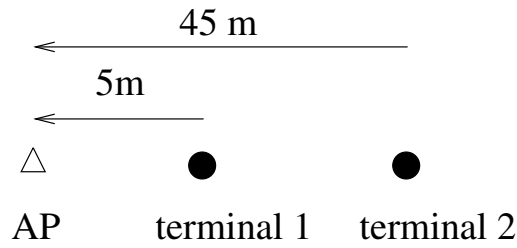
Figure 2.51: Scenario with two terminals to illustrate the near-far effect on the uplink.
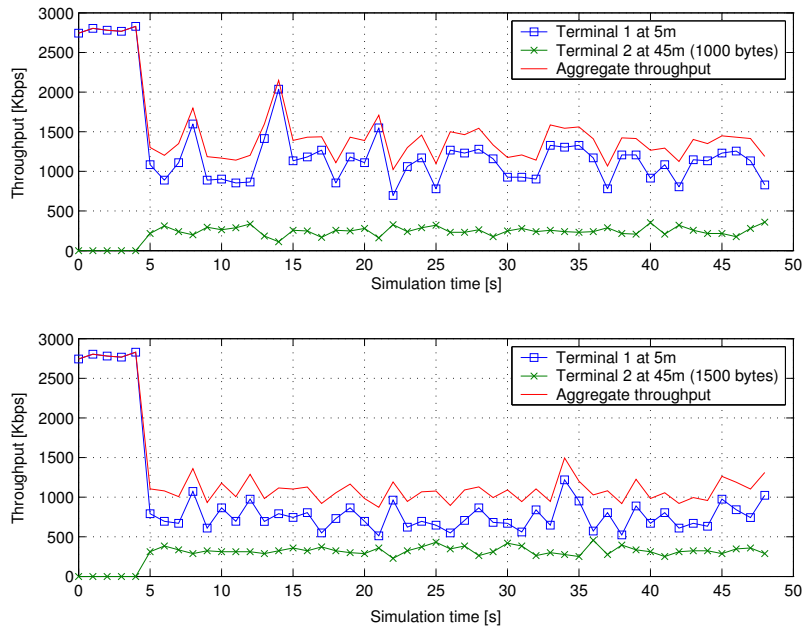
Figure 2.52: Terminals and aggregate throughput vs. simulation time, UDP uplink traffic, 1000 bytes and 1500 bytes terminal 2 packets.

Table 2.11: Summary of possible solutions to the near-far effect.

Solution	Transport	Link	Basic idea
Packet length	TCP/UDP	Uplink	Adapt packet length to the physical mode.
TCP window	TCP	Uplink	Adapt TCP transmission window to the physical mode.
TCP window	TCP	Downlink	Adapt TCP advertized window to the physical mode.
Relay node	TCP/UDP	Uplink & Downlink	Route low data rate packets through a relay.
Scheduling	TCP/UDP	Downlink	Allocate more resources to high data rate terminals.
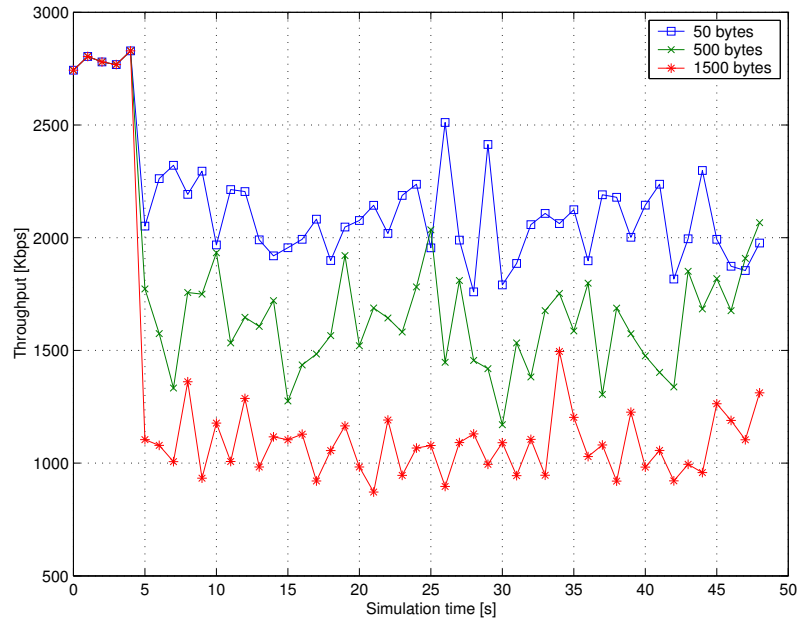
Figure 2.53: Aggregate throughput vs. simulation time, packet size based solution, UDP uplink traffic.
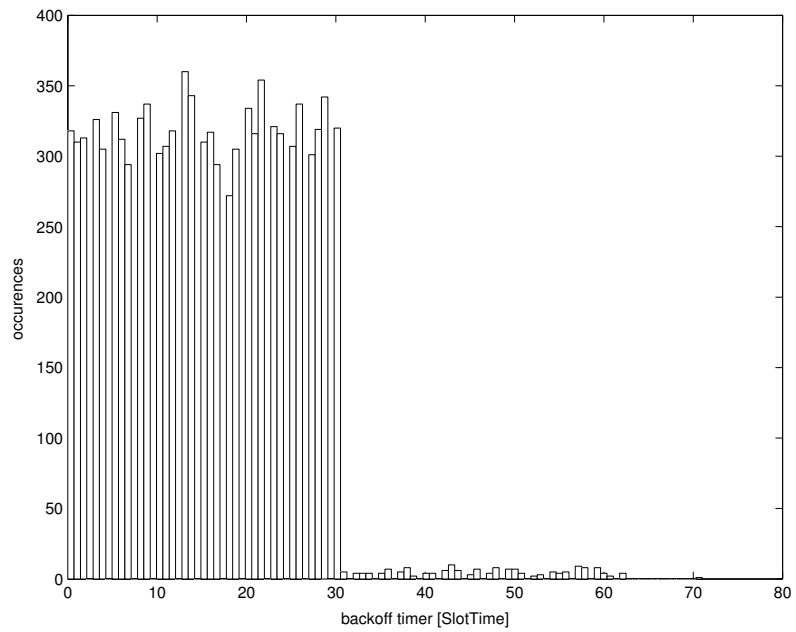
Figure 2.54: Back-off distribution of terminal 1, UDP uplink traffic.

89
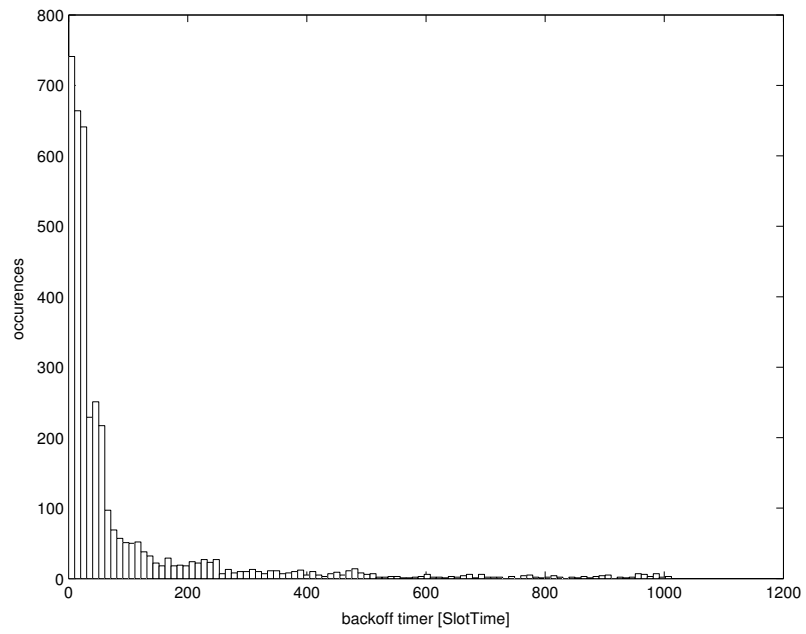
Figure 2.55: Back-off distribution of terminal 2, UDP uplink traffic.

2.6 Coverage Extension

One of the first applications for multi-hop networks seems to be the coverage extension of cellular or WLAN networks. In such a solution, a mobile node outside of the coverage of an AP can still communicate with it provided that its traffic can be routed through another node. This routing node can be either another mobile node in the communication range of the AP, or a fixed relay node belonging to the network infrastructure. There are two main reasons for considering multi-hop connections as an interesting solution:

(i) It is a low cost technology to cover dead zones of a deployment area. The radio propagation is indeed very capricious, especially in indoor environments. High variations of the signal are observable when people are moving around, when doors are closed or opened, when the receiving antenna is moved...(see [26, 27]). In practice, deployment technicians have to rely mainly on experience and a posteriori measurements. Hence, dead zones, i.e., non-covered zones may appear and could be covered by a simple fixed relay.

(ii) It is sometimes difficult to connect an AP to the wireline network. For example, deploying a WLAN in a hospital. In these cases, coverage extension with relay nodes or even a meshed network of AP are interesting solutions.

In this section, we focus on the coverage extension of a WLAN cell in indoor. We first highlight the advantage of deploying fixed relay nodes. We illustrate this solution through the case study of deployment in an office building. Finally, we evaluate throughput decrease and fairness problems implied by the multi-hop communications with IEEE 802.11.

2.6.1 Influence of Fixed Relays

We now investigate the advantage of using fixed relay nodes as a part of the network. To this end, we consider a network consisting in an AP and mobile terminals able to route data to or from the AP (figure 2.56). Terminals are assumed to be mobile in a disk of radius 150 m and all nodes have a nominal transmission range of 50 m (link adaptation is not take into account). A terminal out of range of the AP can still send or receive packets provided that a multi-hop route can be found through other nodes.

Two metrics are considered: the probability of connection and the mean number

of hops to the AP. The latter provides a measure of the available quality of service, since the throughput decreases as the number of hops increases (see section 2.6.3). These metrics for a given terminal depend on its distance to the AP.
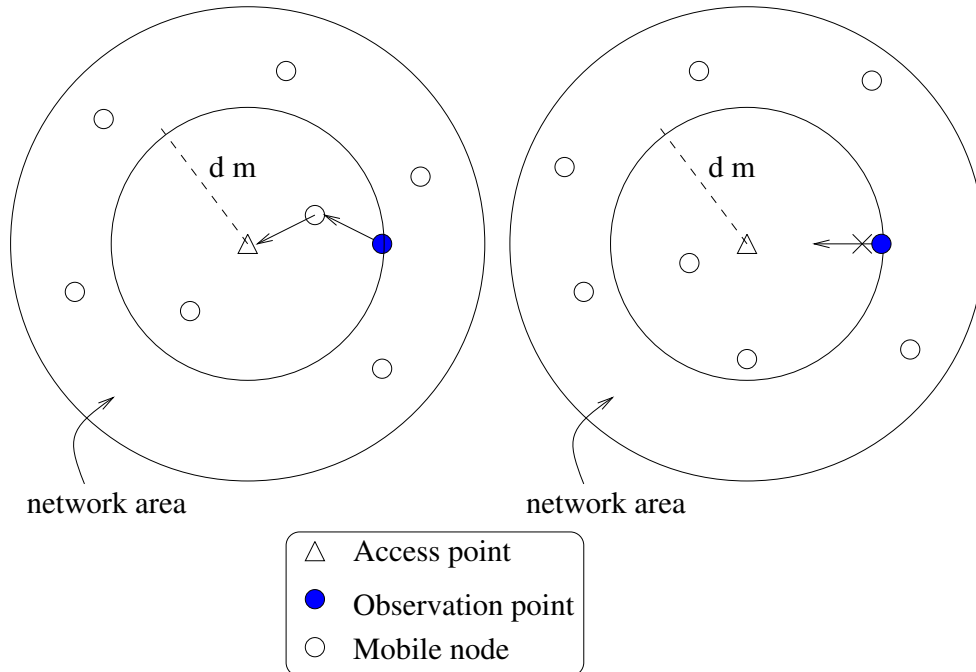
Figure 2.56: Network scenario for coverage extension Monte Carlo simulations.

Monte Carlo simulations have been performed to evaluate the performance of such a system. At a given time instant, nodes are assumed to be randomly distributed according to a Poisson point process. The mean number of nodes is N.

As shown on figure 2.56, an observation point is set in the network at a distance d and the Dijkstra algorithm is run to know the shortest path to the AP. If such a route exists, the observation terminal is connected and the number of hops is recorded. On the left hand side of figure 2.56, the observation point is for example connected in two hops. Otherwise, the observation terminal is not connected, e.g. on the right hand side of figure 2.56. The simulation stops when the 90%-confidence interval is less than 5% of the mean number of hops.

In figure 2.57, the increase of this average value is shown as a function of the distance AP / observation point. The probability of connection (figure 2.58) is highly dependent on the node density.

Let us define the cell range as the distance at which the probability of connection is greater than 90%. If terminals were not able to store and forward packets, as in

a traditional deployment, the cell range would be 50 m because the probability of connection is zero beyond the AP transmission range. In our case where terminals are ad hoc nodes, this probability is slightly decreasing with distance. With our definition, 50 is however the only number of nodes that allows a cell increase. In this case, the cell range is 80 m (a 60% increase with respect to the AP transmission range).
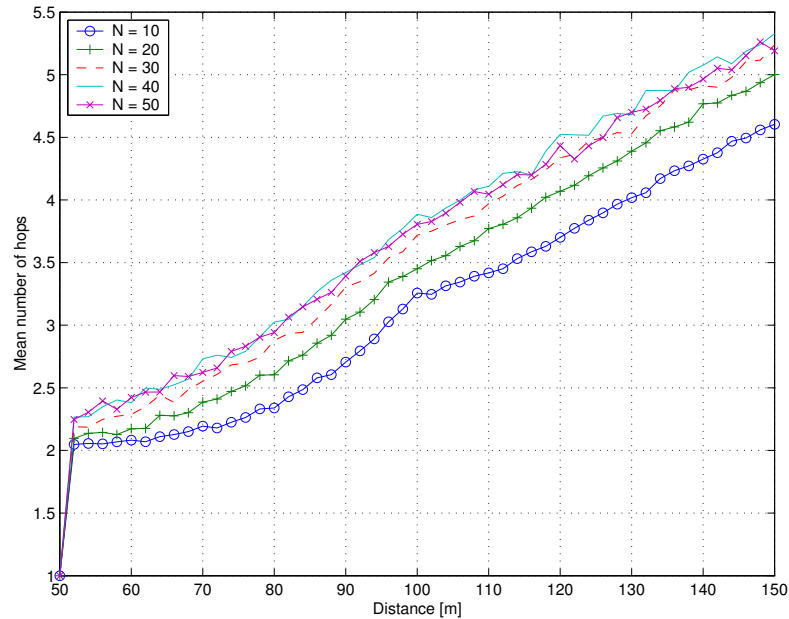
Figure 2.57: Mean number of hops vs. distance to the AP.

We now consider four fixed relay nodes located around the AP, as shown in figure 2.59. The observation point is moved in the direction of a relay. A lower mean number of hops is observed (figure 2.60). The cell range is increased to 100 m for 10, 20, 30, and 40 nodes and to 120 m for $N = 50$ (figure 2.61). In figures 2.62 and 2.63, the observation point is moved on the first bisecting line between two relays. The new cell range in this direction is now 70 m for $N = 10$ and $N = 20$, 72 m for $N = 30$, 96 m for $N = 40$ and again 120 m for $N = 50$ (thus an increase of 140% if we compare to the nominal range of an AP).

As a consequence, using fixed relay nodes increases the network connectivity and slightly decreases the mean number of hops to reach the AP.
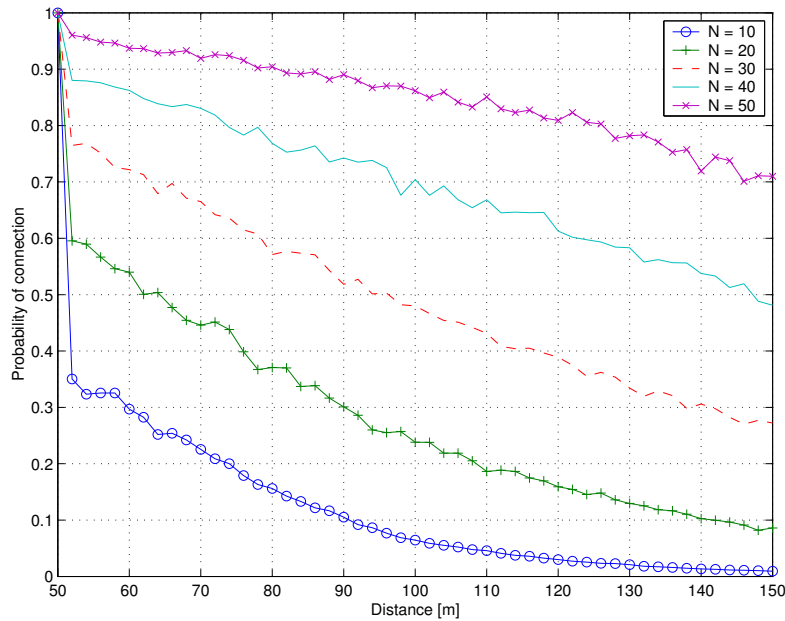
Figure 2.58: Probability of connection vs. distance to the AP.

2.6.2 Coverage Extension in a Office Building

In this section, we show an example of coverage extension of an IEEE 802.11b cell in an indoor environment thanks to multi-hop communications. By adding some relay nodes to the cell, that do not need to be connected to the wireline network, it is shown that we can overcome the rapidly decreasing signal strength in an office environment, reduce the number of dead spots, and so the number of access points. Hereafter, we pay a special attention to the computation of the carrier to interference ratio (CIR) in an ad hoc cloud. Then, Monte Carlo simulations provided in [1] are recalled as an example of application of the proposed algorithm.

Interference Computation in an Ad Hoc Cloud

An interesting metric to compare the coverage areas with and without relays is the received CIR for a given node. Hence, we consider an ad hoc cloud of nodes and relays connected to a single AP, eventually through several hops. In such a network, where the channel is shared among all entities (AP, relays, and node), simultaneous communications can take place if interferences are sufficiently low. The proposed algorithm tries to answer the following question: How to compute the carrier signal C, and the interference level I for a given communication, taken into account that simultaneous communications are possible?

94
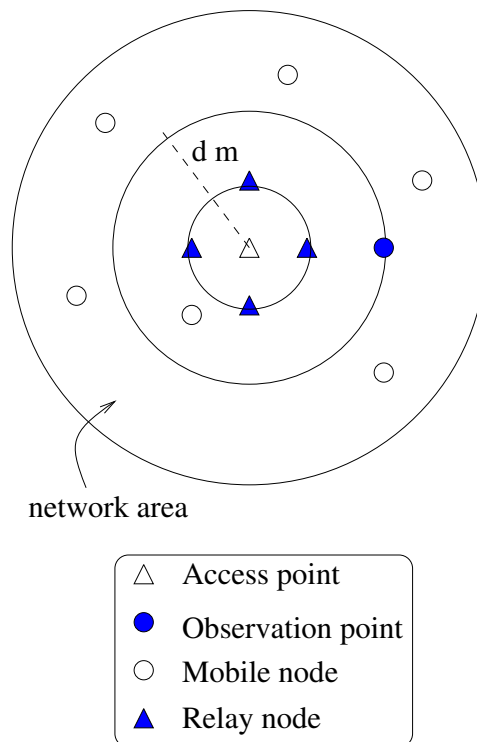
Figure 2.59: Network scenario with relay nodes for coverage extension Monte Carlo simulations.

Definition of the Graph

In order to better explain the algorithm, we focus on an example drawn in figure 2.64. The ad hoc cloud consists in an AP that is the gateway to the wireline network, one or several relay nodes (R) that belongs to the network architecture, and mobile stations (MS). The set of Rs is denoted S_R and the set of MSs, S_{MS}. We assume that the network is static in the time frame of a data packet transmission.

The network is represented by an unidirectional graph $G = (V, E)$, where V is the set of vertices or nodes, and E is the set of edges or flows. The vertices are made of the AP, the Rs and the MSs. If u and v belongs to V, (u, v) belongs to E if and only if a communication between u and v is possible at the predefined basic rate. The basic rate is the physical mode of transmission of the RTS/CTS messages, e.g., 1 Mbps or 2 Mbps for IEEE 802.11b. Now, a communication is possible if and only if the mean PER is below a PER target.

We will neglect the effect of physical carrier sensing and only consider the virtual carrier sensing. With this assumption, an unidirectional graph is a sufficient concept. Physical carrier sensing can be taken into account by considering a directional graph
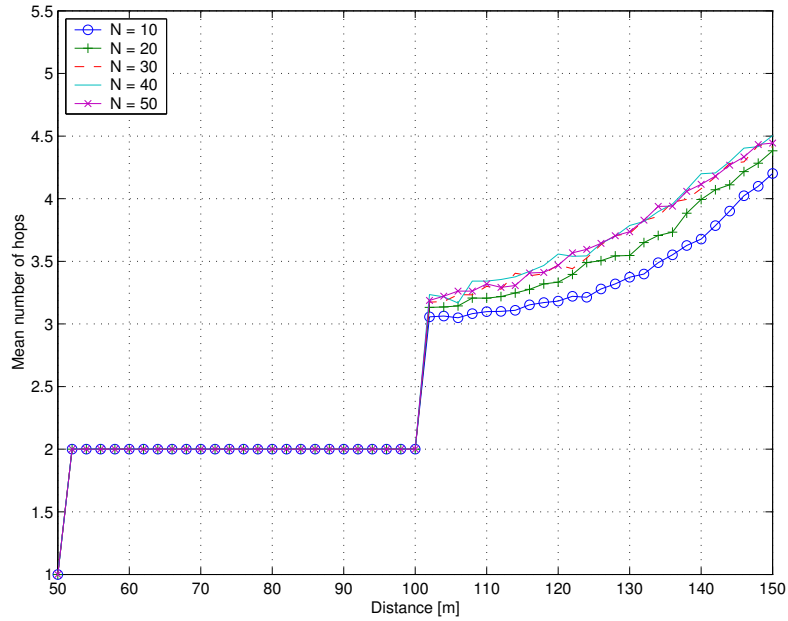
Figure 2.60: Mean number of hops vs. distance to the AP, four fixed relays, observation point moved in the direction of a relay.

and by slightly modifying the algorithm. We now focus on a given MS, MS^*.

Computation of C

First of all, a new definition of the notions of uplink and downlink is needed. We assume that the routing protocol has found a unique route from the MS^* to the AP and that this route is the same from the AP to MS^*. Moreover, without loss of generality and for the sake of simplicity, we assume that the routing protocol chooses randomly one of the shortest path between two nodes. Practically speaking, this path can be computed using the Dijkstra algorithm.

The uplink is the link between MS^* and the next node on the route towards the AP. The downlink is the last hop on the route from the AP to MS^*. Let us denote $R(MS^*)$ the last node on this path. According to the routing rule, $R(MS^*)$ is unique.

As an example, the downlink of MS_4 is obviously R_6-MS_4 and the uplink is MS_4-R_6, so that $R(MS_4) = R_6$. For MS_2, the shortest path to the AP goes through R_3, so that $R(MS_2) = R_3$.

Now, the computation of C becomes obvious for both uplink and downlink: It is the received power at MS^* (downlink) or $R(MS^*)$ (uplink).
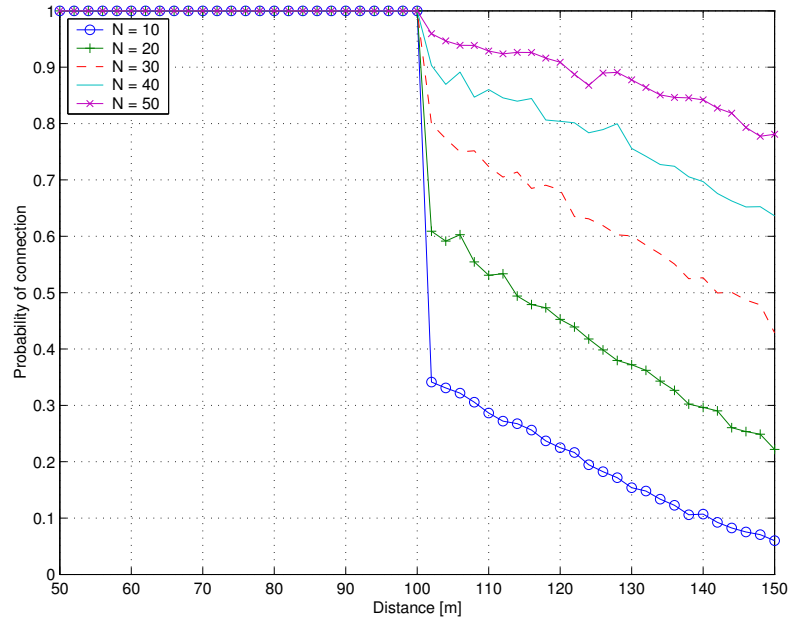
Figure 2.61: Probability of connection vs. distance to the AP, four fixed relays, observation point moved in the direction of a relay.

Computation of I

We try here to determine the set Θ of flows that take advantage of spatial reuse in the ad hoc cloud and transmit simultaneously with the considered flow, i.e., involving MS^*.

When a flow F won the channel thanks to the exchange of RTS/CTS short control packets, all flows at a distance of two hops from F are not allowed to transmit. Involved pairs have indeed heard and decoded RTS and/or CTS and prevent themselves to send any packet. We will denote $H2(F)$ the set of flows that are two-hops away from F and thus that cannot transmit simultaneously with F.

For example, look at the downlink flow F_{14}. R_6 has sent a RTS that have been received by MS_3, MS_4, R_4, and MS_2. All these nodes, except MS_3, prevent themselves to send a RTS for another communication or to respond to a RTS, because of the virtual carrier sensing. So, $H2(F_{14}) = \{F_{15}, F_{13}, F_{12}, F_{11}, F_7, F_8\}$. Note that $H2(F)$ is the same for up and downlink because of the symmetry of the RTS/CTS exchange.

Now, $H2(F)$ can also be seen as the set of contending flows for F. If we assume that the MAC protocol is fair[5], the probability for F to access the channel should

---

[5]This assumption is needed for a first simple approach, although it is questionable, especially

97
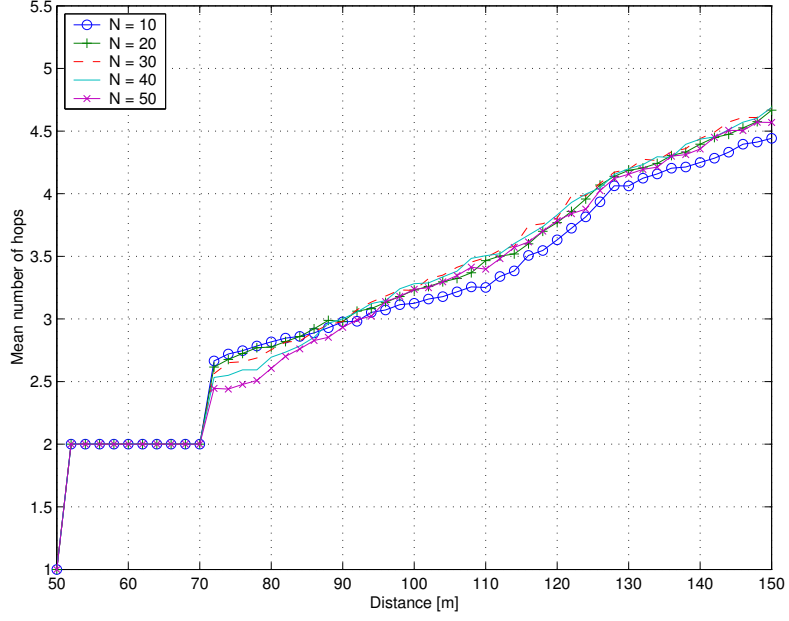
Figure 2.62: Mean number of hops vs. distance to the AP, four fixed relays, observation point moved between two relays.

be proportional to $1/d(F)$, where $d(F) = |H2(F)|$ is the cardinal of $H2(F)$, e.g. $d(F_{14}) = 6$. This assumption is in compliance with the main result of [129] that claims that the maximum probability of a successful transmission is upper bounded by $0.9278/N$, where N is the average number of nodes in the transmission range.

These considerations lead to the following algorithm for the computation of I for the flow F^*. Note that direct communications between MSs may not be allowed. In this case, the corresponding flows, e.g., F_{16}, are not considered as contenders for the shared medium.

1. $\Theta = \{F^*\}$, $T = E \setminus \{\{F^*\} \cup \{F_i | F_i \text{ not allowed}\}\}$.

2. For each F in V, build $H_2(F)$ and compute $d(F) = |H_2(F)|$.

3. $T = T \setminus \{H_2(F^*)\}$.

4. While $T \neq \emptyset$,

   (a) Choose randomly F_i in T with probability
       $\alpha/d(F_i)$, α is such that $\alpha \sum 1/d(F_i) = 1$.

   (b) $\Theta = \Theta \cup \{F_i\}$.

---

in multi-hop networks.
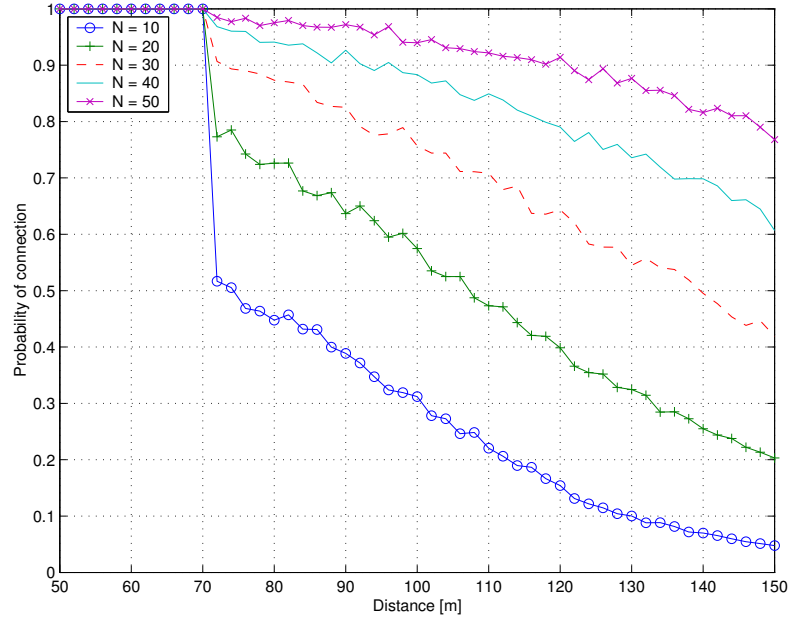
Figure 2.63: Probability of connection vs. distance to the AP, four fixed relays, observation point moved between two relays.

(c) $T = T \setminus \{F_i\}$.

(d) $T = T \setminus \{H_2(F_i)\}$.

5. For each F in Θ, F is considered to be uplink with probability p_{UL}, otherwise it is a downlink.

6. I is the aggregate received power from all flows F in Θ.

We will now run the algorithm for the given example for flow $F14$ downlink.

1. $\Theta = \{F_{14}\}$, $T = \{F_0, F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8, F_9, F_{10}, F_{11}, F_{12}, F_{13}, F_{14}, F_{15}, F_{16}\}$.

2. Determination of $H2(F)$: see table 2.12.

3. $T = \{F_0, F_1, F_2, F_3, F_4, F_5, F_6, F_9, F_{10}, F_{16}\}$

4. Flows F_i are randomly chosen:

   - $F_i = F_{10}$, $\Theta = \{F_{14}, F_{10}\}$, $T = \{F_1, F_2, F_3, F_4, F_5, F_6\}$.
   - $F_i = F_1$, $\Theta = \{F_{14}, F_{10}, F_1\}$, $T = \{F_5, F_6\}$.

99
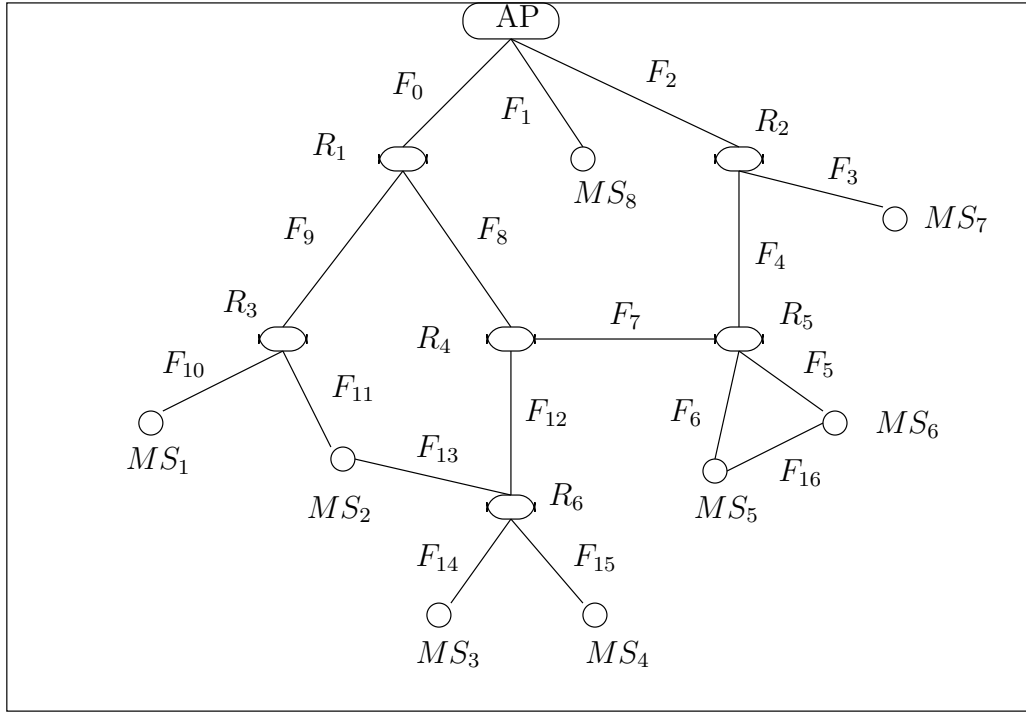
Figure 2.64: Example of multi-hop cloud with an AP, six relays, and eight MS.

- $F_i = F_5$, $\Theta = \{F_{14}, F_{10}, F_1, F_5\}$, $T = \emptyset$.

5. The link direction is chosen randomly: F_{14} is downlink, F_{10} uplink, F_1 is uplink, F_5 is downlink.

6. Interference at MS_3 is the sum of the powers received from MS_1, from the AP and from R_5.

Although the complexity of the algorithm is about $O(N^2)$, where N is the number of nodes in the ad hoc cloud, the time processing is reasonable in the practical scenarios. An example is given below.

Table 2.12: Evaluation of $H2(F)$ for multi-hop cloud.

Flows	$H2(F)$	d(F)
F_0	$\{F_1, F_2, F_3, F_4, F_7, F_8, F_9, F_{10}, F_{11}, F_{12}\}$	10
F_1	$\{F_0, F_2, F_3, F_4, F_8, F_9\}$	6
F_2	$\{F_0, F_1, F_3, F_4, F_5, F_6, F_7, F_8, F_9\}$	9
F_3	$\{F_0, F_1, F_2, F_4, F_5, F_6, F_7\}$	7
F_4	$\{F_0, F_1, F_2, F_3, F_5, F_6, F_7, F_8, F_{12}, F_{16}\}$	10
F_5	$\{F_2, F_3, F_4, F_6, F_7, F_8, F_{12}, F_{16}\}$	8
F_6	$\{F_2, F_3, F_4, F_5, F_7, F_8, F_{12}, F_{16}\}$	8
F_7	$\{F_0, F_2, F_3, F_4, F_5, F_6, F_8, F_9, F_{12}, F_{13}, F_{14}, F_{15}, F_{16}\}$	13
F_8	$\{F_0, F_1, F_2, F_4, F_5, F_6, F_7, F_9, F_{10}, F_{11}, F_{12}, F_{13}, F_{14}, F_{15}\}$	14
F_9	$\{F_0, F_1, F_2, F_7, F_8, F_{10}, F_{11}, F_{12}, F_{13}\}$	9
F_{10}	$\{F_0, F_8, F_9, F_{11}, F_{13}\}$	5
F_{11}	$\{F_0, F_8, F_9, F_{10}, F_{12}, F_{13}, F_{14}, F_{15}\}$	8
F_{12}	$\{F_0, F_4, F_5, F_6, F_7, F_8, F_9, F_{11}, F_{13}, F_{14}, F_{15}, \}$	13
F_{13}	$\{F_7, F_8, F_9, F_{10}, F_{11}, F_{12}, F_{14}, F_{15}\}$	8
F_{14}	$\{F_7, F_8, F_{11}, F_{12}, F_{13}, F_{15}\}$	6
F_{15}	$\{F_7, F_8, F_{11}, F_{12}, F_{13}, F_{14}\}$	6
F_{16}	$\{F_2, F_3, F_4, F_5, F_6, F_7, F_8, F_{12}\}$	8

Scenario and Performance Metric

For the simulations, an office building has been considered. Twenty terminals are assumed to be active. The AP and four relays form a string topology (figure 2.65). Other simulation parameters are shown on table 2.13. It is clear that the AP location is not an optimal choice. It has been deliberately placed at the extremity of the building to show that deep indoor coverage is possible with relay nodes.

The path loss model is the log-distance model previously described (section 2.2.2) with a carrier frequency of $f = 2.412$ GHz, and the link adaptation mechanism is used. Coverage prediction results from link budget calculations over a high number of iterations. Each iteration represents a photography of the observed area. Interfering MS positions, shadowing, and communication directions are randomly selected from one snapshot to another. At each iteration, C and CIR are computed on a grid of MS receivers in the network area.

A metric is introduced in [1]. This is a measure of the coverage gained thanks to the deployment of relay nodes. The reduction of non covered area (RNC) is the proportion of non covered area with an AP alone that is now covered thanks to the
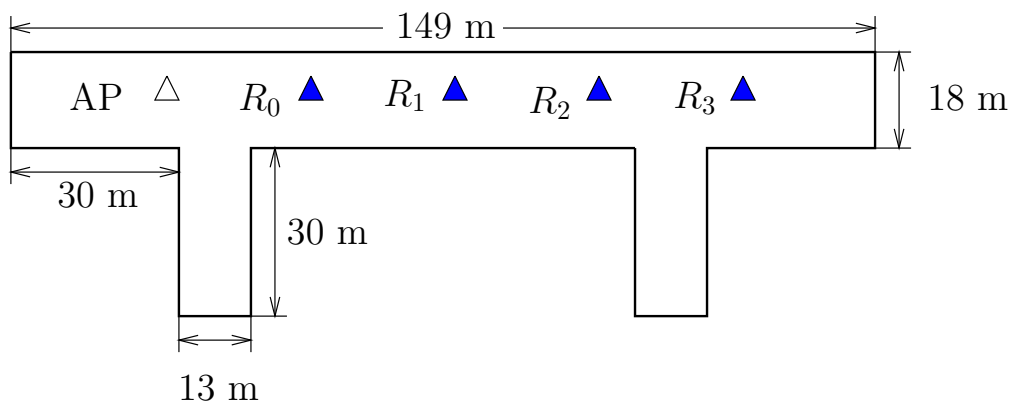
Figure 2.65: Scenario for coverage extension in an office building.

Table 2.13: Simulation parameters for coverage extension.

Parameter	Value
Transmit power	15 dBm
Transmit antenna gain	0 dBi
Receive antenna gain	0 dBi
Receiver sensitivity threshold	−94 dBm
Shadowing standard deviation	4 dB
Thermal noise power	−101 dBm
Uplink traffic load	30%
Downlink traffic load	70%
Number of MS	20

relay nodes:

$$RNC = (\%NC - \%NCR)/\%NC, \tag{2.9}$$

where $\%NC$ is the proportion of indoor area non covered by the AP alone and $\%NCR$ is the proportion of indoor area not covered by the AP with the relays. Note that $RNC = 1$, when the whole indoor area is covered.

Simulation Results

The percentages of not covered areas if relays are used or not are now compared. The software described in [62] provides coverage maps (C or CIR) resulting from the statistical average of the computations per snapshot. By applying the link adaptation mechanism to those average values, they are converted in expectable physical modes.
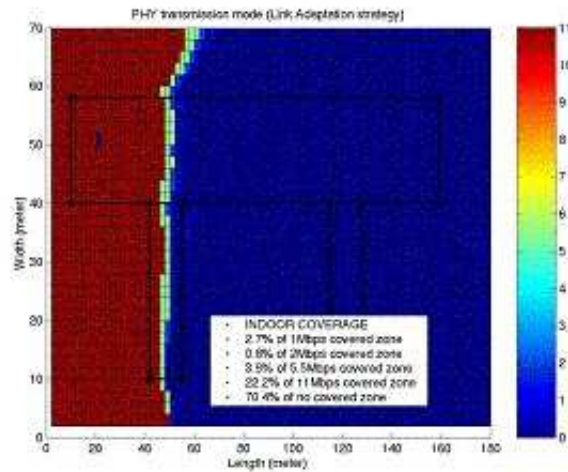
Figure 2.66: Physical modes map, a single AP without relay nodes.

By comparing the Physical modes maps (figures 2.66 and 2.67 are an example of snapshot) resulting from the two cases, AP alone and AP with relays, the RNC brought by the added relays is computed. In this example, $RNC = 0.95$. Moreover, the indoor covered area is tripled (29.6% without relay nodes versus 96.4% with relays). This result must be counter balanced by the fact that the user throughput is reduced when multiple hops are needed to reach the MS (downlink) or the AP (uplink). However, the cost of installation is reduced because relays do not need to be connected to the wireline network and the deployment is very flexible.

Note that there is only a small increase of the higher data rates coverage: The 11 Mbps physical mode covers 22.2% of the indoor area without relays and 24.1% with relays, whereas the 2 Mbps covers respectively 0.8% and 33%. In fact, 11 Mbps rates are almost not visible around relays. This phenomenon can be explained by the assumed routing algorithm that chooses routes with the minimum number of hops. Indeed, a MS close to the first relay is still in the communication range of the AP, and so will preferably be attached to the AP with a lower physical mode, rather than to the relay with a high physical mode. Note that a mono-frequency ad hoc cloud has been considered. A multiple frequency system is assumed in [31]. Let us now investigate what price has to be paid for coverage extension.

103
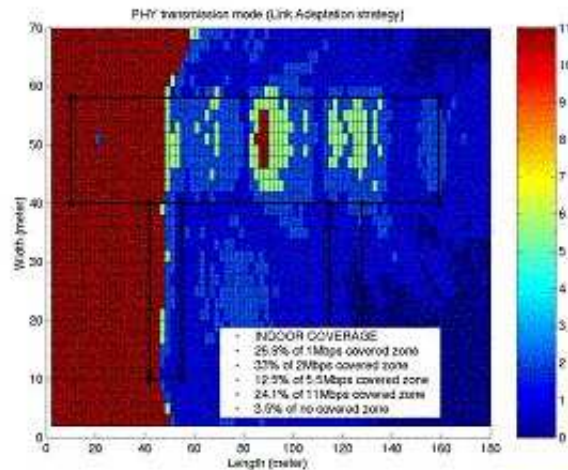
Figure 2.67: Physical modes map, an AP and four relay nodes.

2.6.3 Throughput Decrease and Fairness Issue

In this section, we show the throughput decrease induced by multi-hop communications and we underline the fact that IEEE 802.11 may not be an optimal solution in this case.

To illustrate the throughput degradation due to multi-hop connections, we first consider a string topology (figure 2.68) of nodes operating with IEEE 802.11 at 2 Mbps. The link adaptation is not taken into account in this simple simulation, all packets are sent at the data rate of 2 Mbps. The eight nodes are static and can communicate only with their direct neighbors. The carrier sensing range is twice that of the transmission range, i.e., a transmission interfere up to two hops away from the sender.
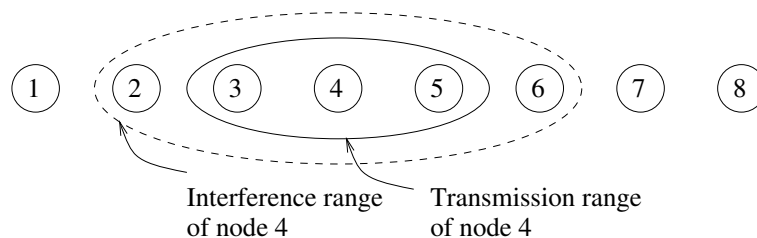
Figure 2.68: String topology with eight nodes.

Besides, a single constant bit rate source over UDP with 512 byte packets in node 1 is considered. The sending rate is chosen, so that the buffer of node 1 is

never empty, i.e., is backlogged. The destination is successively node 2, 3, 4... Thus, the distance between source and destination increases from one to seven hops.
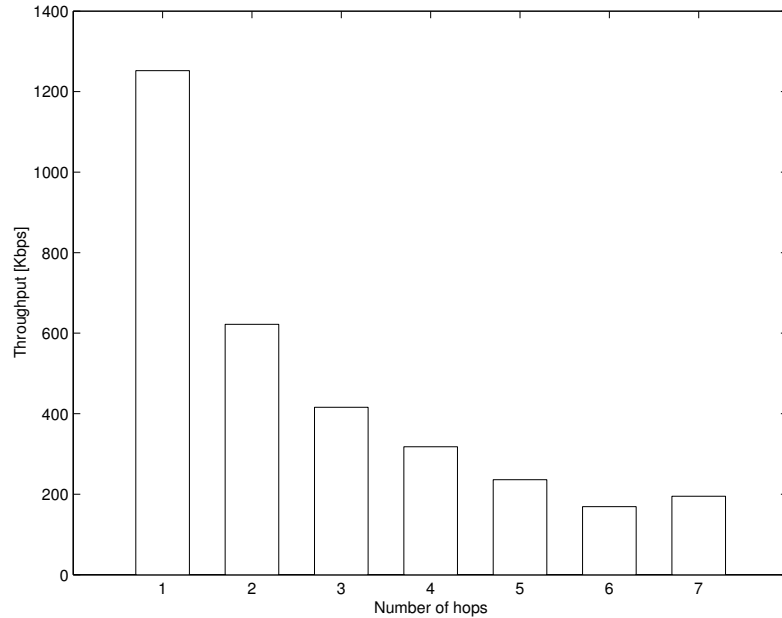
Figure 2.69: Throughput vs. number of hops in the string topology.
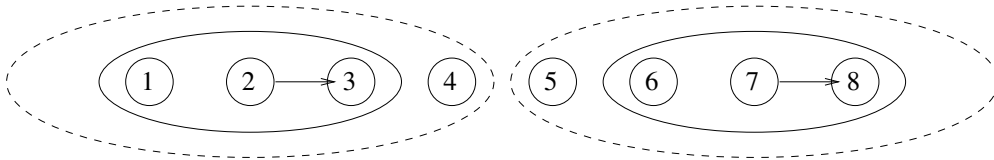
Figure 2.70: Spatial reuse of resources in the string topology.

In figure 2.69, the maximum achievable throughput is shown as a function of the number of hops. The main result is that until six hops, the capacity of the system is approximately divided by the number of hops. Only after seven hops, IEEE 802.11 is able to take advantage of the spatial reuse of resources: Simultaneous transmissions are possible without mutual interference, e.g. 2-3 and 7-8 in figure 2.70. In this case, the achievable throughput is slightly higher than for six hops. This decrease of the end-to-end throughput when the number of hops increases has also been observed by experimentation [26] for both TCP and UDP (figure 2.71).

These trends are confirmed if we consider a more realistic model for indoor environment and a single relay capable of relaying traffic to or from the AP, which is the scenario of section 2.6.2. A file transfer of 10 Mbytes from the AP to the terminal
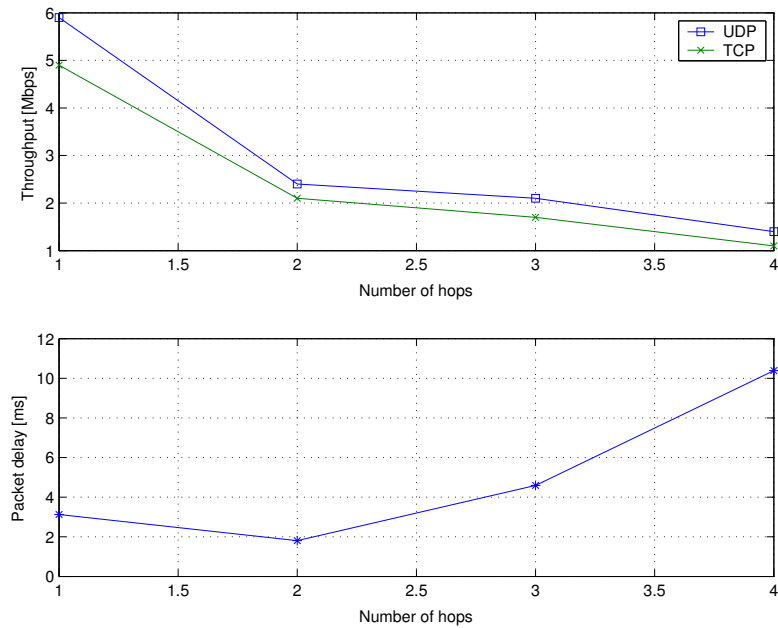
Figure 2.71: UDP and TCP throughput, and packet delay vs. number of hops, field experimentation.

is simulated. The terminal is moved in the direction of the relay (figure 2.72). The achievable throughput is evaluated as a function of the distance AP-terminal on the one hand, and as a function of the distance AP-relay on the other hand (figure 2.73). All nodes are running Dynamic Source Routing (DSR) as routing protocol. The advertized window of TCP is set to 64 segments of 1024 bytes.

Up to a distance AP-relay of 45 m, a clear increase of the covered area is observed. At 50 m, the data rate from the AP to the relay is too weak to enable an efficient relaying. Note that for a AP-relay distance between 15 and 35 m, there is an increase of the throughput between 45 and 50 m (AP-terminal). At 50 m, the route to the terminal goes preferably through the relay node, while at 45 m a direct link to the AP is preferred. In the first case, link adaptation allows two high data rate hops, while in the second case only a low data rate is used. This explains a better performance at a higher distance.

In figure 2.74, the 500 Kbps boundary of the cell is considered, i.e., the distance at which the achievable throughput is less than 500 Kbps. With this assumption, the optimal location of the relay node is around 35 m, where the gain is approximately 90% in distance in the direction of the relay.

Besides the throughput degradation due to the number of hops, performance
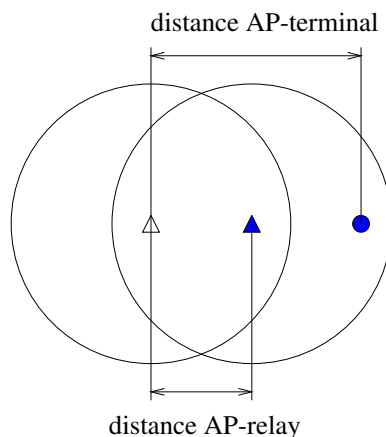
Figure 2.72: Topology for the throughput degradation study in presence of a relay node.

is further impacted by the exposed terminal problem which is not solved by IEEE 802.11. Figure 2.75 shows a typical exposed terminal topology with a 2 Mbps physical layer and backlogged transmissions. Although transmissions could be simultaneous in this scenario, the throughput achieved by IEEE 802.11 is approximately 1100 Kbps with 512 byte packets, i.e., less than for a one-hop transmission (see figure 2.69). This is due to the fact that in the exposed terminal topology, spatial reuse of resources is not exploited by the MAC protocol.

Let us investigate the fairness issue in a multi-hop environment with IEEE 802.11. We consider the string topology (figure 2.76), similar to that proposed in [202] to highlight the unfairness of IEEE 802.11. Nodes 0 and 4 transmit 512 byte packets at a constant bit rate of 1200 Kbps, i.e., beyond the channel capacity. Figure 2.77 shows the throughput of each connection as a function of the simulation time. We see that the link 4-2 is almost starved. DCF allocates few resources to the communication between nodes 4 and 2.

As a conclusion, using multi-hop networks is an interesting solution to extend the coverage of a WLAN cell. This extension is however obtained at the price of a throughput degradation. This lost is due to the fact that all transmissions are using the same shared medium. Thus, the channel is occupied n times, when a packet is relayed over n hops. IEEE 802.11 itself presents weaknesses: The exposed terminal problem is not solved and it exhibits fairness problems. However, we have identified two sources of capacity increase. Two high data rate links may be preferred to a single low data rate one. This is an interesting result when link adaption is used. Moreover, the spatial reuse of resources that allows simultaneous transmissions on

107
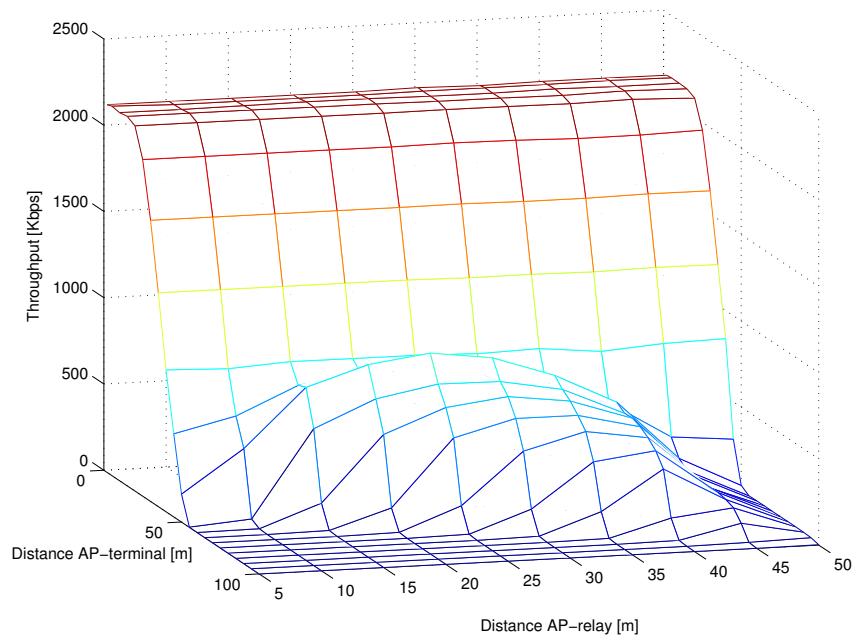
Figure 2.73: TCP throughput vs. distance AP-terminal and AP-relay.

the same channel can be exploited to increase the capacity of the network.
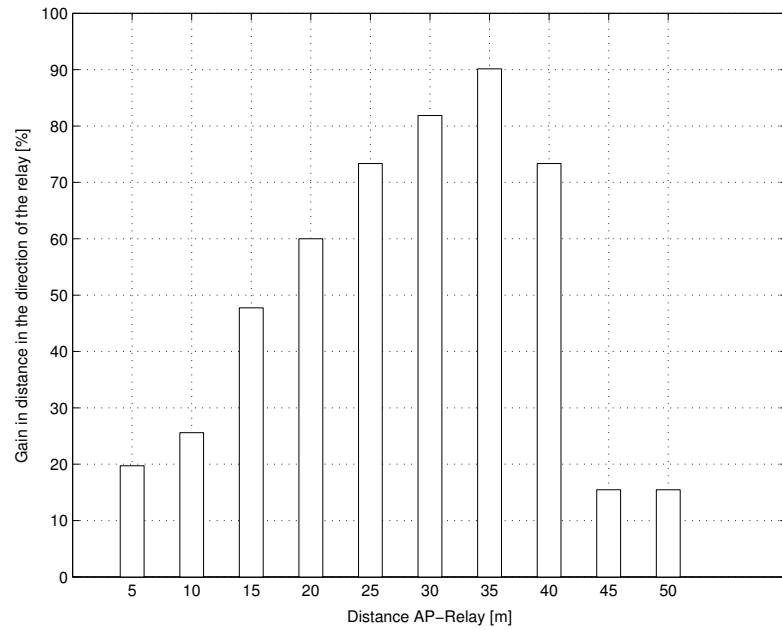
Figure 2.74: Gain in distance in the direction of the relay vs. distance AP-relay, cell boundary set at 500 Kbps.
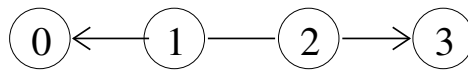
Figure 2.75: Exposed terminal topology.
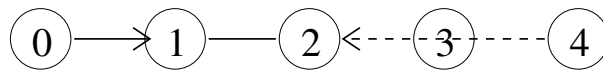
Figure 2.76: String topology with five nodes.
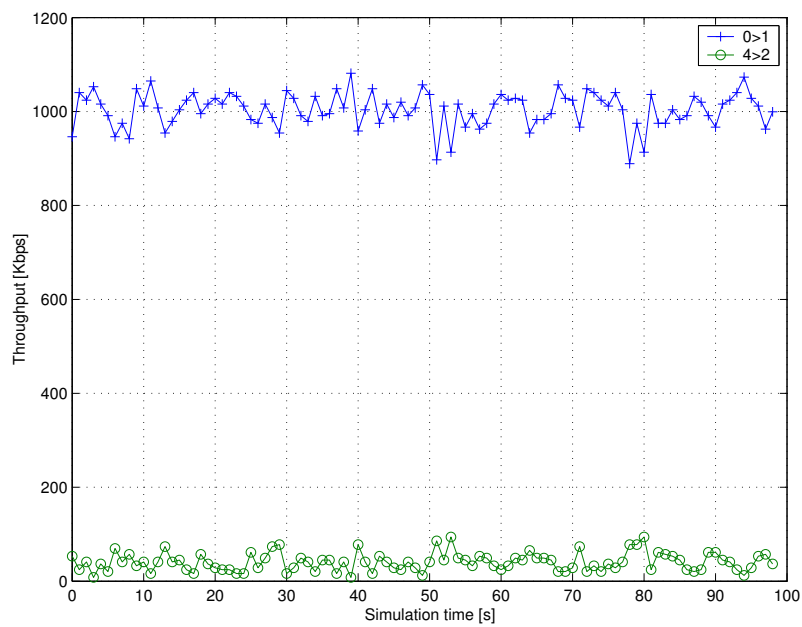
Figure 2.77: Connection throughput vs. simulation time for the five nodes string topology.

2.7 Conclusion

In this chapter, we have studied the performance of IEEE 802.11b DCF. In a first phase, we concentrated on single-hop networks, from an AP connected to an infrastructure network. Then, we considered the possibility to extend the coverage of this AP thanks to the multi-hop concept.

Let us recall the main conclusions of this capacity study:

> The DCF protocol is characterized by a high overhead that considerably reduces the throughput of higher layers with respect to the available physical data rate.

> IEEE 802.11 presents a very good performance for TCP based applications like FTP or WWW in terms of throughput and delays. Recall that the standard has been originally designed for this type of traffic.

> DCF has a poor performance with voice over IP traffic. This is mainly due to the overhead, that is particularly pronounced with small voice frames.

> A case study has shown that IEEE 802.11 can be deployed in rural areas to provide high speed Internet at a low cost. We have proposed dimensioning values for the satellite link and showed by simulations that up to sixty subscribers can access the Internet via an AP. This proves the viability of such a solution with respect to an ADSL deployment.

> Network designers should be informed of the near-far effect that reduces the aggregate throughput of a cell and can cause sharp decline of high data rate users. Three solutions have been proposed to combat this phenomenon: the relay and the TCP window based solutions for the downlink and the packet size adaptive solution for the uplink.

> Coverage extension has been studied as a low cost and easily deployable solution to extend the range of an AP. A specific algorithm for the computation of the carrier to interference ratio in an ad hoc cloud has been proposed in this context. This is an interesting application of the ad hoc concepts to AP centric networks.

> This study has shown that a trade-off has to be found by network designers between coverage and achievable throughput. Throughput is indeed degraded

111

as the number of hops to the AP grows. We have also seen that IEEE 802.11 exhibits weaknesses in multi-hop networks in terms of capacity and fairness.

Although the commercial deployment of IEEE 802.11b and g has been on going for several months, further improvements and differentiating technologies are of interest to network manufacturers. Thus, a lot of work remains:

- compare the presented performance with that of the other protocols of the IEEE 802.11 family, e.g., IEEE 802.11g, a, n

- study IEEE 802.11e that introduces service differentiation and thus is of particular interest for mixed traffic (TCP and UDP)

- propose methods for fine tuning of parameters of IEEE 802.11 and new algorithms for VoIP traffic. Among the possible solutions to improve the capacity, we can notice: frame concatenation, header compression, efficient use of RTS/CTS, adequate number of frame retransmissions

- study the hand-over issue in IEEE 802.11. Hand-over durations are indeed today not compatible with VoIP traffic

We have seen from the two preceding chapters that IEEE 802.11 provides satisfying performance in single-hop networks but is not an optimal solution for ad hoc networks. This is particularly true in multi-hop environments and at high input loads where enhancements are possible in terms of throughput and fairness. In the next chapter, we propose an alternative to IEEE 802.11.

Chapter 3

CROMA Protocol Description and Performance Evaluation

3.1 Introduction

With the growing interest of the research community for ad hoc networks, IEEE 802.11 has been used extensively as physical and MAC layer for the study of multi-hop communications. In particular, with the availability of low cost WLAN cards, most of the test-beds in this field use the IEEE standard.

Chapter 1 has shown that IEEE 802.11 is the heir of a family of protocols that are contention-based, as opposed to the category of conflict-free MAC protocols. Chapter 2 underlined the good performance of DCF in single-hop networks, especially for TCP based traffic. It has also been shown that the MAC layer of IEEE 802.11 may not be the optimal choice for multi-hop networks. And it leaves ample room for improvement in channel utilization and fairness, in particular at high input load.

In this chapter, we turn to the second family of protocols to improve the performance of IEEE 802.11 in ad hoc networks because conflict-free schemes may be preferred in heavy loaded scenarios. To further favor throughput thanks to spatial reuse of resources, we focus on a link allocation solution. Finally, the necessity to address varying topologies and traffic patterns gives rise to an adaptive slot allocation protocol. Our proposition is called CROMA for Collision-free Receiver-Oriented MAC. The detailed description of CROMA and the explanation of the acronym are given in section 3.2.

It is based on two main ideas. The first one is that TDMA based MAC protocols can provide a very good utilization of the shared radio resource, especially at high input loads. Node synchronization is however needed. The second idea is that the receiver communication range is the most relevant area of contention. This observation explains the receiver-oriented feature of CROMA. Receivers act as local and temporary base stations and can manage one or several communications on a single slot.

The analytical study presented in section 3.3 for a fully connected network confirms the improved channel utilization of CROMA. This trend and the good behavior of CROMA in terms of fairness are also shown through extensive simulations in multi-hop environments (section 3.4). Finally, a multi-slot extension of CROMA is presented in section 3.5 that makes the protocol independent on the frame length.

Most of the results of this chapter have been presented by the author in [3, 4, 5, 8].

3.2 Protocol Description

The Collision-free Receiver-Oriented MAC (CROMA) is a medium access protocol for mobile ad hoc networks that schedules transmissions in a slotted environment. It is a dynamic and distributed protocol that operates on a single-frequency channel. The utilization of omni-directional antennas is assumed.

In CROMA, time is divided into frames, further divided into a fixed number L of time-slots. Each slot can be temporarily and locally attributed to the receiver of a communication link to account for varying topology and traffic patterns. A receiver occupying a slot is allowed to poll several senders among its neighbors. The number of simultaneous communications for each slot is however limited by the protocol to a pre-defined value K.

CROMA is a receiver-oriented protocol since a slot in the frame is associated to a single receiver. CROMA is also collision-free because it ensures that unicast packets cannot collide.

CROMA doesn't rely on a traffic prediction algorithm at the receiver. Indeed, a requesting node has to reserve resources at its intended receiver during a random access phase. This reservation is needed only at the beginning of a packet train (or message). When a receiver has no longer traffic to poll, communications are released and the slot is free for another receiver.

3.2.1 Frame Structure

CROMA divides time into frames that are, in turn, divided into L equal time-slots. All mobile nodes are assumed to be perfectly synchronized.

Synchronization is a very critical issue for CROMA as for all distributed TDMA systems. However, as in [187] and [204], our focus is on protocol description assuming a perfect synchronization between nodes.

Throughout this chapter, the following terminology has been used: A *requesting node* is a node that has data packets to send but has not yet succeeded in the reservation phase, its *intended receiver* is the destination node of these data packets. A *sender* is a node which succeeded in the reservation phase and is ready to transmit data packets when polled by the receiver. A *receiver* is a node that polls senders on a given slot.

Each time-slot is divided in three parts: two mini-slots, called *REQ-mini-slot* (Request) and *RTR-mini-slot* (Ready-to-Receive) for signaling, and a *DATA-mini-slot* for data transmission (see figure 3.1).
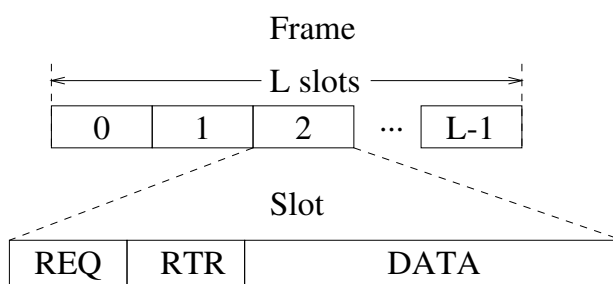
Figure 3.1: Frame structure of CROMA.

The REQ-mini-slot is used by requesting nodes during the random access phase for sending a REQ to its intended receiver. The RTR-mini-slot is used by the intended receivers to acknowledge requests as well as previous data transmissions, and to poll one of the senders having managed a successful reservation. During the DATA-mini-slot, the sender that has been polled in the preceding RTR-mini-slot transmits a data packet. These data packets are of fixed length, fragmentation and reassembly are done by higher layers.

With this frame structure, we prove in appendix D the correctness of CROMA, i.e. that data packets cannot collide.

3.2.2 CROMA from an Example

Before going into more detail in the protocol description, let us illustrate one of the key features of CROMA, i.e., to allow multiple reservations on the same slot. The receiver maintains a list of senders that managed a successful reservation and will poll them in the successive frames.

This feature is illustrated in figure 3.2, which shows two successive reservations on the same slot i. In a given frame j and slot i, node A sends a REQ packet with its address to node B which replies with an RTR containing a field to acknowledge the reservation (ackreq), and a field to poll node A (pol). The RTR is also received by node C that is now aware of a communication on slot i with B as receiver. During the data phase, A, that has just been polled by B, is allowed to transmit a packet to B with its address A and a sequence number (sn) 0. We say that B has got the floor on slot i. In frame $j + 1$, slot i, C sends a request for reservation to node B which acknowledges through a RTR the reservation with the field ackreq as well as the packet transmitted by node A in frame j. Simultaneously, node B polls node C which sends its first packet. In frame $j + 2$, B now polls A. With the RTR, it also acknowledges the data packet of C with sequence number 0. In frame $j + 3$, node B polls node C and acknowledges the data packet of A with sequence number 1, and so on.

3.2.3 The Choice of a Receiver-oriented Protocol

The choice of a receiver-oriented protocol is justified by the following arguments:

(i) This is a "natural" choice since only the zone that has to be secured with respect to collisions is the zone around the receiver, and thus, the spatial reuse of the radio resources is favored.

(ii) This choice allows the multiplexing of several communications on a single slot. That implies finer flow control and QoS negotiation. If a slot is associated to a sender, it cannot easily multiplex communications with different receivers since they may not be available because of a hidden terminal.

(iii) If a slot is associated to a receiver, a current communication on a given slot does not prevent a random access on this slot. More bandwidth for the contention for the channel implies less collisions and interference. If a slot is associated to a sender, it has to send at each frame a control packet (RTS) to give the address of
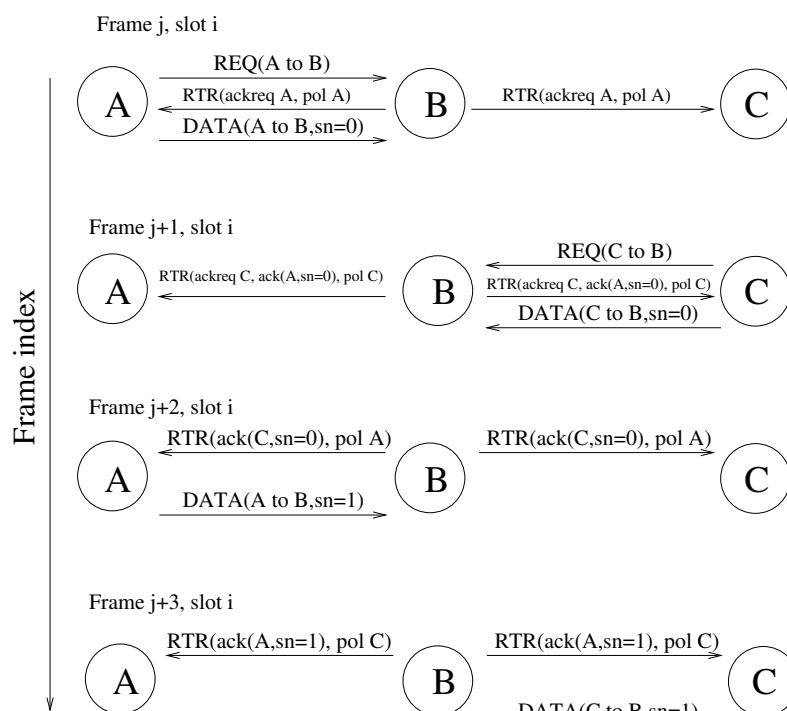
Figure 3.2: Example of two parallel connections on a slot with CROMA.

its intended receiver. Moreover, the receiver has to respond with another control packet (CTS) in order to avoid the hidden terminal problem. In CROMA, once the reservation has been achieved, the REQ is not used anymore for the duration of the communication, and the REQ-mini-slot can be used for new reservations.

3.2.4 Packet Formats

This section describes the different packet formats and the MAC header of the data packets. It gives also the definition of all the MAC fields. Their signification will be detailed in the protocol description (sections 3.2.5, 3.2.6, and 3.2.7).

Common Parts

The control packet formats and the MAC header of the data packets are shown in figure 3.3. Generic information, e.g. protocol version, is given in the *frame control* field (*fc*). The field *fcs* (frame check sequence) contains a CRC (cyclic redundancy code) calculated on all the fields of the MAC header and on the frame body. The field *source.ad* gives the Ethernet address of the packet source.
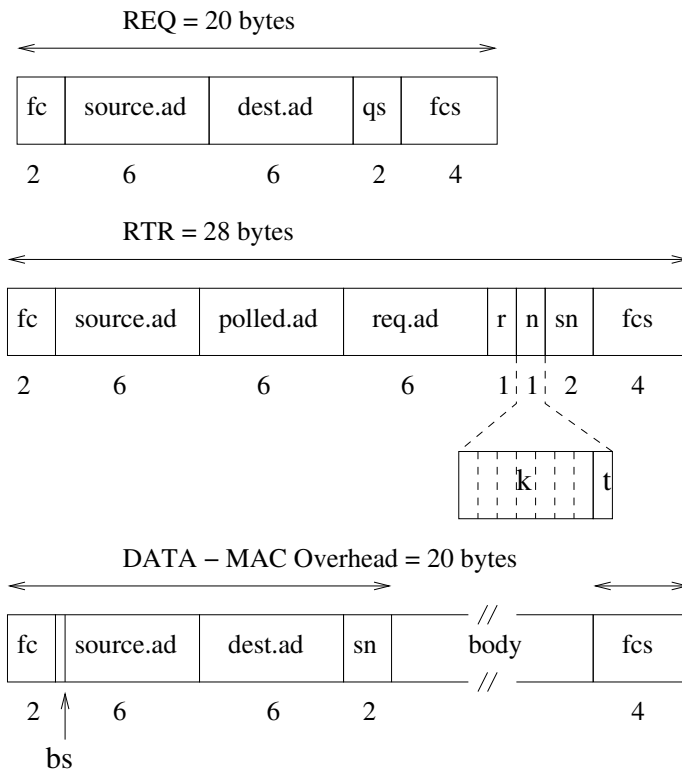
Figure 3.3: Packet formats of CROMA.

All packets, including data packets, have a pre-determined size, and each mini-slot is just long enough to allow the transmission of the associated packet. For example, the REQ-mini-slot allows the transmission of a REQ and includes the time for additional bits from the physical layer, the transmit-to-receive turn around time, and a small time interval to take into account the propagation delays. Note that it is preferred that the size of the control packets is short compared to the length of the data packets (e.g. 512 bytes).

REQ Control Packet

In a REQ, the field *dest.ad* gives the Ethernet address of the destination of the packet (the intended receiver). The field *qs* is used by a requesting node to indicate to the intended receiver the requested quality of service for the communication. This field may be used by higher layers to negotiate the QoS. It could be used in future versions of the protocol.

RTR Control Packet

A RTR has three different functions, as illustrated in section 3.2.2 and in figure 3.2: respond to a REQ, poll the different senders on the current slot and acknowledge data packets.

In the RTR, the fields *req.ad* and r are used to reply to the requests sent on the same slot (during the REQ-mini-slot). If a request is correctly received and accepted, it is acknowledged by putting the address of the requesting node in the field *req.ad* and the value ACK in the field r. If a request has been correctly received, but the communication cannot be established, the field r is set to $NACK$. This situation is possible if the requested QoS is not allowed or if the number of current communications has reached its maximum, K. If the receiver detects a collision of REQs, r is set to COL. If the receiver didn't received any request, or if the request cannot be decoded because of the channel conditions, r is set to $NOTRECVD$. The values $NACK$, COL, and $NOTRECVD$ are useful information for the requesting nodes to reschedule their requests.

The field *polled.ad* is used by a receiver to poll a sender that previously managed to establish a connection on this slot. If a sender reads its address in the field *polled.ad*, it is allowed to send a data packet during the DATA-mini-slot of the same slot, just after receiving the RTR.

The acknowledgment of data packets is done thanks to the field sn that stands for *sequence number*. Each node maintains a counter that is incremented for each new data packet. Receivers keep the last received sequence number. If in time-slot i of frame j, a receiver has received a data packet with sequence number m, it sets the field sn to m in the RTR of the slot i of frame $j + 1$ and so, acknowledges the previous data packet.

The byte n of a RTR gives information about the slot utilization. It is divided into seven bits that indicate the number k of current communications, and one bit t to inform that the receiver will not accept requests on this slot anymore. More details on the use of the bit t for fairness purposes are given in section 3.2.9. If k has reached the maximum K or if the bit t is set to 1, no more request can be done on this slot.

Data Packets

In data packets, the header includes a *buffer status* bit (*bs*) indication to inform the receiver about the sender buffer occupancy. *bs* is related to the multi-slot communication feature of CROMA that will be introduced in section 3.5. The field *dest.ad* gives the address of the packet destination.

As previously explained, each sender maintains a counter that is incremented for each new packet. This sequence number, put in the field *sn*, is used by the receiver to acknowledge the received packet. Let us recall that data packets have a fixed size resulting from higher layer segmentation or aggregation (512 bytes in simulations).

3.2.5   Reservation

Any communication between two nodes must be preceded by a preliminary reservation phase. In the reservation phase, requesting nodes contend to get access to a receiver. This random access, performed during the REQ-mini-slots, consists of five sub-phases: the listening of an entire frame, the time-slot selection, the REQ transmission on the chosen slot, the listening of the RTR, and the retry for a new reservation phase in case of failure (with or without random back-off). These five sub-phases are now detailed.

Frame Listening

The first phase consists in listening to the RTR-mini-slots during an entire frame, and recording the state of each slot. This listening process starts at the beginning of the reservation phase and lasts until a successful reservation.

A slot can be in several states:

FREE: No activity has been sensed during the RTR-mini-slot, i.e., no receiver has got the floor on this slot. A request will be possible on this type of slot.

OCC-A-COL-k: i.e., occupied, available, collision, and k communications. In this case, the *source.ad* of the RTR is the address of the intended receiver, a collision has been detected by the receiver during the REQ-mini-slot ($r = COL$ in the RTR), and there are currently $k < K$ communications on the slot. A request will be possible on this slot.

120

OCC-A-NCOL-k: i.e., occupied, available, no collision, and k communications. In this case, the *source.ad* of the RTR is the address of the intended receiver, no collision has been detected by the receiver during the REQ-mini-slot ($r \neq COL$ in the RTR), and there are currently $k < K$ communications on the slot. A request will be possible on this slot.

OCC-NA: i.e., occupied and not available. This is the case if a RTR has a *source.ad* different from the address of the intended receiver, if the requesting node detected a collision during the RTR-mini-slot, if it could not decode the field *source.ad* in the RTR, or obviously if the requesting node is itself a receiver on this slot. This is also the case if the field k of byte n has reached the maximum number of communications on a slot or if the bit t of byte n is equal to 1. A request won't be possible on this slot.

It is important to emphasize that the slot states are updated continuously during the whole reservation phase. In order to reduce the energy consumption, slot states updates can be limited to a few frames before starting the reservation procedure.

Time-slot Selection

The choice of the time-slot depends on the scheduling policy. This policy may have several objectives, e.g., maximize the slot utilization, limit the amount of interference in the network, establish connections which are robust with respect to node mobility. As we focus on highly loaded networks, we present here a simple policy that favors free slots first and therefore, aims at maximizing the slot utilization:

1. If there is at least one slot in state *FREE*,
   choose one randomly and exit, otherwise go to step 2;

2. If there is at least one slot in state *OCC-A*,
   select the slots having the lowest value of k. Among slots in this set:

   (a) If there is at least one slot in state *OCC-A-NCOL*,
       choose one randomly and exit;

   (b) Otherwise, choose one slot in state *OCC-A-COL* randomly and exit;

   Otherwise restart the reservation phase at the next frame.

REQ Transmission and RTR Generation

On the chosen slot, the reservation is done by sending a REQ during the REQ-mini-slot. Two cases must now be considered:

(i) The sender has chosen a free slot. If the intended receiver can decode the REQ, it replies to the request by sending an RTR in the same slot and by using the fields *req.ad* and *r* of this packet, as explained in section 3.2.4. Otherwise, the intended receiver doesn't reply. Note however that the intended receiver may be aware that the slot is occupied, which can happen in a hidden terminal configuration. In this case, the receiver doesn't answer.

(ii) The sender has chosen a slot that is already occupied by the intended receiver. In this case, the intended receiver replies with an RTR.

RTR Listening and Decision

A requesting node that has sent a REQ during the first mini-slot of the chosen slot listens to the following RTR-mini-slot. This implies that all nodes need to monitor the channel at any REQ and RTR-mini-slot in order to be able to answer an incoming request. Table 3.1 gives a summary of the decisions of the requesting node after the RTR-mini-slot. CROMA allows the receiver to reply with COL if the physical layer is able to provide this information.

If the field *req.ad* has been set to its address and r to ACK, the requesting node enters the transmission phase. If r indicates a collision, the random back-off algorithm is started. In all other cases, the requesting node is allowed to restart the reservation phase at the next frame. The random back-off algorithm is thus only used when a heavy load is detected for the intended receiver.

It can be noted that if a sender detects a collision during the RTR-mini-slot, it releases its current communication because this means that there is a conflict between two receivers for this slot (see section 3.2.7).

Back-off Algorithm

The back-off algorithm starts when a requesting node has been informed that a collision occurred. An integer BO is randomly chosen between 1 and $BACKOFFWND$. This is a timer that is decremented at the beginning of each frame and each time

Table 3.1: Decision of a requesting node after listening to the RTR-mini-slot.

Reception	req.ad	r	Decision
RTR decoded	my_address	ACK	enter the transmission phase
	my_address	$NACK$	retry on next frame
	not my_address	-	retry on next frame
	broadcast_address	$NOTRECVD$	retry on next frame
	broadcast_address	COL	start back-off algorithm
RTR not received nor correctly decoded	-	-	retry on next frame

the requesting node senses a slot in state *OCC-A* or *FREE*. As soon as *BO* reaches 0, a slot is chosen on the forthcoming frame according to the scheduling policy for a new request. With this algorithm, the load on the available slots is taken into account.

The parameter $BACKOFFWND$ is increased by a multiplicative factor at each successive retransmission and decreased by one at each success. However, there are a lower and an upper bounds for it, called *BOmin* and *BOmax*, e.g., 2 and 32. An adequate design choice for CROMA could be for the multiplicative factor 1.5 (also used by MACAW [54]) which is less than the value chosen by IEEE 802.11 (2) because with CROMA we expect less contention at high input loads. Moreover, the slot duration in CROMA implies a higher channel access time when the back-off algorithm is used.

Note that each node has a back-off algorithm instance for each destination in accordance with the queue structure presented below (see section 3.2.8).

3.2.6 Transmission

A sender starts the transmission phase immediately after a successful reservation. All receivers having reserved resources during the reservation phase poll the associated senders. When a sender recognizes its address in the field *polled.ad* of the RTR, it sends a data packet during the DATA-mini-slot.

Each sender maintains a counter of its transmissions that is incremented at each new packet. This sequence number is copied in the field *sn* of the packet header. The receiver is thus able to acknowledge the last correctly received data packet by copying its number in the RTR. The sender maintains the sent data

packet until its acknowledgment. If the next RTR is not received or if this RTR does not acknowledge the stored packet, a retransmission is necessary. After M retransmissions the stored packet is thrown away and the communication is released. This loss can be handled by an upper layer.

Figure 3.4 shows an example of a transmission phase with one receiver and three senders. It only shows slots i of successive frames. On the upper part of the figure, the RTRs of the receiver are represented with the fields *polled.ad* and *sn*. A cyclic polling is shown for the sender scheduling. *sn* field of data packets are shown too.

It is clear that on a given slot each receiver acts as a local base-station with respect to its associated senders. Thus, the polling mechanism allows a high flexibility for the scheduling of different flows by higher layers and provides a basis for the QoS algorithms. Moreover, several parallel communications are possible on a given time-slot.
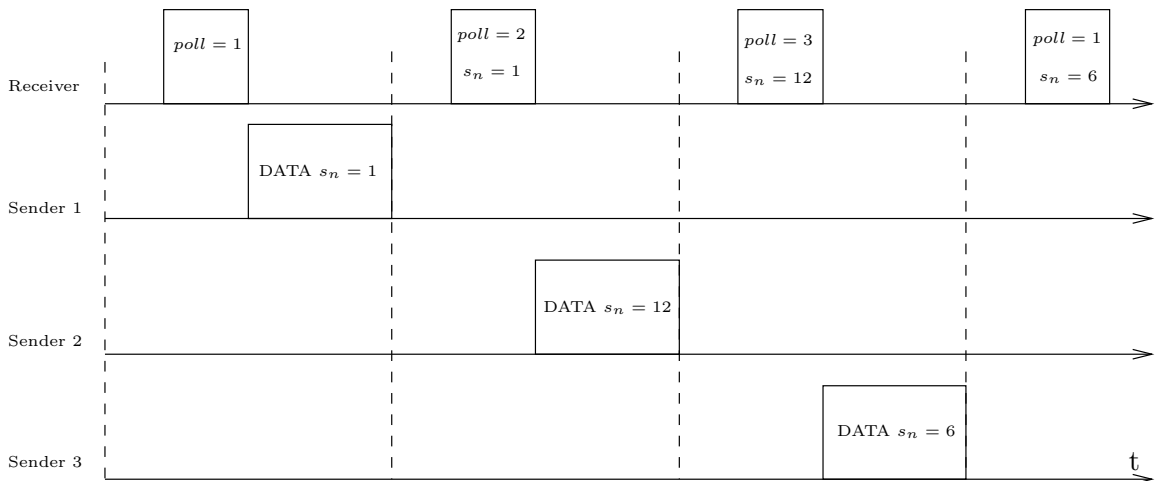
Figure 3.4: Polling during the transmission phase.

Note that the presented version of CROMA doesn't include any specific mechanism for broadcast packets. Section 3.2.10 proposes a simple solution for such packets.

3.2.7 Release

A transmission can be interrupted in one of the four following cases:

(i) The sender informs the receiver that it has sent the last packet by setting the

field *sn* of the packet header to the value *EOT* (end of transmission). If the last packet is correctly received, the receiver does not re-schedule the sender anymore. However, it acknowledges the last packet with its next RTR.

(ii) If a receiver polls W times a sender without receiving any packet, the communication is released. Also if a sender doesn't receive any polling on a given slot during W' frames, the communication is considered to be broken. In absence of QoS mechanism and thus with a fair polling, it is possible to choose $W = W' = K$.

(iii) During a communication, a sender may receive several RTRs, i.e., there is a collision of RTRs. In this case, the sender considers that the current communication on this slot is released.

(iv) After M retransmissions of a packet without acknowledgment, a sender considers that its communication with the intended receiver is broken.

3.2.8   Queue Management

CROMA includes a specific queue management that allows to take a full advantage of the slotted structure of the protocol. As shown in figure 3.5, data packets are sorted in each node according their next hop destination, i.e., there is one queue per potential receiver. One queue may be reserved for broadcast packets. These queues are either connected, i.e., a connection has been established with the receiver, or disconnected, i.e., a reservation phase is needed for that receiver if the node has a packet to send. The broadcast queue is always considered to be disconnected.

A connected queue is associated with a given time slot in the frame. The MAC layer monitors the RTR-packets on this slot and sends a data packet from the connected queue when the node is polled. When the connection is released, the queue is disconnected.

Data packets in the disconnected queues wait for a new connection with their respective destination. The MAC layer chooses one of these destinations and creates the corresponding REQ-packet that will be sent in the next frame. When a queue is connected, a new REQ-packet is created for it. An instance of the back-off algorithm is running for each non empty disconnected queue.
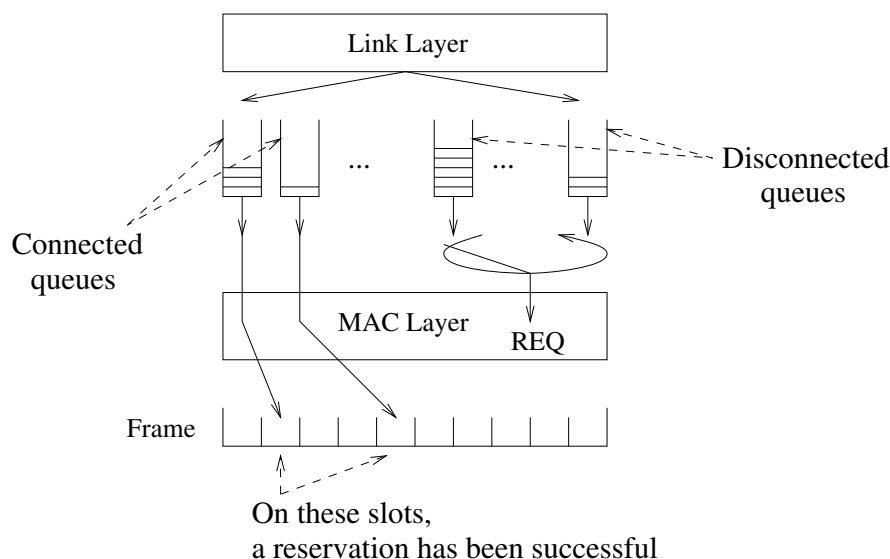
Figure 3.5: Queue management.

3.2.9 Fairness Issue

CROMA includes a mechanism to ensure a local fairness among data flows. On a given time-slot, fairness among incoming flows is ensured by the receiver of the slot by means of the RTRs based on different polling strategies.

However, if the number of slots in the frame is small compared to the number of potential receivers, situations of unfairness can arise and flows can be completely starved. The bit t in the RTRs is used to avoid such situations.

A receiver having the floor on a given slot counts the number of consecutive full frames. A frame is full from the point of view of a receiver, if it senses activity at each slot of the frame. In this case, it detects a potential blocking situation for pair of nodes that cannot communicate because there are no free slots anymore. If the number of monitored full frames reaches $MAX\_FULLFRAMES$, the receiver sets the bit t to 1 indicating that it will not accept new requests and that the current communications have to be released.

A sender detecting a bit t set to 1, sets the field sn of its next packet header to EOT and stops sending packets to the receiver. This release is done even if the sender has still packets to transmit. In this case, a new reservation is needed after a back-off. A requesting node detecting a bit t set to 1 in a RTR updates the slot state to OCC-NA.

Blocking situations leading to unfairness are avoided. However, since the re-

126

ceivers continuously monitor every RTR-mini-slot in the frame, a trade-off with
respect to energy consumption needs to be evaluated.

3.2.10   Broadcast Packets

In the current version of the protocol, there is no specific mechanism to send broad-
cast packets. Such packets have to be unicast to each neighbor of the sender. This is
not an efficient approach, especially to flood packets through the network as required
by many reactive routing protocols.

A simple solution to this problem with the current version of the protocol is the
following. A sender wishing to send a broadcast packet, selects a slot in the frame
according to a scheduling policy similar to the one described in section 3.2.5. Then,
it sends successively a REQ, a RTR, and the broadcast packet in the chosen slot as
shown in figure 3.6. The REQ includes in the header a field informing the neighbors
of the sender that a broadcast packet will be sent. Neighbors that successfully
receive the REQ respond with an RTR in the immediately following mini-slot.
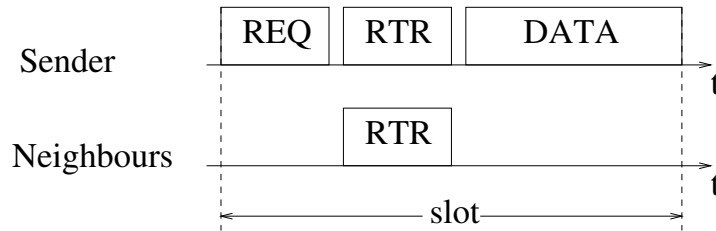
Figure 3.6: Broadcast packet.

This algorithm ensures that no collision between a broadcast packet and a unicast
packet can occur. Neighbors of the sender that decode the first REQ are not allowed
to transmit any packet in the slot. Moreover, all the on going communications on
the slot at a distance of one hop are released thanks to the RTR sent by the sender.
All communications at a two-hop distance on the slot are also stopped thanks to
the RTR sent by the neighbors of the sender. However, collisions between broadcast
packets are still possible.

It can be noted that the sender of a broadcast packet may choose an already
occupied slot for its transmission. In this case, the running communication will be
released. That means that broadcast packets have priority over unicast packets.
This is a reasonable assumption if we remember that broadcast packets are mainly
sent by the routing layer, e.g. for route discovery.

Note that this mechanism could be easily applied to IEEE 802.11 in order to reduce the number of collisions between unicast and broadcast packets. The performance evaluation of CROMA doesn't consider broadcast packets. Thus, this important aspect is left for further research.

3.3   Detailed Analytical Study of CROMA

In this section we calculate the aggregate throughput, i.e., the slot utilization of the protocol CROMA in a fully connected network. Following [187], it can be claimed that this topology is the worst case in terms of interference, contention, and spatial reuse.

3.3.1   Model for Slot Utilization Analysis

First of all, we describe our analytical model for the slotted MAC protocol CROMA. From this model we will derive the slot utilization of CROMA as a function of the probability p of request between a given source-destination pair.

1. We consider a fully connected network of N synchronized nodes.

2. All packets are of constant length and are transmitted over an assumed noiseless channel.

3. There are L slots per frame.

4. The maximum number of connections on a slot is K, i.e., when a receiver is already polling K different senders on a slot, no new REQ is allowed.

5. A receiver can only be associated with a single slot. This assumption can be in practice relaxed, but for the sake of tractability of the model, we limit the analysis to this case.

6. A node can be a sender on several slots of the frame (thus for different receivers). While being in communication on a slot, a node can send a REQ on another slot of the frame to start another connection.

7. The traffic between any two nodes s and d is an ON/OFF traffic.

8. The ON periods are modeled by bursts of packets following a geometrical distribution. The length of a message follows a geometrical law with parameter q. Thus, the average message length (AML) is $1/(1-q)$.

9. The OFF periods are modeled by series of slots without transmission following a geometrical distribution. If a source s doesn't communicate with a destination d, there is a probability p that s wants to communicate with d at the next frame.

10. A non-persistent policy is assumed for retransmissions after a failure, i.e., we can consider a fixed probability p to start a communication.

The system states are described by the vectors of the number of parallel connections on each slot at the end of the frame, $(a_0, a_1, ..., a_{L-1})$, where

- a_i is the number of current connections on slot i.

- $0 \leq a_i \leq MIN(K, N-1)$ (see assumptions 1 and 4).

Let S be the number of occupied slots in the frame (see assumptions 3 and 5):

$$S = \sum_{i=0}^{L-1} 1_{\{a_i > 0\}} \leq MIN(N, L) \ . \tag{3.1}$$

For the sake of simplicity, the states describe neither the receiver associated to each slot, nor the list of associated senders. The vector $(a_0, a_1, ..., a_{L-1})$ is a discrete-time stochastic process, whose state space is also discrete. Moreover, this process is independent of its history because of the memoryless property of the geometric law. Consequently, this process is a discrete time Markov chain (DTMC). Since the graph is aperiodic and finite, the chain is always ergodic.

From one frame to another, the possible transitions on slot i are:

- $a_i \rightarrow a_i + 1$ ($a_i < K$): A reservation has been successful on slot i AND no communication has come to an end.

- $a_i \rightarrow a_i$: (There is no successful reservation AND all existing communications continue) OR (there is a successful reservation AND this is the end of an existing communication).

- $a_i \rightarrow a_i - 1$ ($a_i > 0$): There is no successful reservation AND this is the end of a communication.

A transition probability between the two states $(a_0, a_1, ..., a_{L-1})$ and $(b_0, b_1, ..., b_{L-1})$ is assumed to be the product of the transition probabilities associated to each slot:

$$P\left((a_0, a_1, ..., a_{L-1}) \rightarrow (b_0, b_1, ..., b_{L-1})\right) = \prod_{i=0}^{L-1} P(a_i \rightarrow b_i). \qquad (3.2)$$

3.3.2 Analysis for $L = 1$

In this section, $L = 1$. In this simple case, we can derive a closed-form formula for the slot utilization.
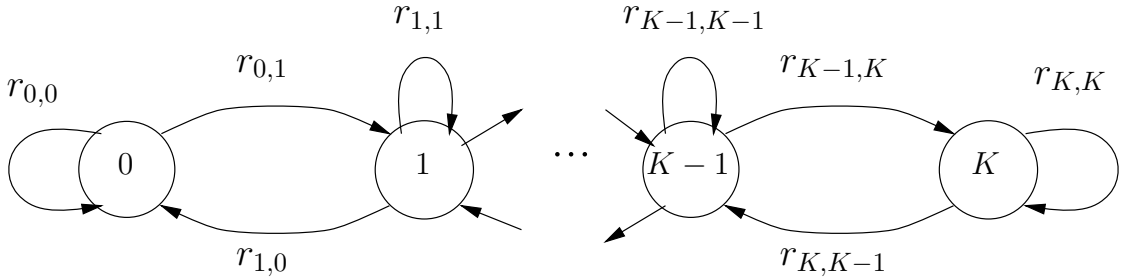
Figure 3.7: Discrete time Markov chain representing the state of the slot, for $K \leq N$.

The system is described by the number of parallel connections on the considered slot at the end of the frame (the DTMC is shown in figure 3.7). In order to compute the transition probabilities $r_{i,j}$ of this Markov chain, we notice that the probability for a source-destination pair to enter a ON period is p. Thus, the probability that a node sends a request on a free slot is the probability that this node has a request for at least one of the destinations:

$$p' = 1 - (1 - p)^{N-1} . \qquad (3.3)$$

Thus, on a free slot, a successful reservation occurs if and only if a single node among N is sending a request during the REQ-mini-slot. Consequently the probability to have a successful reservation on a free slot is:

$$\theta(0) = \binom{N}{1} p'(1 - p')^{N-1} . \qquad (3.4)$$

On an occupied slot with n connections, a receiver has got the floor on the slot and successively polls n senders that managed a successful reservation. Here, a successful reservation occurs if and only if a single node among the $N - (n + 1)$

nodes not currently in connection is sending a request. Therefore, the probability to have a successful reservation on an occupied slot is:

$$\theta(n) = \begin{pmatrix} N - (n+1) \\ 1 \end{pmatrix} p(1-p)^{N-(n+1)-1} . \tag{3.5}$$

In state $0 \le n < K$, there is a transition to state $n+1$ if and only if a successful request is received and this is not the end of a current communication. The transition state $r_{n,n+1}$ is thus given by:

$$r_{n,n+1} = \theta(n)q . \tag{3.6}$$

In state $0 < n < K$, there is a transition to state $n-1$ if and only if there is no successful request and this is the end of a communication, so:

$$r_{n,n-1} = (1 - \theta(n))(1 - q) . \tag{3.7}$$

From these two equations, we obtain directly $r_{n,n}$ for $0 < n < K$:

$$r_{n,n} = 1 - r_{n,n+1} - r_{n,n-1} . \tag{3.8}$$

In state 0, the slot is free and so $r_{0,1} = \theta(0)$ and $r_{0,0} = 1 - r_{0,1}$. In state K, $r_{K,K} = 1 - r_{K,K-1}$. The transition matrix is given by:

$$P = \{r_{i,j}\}_{0 \le i,j \le K} . \tag{3.9}$$

The steady state probabilities are obtained by solving the steady state equations $\vec{\pi} = \vec{\pi}P$, expressing all the probabilities as a function of π_0:

$$\pi_n = \frac{\pi_0}{1-q} \left[\frac{q}{1-q} \right]^{n-1} \prod_{k=0}^{n-1} \frac{\theta(k)}{1 - \theta(k+1)} , \tag{3.10}$$

for all $n \in \{1, \cdots, K\}$. The system is totally described with the normalizing constraint: $\sum_{n=0}^{K} \pi_n = 1$. The slot utilization of the protocol is given by $U = 1 - \pi_0$:

$$U = 1 - \frac{1}{1 + \sum_{n=1}^{K} \frac{1}{1-q} \left[\frac{q}{1-q} \right]^{n-1} \prod_{k=0}^{n-1} \frac{\theta(k)}{1-\theta(k+1)}} . \tag{3.11}$$

Figure 3.8 shows the slot utilization of CROMA, U, as a function of the probability p for $K = 3$, $N = 5$ and different average message lengths ($AML = 2$, 10 and

131

100 packets). Dotted curves have been obtained by simulations. These simulations reproduce the assumptions of our model. It is easy to conclude that: (i) The slot utilization with CROMA approaches 1 with increasing message length. (ii) The approximations of the analysis have little impact on the performance evaluation.
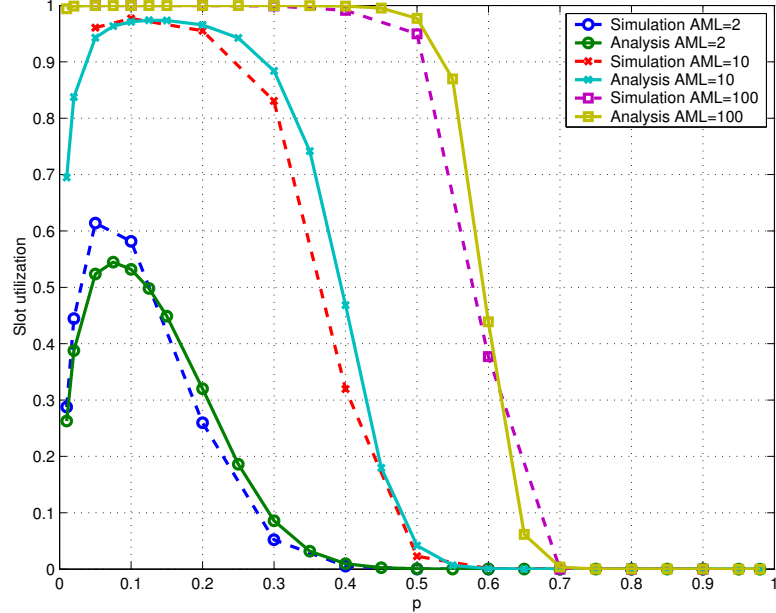
Figure 3.8: Slot utilization vs. input load, $L = 1$, $N = 5$, $K = 3$.

From the DTMC, the average number of connections, N_c on the slot can also be derived:

$$N_c = \sum_{n=1}^{K} n\pi_n \ . \tag{3.12}$$

Figure 3.9 shows the average number of connections for different AML values. This mean number is clearly related to the transmission delay. In fact, the higher the number of connections on a slot, the smaller is the resource allocated to a single connection. Thus, a trade-off has to be made between slot utilization and delay.

3.3.3 General Case Analysis

In this section, we extend the previous results to the general case with L slots per frame. We first compute the transition probabilities, while distinguishing an occupied slot, a free slot and a full slot. For the sake of clarity, we only consider the case $K \leq N$.

Let us consider a slot i occupied by the receiver d (this is the case, where $0 <$
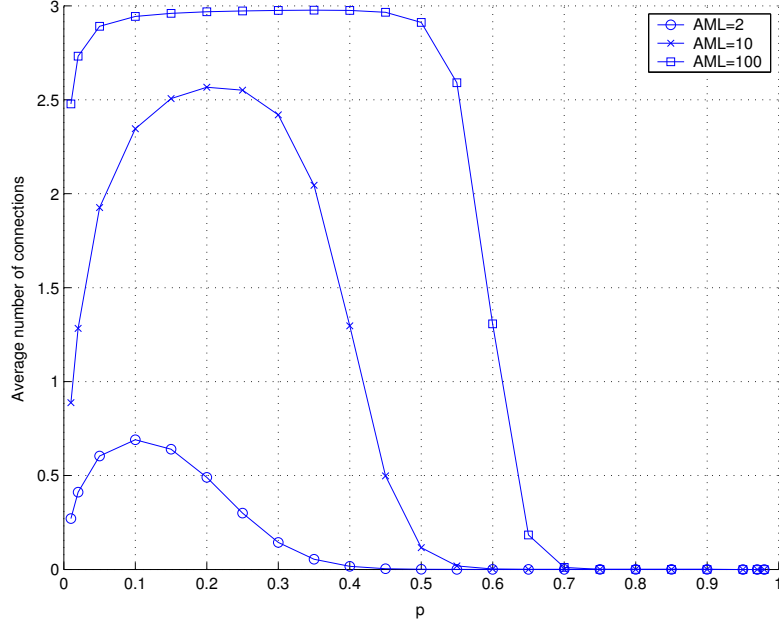
Figure 3.9: Average number of connections vs. input load, $L = 1$, $N = 5$, $K = 3$.

$a_i < K$). The number of nodes that are likely to send a REQ to d are nodes that are currently not in communication with d, their number is $N - 1 - a_i$. The probability for such a node s to send a REQ on slot i is p (see assumption 9). Thus, the probability of a successful reservation is:

$$\theta_i = \begin{pmatrix} N - 1 - a_i \\ 1 \end{pmatrix} p \left(1 - p\right)^{(N-1-a_i)-1} . \tag{3.13}$$

Note that if $K = N$ and $a_i = N - 1$, all nodes have a connection with the considered receiver, so that there is no REQ on this slot, and $\theta_i = 0$. The probability that a message is coming to an end is (see assumption 8): $1 - q$. We can now derive the transition probabilities for slot i:

$$P(a_i \rightarrow a_i + 1) = \theta_i q \tag{3.14}$$

$$P(a_i \rightarrow a_i) = \theta_i (1 - q) + q(1 - \theta_i) \tag{3.15}$$

$$P(a_i \rightarrow a_i - 1) = (1 - \theta_i)(1 - q) . \tag{3.16}$$

Let us now consider a free slot i ($a_i = 0$). There are $S = \sum_{i=0}^{L-1} 1_{\{a_i > 0\}}$ occupied slots in the frame corresponding to S receivers, since a receiver is associated to a single slot (see assumption 5).

133

On the considered free slot i, N senders are likely to send a REQ for $N - S$ possible receivers. Indeed, a node is allowed to send traffic to several receivers in parallel on different slots, so all nodes are likely to start a new communication on i. Moreover, requests on i can be addressed to any of the $N - S$ nodes that are not receivers on another slot because i is not attributed.

Considering a node s, the probability that s has n REQ for the $N - S$ possible receivers is

$$p_1(n) = \binom{N - S}{n} p^n (1 - p)^{N - S - n} \tag{3.17}$$

if s also belongs to the S receivers, and

$$p_2(n) = \binom{N - S - 1}{n} p^n (1 - p)^{N - S - n - 1} \tag{3.18}$$

otherwise. Thus, the probability that s has n requests is:

$$p(n) = p_1(n) \frac{S}{N} + p_2(n) \frac{N - S}{N} . \tag{3.19}$$

Now, the probability that s sends a REQ on the free slot i is:

$$\begin{aligned} \beta &= \sum_{n=1}^{N-S} Pr\big[\text{s sends a REQ on i}\big|\text{s sends n REQ}\big] p(n) \tag{3.20} \\ &= \sum_{n=1}^{N-S} min\left(\frac{n}{L - S}, 1\right) p(n) . \end{aligned}$$

Finally, there are N possible senders like s, so the transitions probabilities for i are:

$$\begin{aligned} P(0 \rightarrow 1) &= \binom{N}{1} \beta (1 - \beta)^{N-1} \tag{3.21} \\ P(0 \rightarrow 0) &= 1 - P(0 \rightarrow 1) . \tag{3.22} \end{aligned}$$

For a full slot $(a_i = K)$, the transition probabilities are obvious:

$$\begin{aligned} P(K \rightarrow K) &= \theta_i(1 - q) + q(1 - \theta_i) \tag{3.23} \\ P(K \rightarrow K - 1) &= 1 - P(K \rightarrow K) . \tag{3.24} \end{aligned}$$

The steady state equations $\vec{\pi} = \vec{\pi}P$ can be solved using any numerical method, e.g., the iterative method of Gauss-Seidel (see [51] or [183]).

Figure 3.10 shows the slot utilization of CROMA as a function of p for different average message lengths. Analysis and simulations (dotted lines) are presented and the figure shows that the two approaches match pretty well. As for $L = 1$, we can see that CROMA can achieve very high slot utilization for high AML. Note that values of p near 1 are not realistic in a real implementation because of the back-off algorithm. This mechanism has indeed the effect of reducing the probability of request. Figure 3.11 shows the influence of K on system performance. There is a
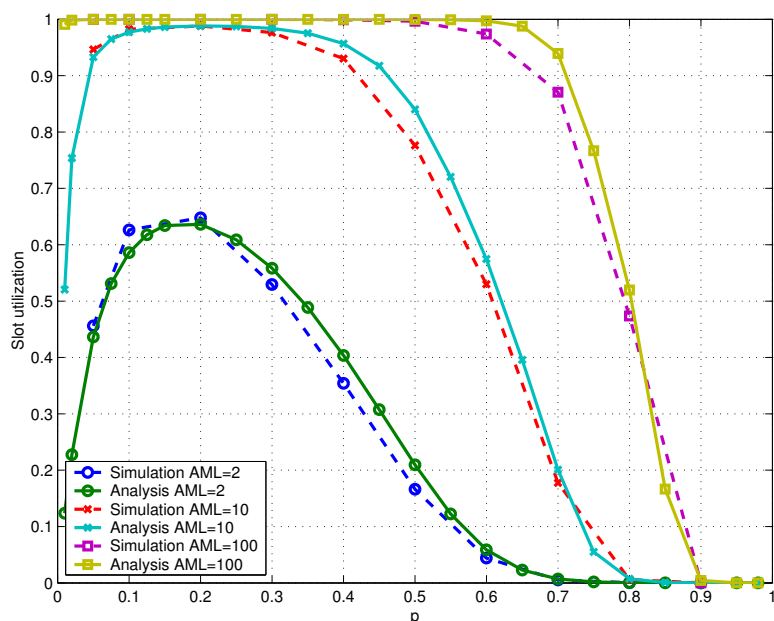
Figure 3.10: Slot utilization vs. input load, $L = 3$, $N = 5$, $K = 3$.

clear gain of channel utilization as K increases. However, this is obtained at the cost of increased delays. This is shown in figure 3.12, where the average number of connections per slot is plotted. A higher number of connections per slot implies a higher delay for the burst transmissions.

Through the analysis presented above, we have been able to elucidate further the results already presented by the author in [8]. The conclusions of the analysis are given below.

> The presented analysis provides an efficient tool for the channel utilization study. Theoretical results have been validated by simulations.
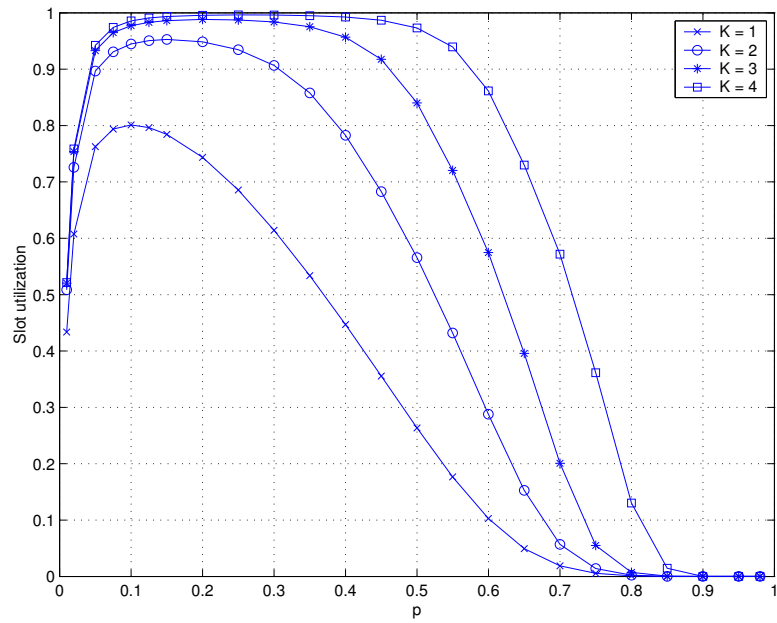
Figure 3.11: Slot utilization vs. input load, influence of K, $L = 3$, $N = 5$, $AML = 10$.

> The slot utilization of CROMA approaches 1 with increasing message length.

> Increasing the number of possible communications on a slot provides better slot utilization at the expense of packet delays.
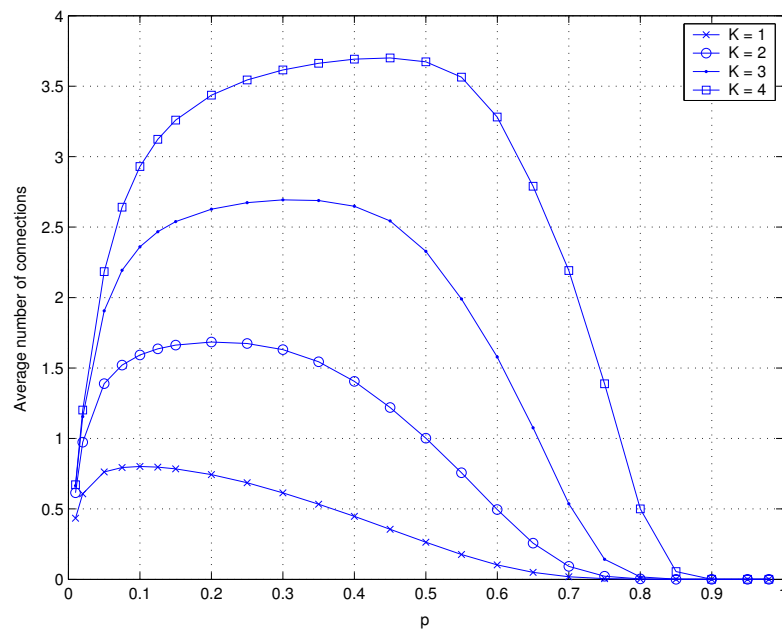
Figure 3.12: Average number of connections vs. input load, influence of K, $L = 3$, $N = 5$, $AML = 10$.

3.4 Multi-hop Networks

In this section, the performance of CROMA and of IEEE 802.11 (DCF mode) are compared through simulations.

3.4.1 Methodology

Studying MAC protocols in a multi-hop environment leads to the problem of choosing an appropriate node topology. Literature on ad hoc networks has solved the issue by considering on the one hand typical networks, like the string network [202], or the grid network [186], and on the other hand randomly generated networks, static [139, 204] or mobile [69]. In this chapter, we adopted a dual approach whereby CROMA is evaluated in conjunction with a network having a classical and challenging topology, and networks with random topologies.

The metrics used to evaluate the performance of the MAC protocols are:

End-to-end delay: This is the average time spent by a packet from the traffic generator of a source to the reception module of the destination.

End-to-end delay jitter: This is the standard deviation of the end-to-end packet delay.

Aggregate end-to-end throughput: This is the average number of bits successfully received by all nodes in the network per second. The input load is the average number of bits transmitted by all nodes per second. If the system is stable the aggregate throughput equals the input load.

Fairness index: This is the widely used index, f, defined in [118]. If a system allocates resources to n contending entities, such that the ith entity receives an allocation x_i, then:

$$f(x) = \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n \sum_{i=1}^{n} x_i^2} \ . \tag{3.25}$$

If all entities get the same amount, i.e., x_i's are all equal, then the fairness index is 1 and the system is 100% fair. Basically, the adequate selection of entities and x_i's is application dependent. In our case, entities are flows of data between source-destination pairs (i, j) and x_i's are their throughput, $T_{i,j}$.

The traffic is ON/OFF with exponential distributions, and the packet size is set to 512 bytes. Moreover, the channel is supposed to be perfect with a physical data rate of 2 Mbps. The transmission area of a node is a disk of radius R. Outside of the transmission area no communication is possible. Simulations have been done using ns2. The finite state machine describing CROMA is given in appendix E. The simulation parameter values are presented in table 3.2. Note that the mean OFF time is fixed and that the mean ON time will vary in simulations. Note also that the inter-mini-slot time and the physical layer overhead are consistent with the IEEE standard. Simulations have been conducted to ensure statistically significant results (90%-confidence intervals have a length typically less than 10% of the metric mean value).

Table 3.2: Main parameter values for simulations.

Parameter	Value
DATA Packet size	512 bytes
BOmin	2
BOmax	64
K=W=W'	3
M	7
MAX_FULLFRAMES	30
Inter-mini-slot time	10 μs
PHY overhead	24 bytes
PHY Data Rate	2 Mbps
ON distribution	Exponential
OFF distribution	Exponential
Peak Rate	256 Kbps
Mean OFF time	0.5 s
Simulation time	200 s
Number of simulations per point	10

3.4.2 A Challenging Topology

Throughput and Delay Analysis

In order to evaluate the performance of CROMA in ad hoc networks, we consider a very simple multi-hop situation that has been used in the literature for the evaluation of MAC protocols, e.g. in [95].

In this configuration, eight static nodes form a regular topology, flows of data

139

are shown on figure 3.13. Four end-to-end communications are running in parallel: 0-1-2-3, 0-5-2-7, 7-6-5-4, and 3-6-1-4, so that several nodes have to receive and/or to relay several flows of data. A line without arrow between two nodes means that they are in the communication range of each other, i.e., the transmissions from one of them can be successfully decoded by the other one. A line with arrow means that at least one flow of data is using this link.
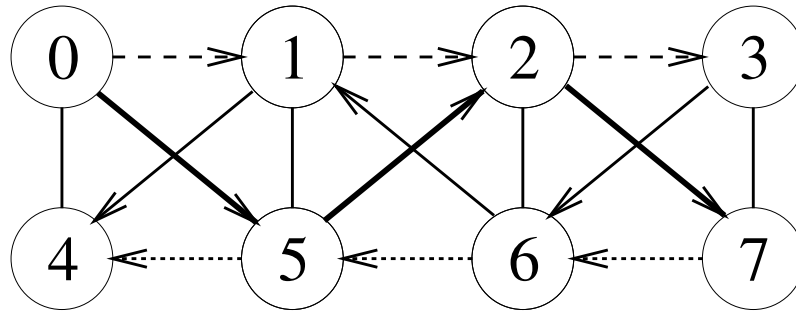
Figure 3.13: A multi-hop topology, the "squares topology".

This configuration is interesting for several reasons:

(i) It exhibits a lot of hidden terminal situations. For example, nodes 6 and 2 are hidden from node 0, nodes 7 and 3 are hidden from node 5.

(ii) Spatial reuse is possible and there are situations of exposed terminal. For example, nodes 1 and 2 are exposed. Several flows can share the same slot, e.g., 1-4 and 2-7, or 2-3 and 5-4.

(iii) Nodes and flows experience different contention situations, nodes 0, 3, 4, and 7 have three neighbors, while nodes 1, 2, 5, and 6 have five neighbors.

Figures 3.14 and 3.15 show the end-to-end packet delay and jitter as function of the input load for IEEE 802.11 and CROMA with two scales (the lower set of curves is a zoom in the low load zone). The different curves for CROMA assume different number of slots per frame.

In the case of low input loads, IEEE 802.11 outperforms CROMA because the low level of contention implies a small number of collisions and small back-off windows. Note that in case of collision, back-off intervals are much smaller with the IEEE 802.11. The timer is indeed decremented every 10 μs instead of every slot with CROMA. At this level of load, the network cannot fully take advantage of the reservation scheme because trains of packets are small.
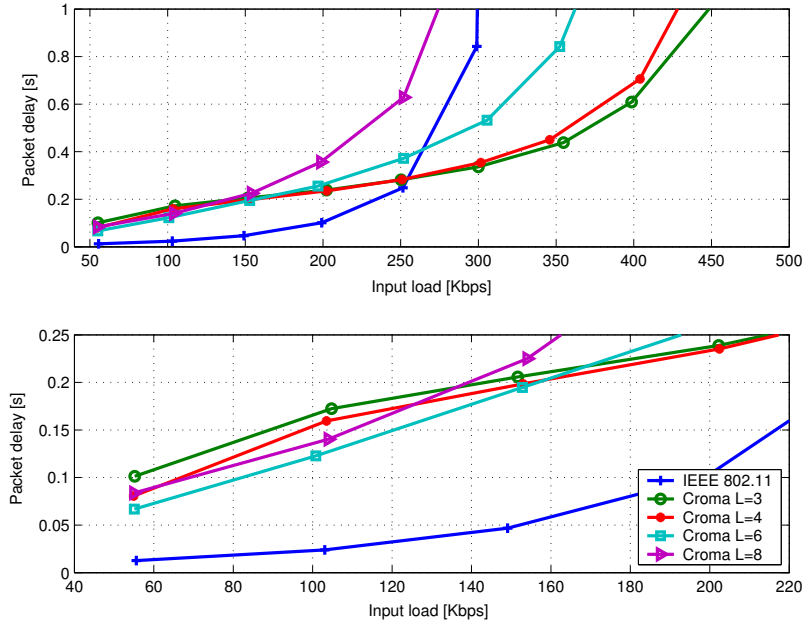
Figure 3.14: End-to-end delay vs. input load, squares topology.

In the case of higher input loads, i.e., above 250 Kbps, IEEE 802.11 nodes experience more contention, and thus more collisions and wider back-off windows: The access delay increases dramatically. On the other side CROMA takes advantage of packet bursts to reduce the number of requests per transmitted packet. If a flow has made a successful reservation, long trains of packets can be transmitted without contention.

Delay and jitter for CROMA $L = 8$ are always greater than for IEEE 802.11. It is clear that CROMA $L = 8$ is not well dimensioned for the considered topology. Actually, the number of slots is too high and the resource is not fully exploited. The multi-slot communication of CROMA presented in section 3.5 aims at solving this kind of effect. Figure 3.16 shows that the slot utilization of CROMA $L = 8$ does not exceed 0.75. This is much less than CROMA $L = 6$ that reaches 1. CROMA $L = 3$ and 4 fully exploit spatial reuse and exceed 1.1.

The reservation scheme, the synchronization, and the ability of CROMA to handle the exposed terminal problem allow the network to achieve high throughputs. Figure 3.17 shows the aggregate throughput as a function of the input load. IEEE 802.11 saturates at a throughput of 300 Kbps. In comparison, CROMA $L = 8$ achieves a maximum throughput of 350 Kbps, although we have seen that it is obviously badly dimensioned for the topology. CROMA $L = 6$ reaches a maximum throughput of 450 Kbps. For less slots per frame, a problem of stability of the
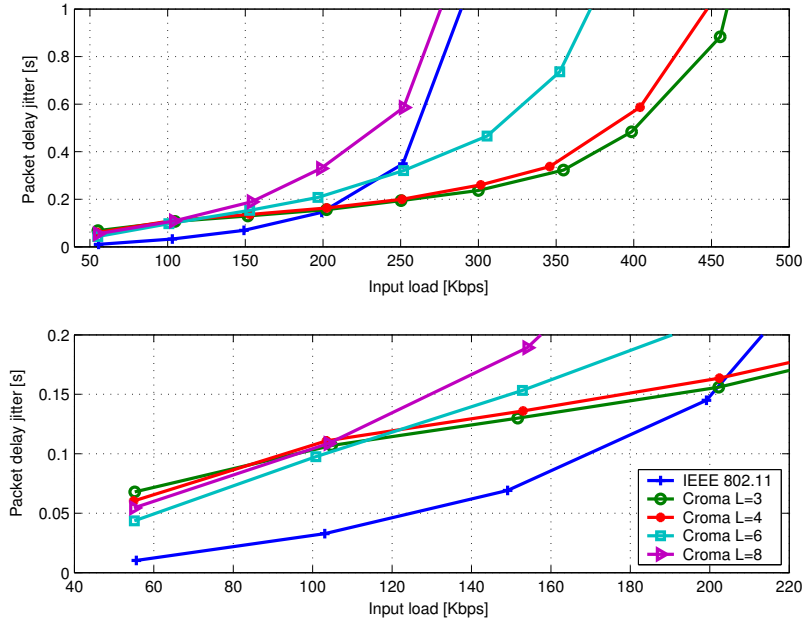
Figure 3.15: End-to-end delay jitter vs. input load, squares topology.

throughput arises.

Although CROMA $L = 3$ and 4 achieve respectively 475 and 500 Kbps, the throughput slightly decreases for input loads higher than 550 Kbps. The reason is that the small number of slots (with respect to the studied topology) implies a slight instability due to the fairness strategy. This strategy indeed imposes frequent communication breaks at high input loads and the frequency is higher when the frame is short. However, curves show a slow decrease leading to acceptable values even at high input loads.

In figure 3.17, the long-term point of operation of a network should be on the first bisecting line. CROMA follows this line far beyond IEEE 802.11. When the throughput curves go away from this line, the network has reached its saturation throughput and is now unstable, i.e., buffers increase indefinitely. The stability of the protocol with respect to peaks of input load can be observed beyond the saturation point.

Note that the gain of CROMA over the standard is expected to be reduced if we consider a more realistic channel model. A widely used model assumes an area of interference beyond the transmission range. In this area, a signal cannot be decoded but can cause a collision. This effect reduces the advantage of solving the exposed terminal problem.
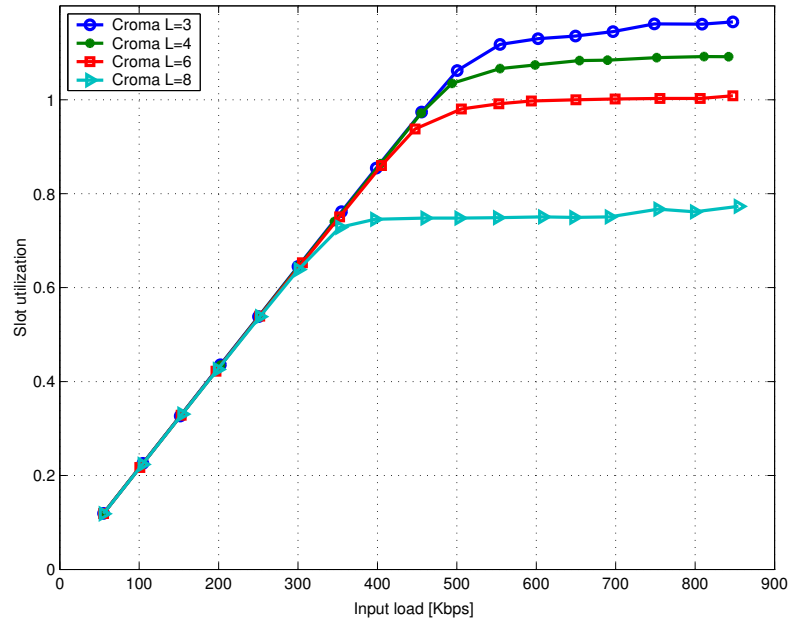
Figure 3.16: Slot utilization vs. input load, squares topology.

At this point, three of the main characteristics of CROMA have already been illustrated:

> CROMA provides very good throughputs at high input loads.

> IEEE 802.11 outperforms CROMA in terms of delay/jitter at low input loads, but not at high input loads.

> The performance of CROMA is dependent on the frame length. This effect is handled by the multi-slot communication feature of CROMA which allows an opportunistic use of free slots (see section 3.5).
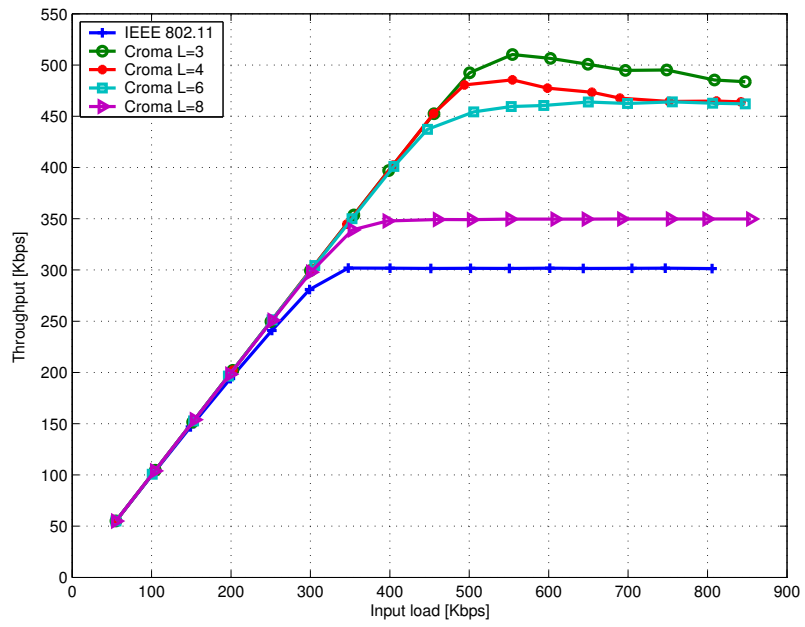
Figure 3.17: Throughput vs. input load, squares topology.

Fairness Analysis

As explained in section 3.2.9, the use of the bit t has been foreseen in CROMA for designing fairness strategies for resource sharing. This is expected to be of particular help when the network input load is high and the number of slots per frame is small for the considered topology/traffic pattern. For example in the topology of the figure 3.13 with $L = 4$, if node 1 hears the RTRs of node 2 on slot 0, node 5 on slot 2, node 6 on slot 3, and sends RTRs on slot 1, 1 cannot send any REQ to 4 since the frame is full. In case of low input loads, this situation is transient and has a low impact on the long-term fairness. In case of high input loads however, the connection 3-4 is completely starved leading to unfairness.
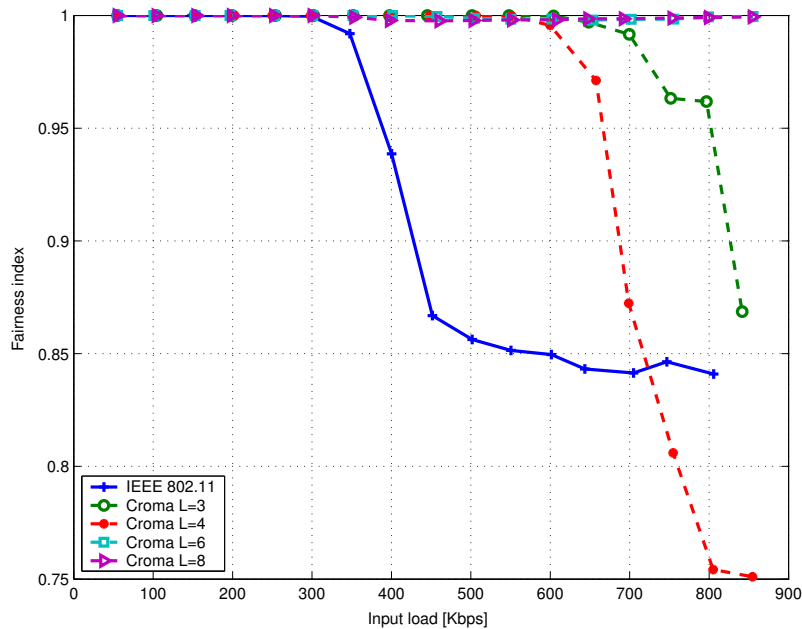
Figure 3.18: Fairness index vs. input load, without bit t, squares topology.

Figures 3.18 and 3.19 compare two strategies using or not the bit t and the fairness indexes of CROMA with different values for L are compared to that of IEEE 802.11.

Two extreme situations for the behavior of CROMA in terms of fairness are shown. On the one hand, if the frame length is over-dimensioned, e.g., $L = 6$ and $L = 8$, the fairness strategy has no effect on the fairness index. Each receiver has locally the possibility to use a slot quite regularly. Moreover, some slots are eventually left unused, even at high input loads. In this case, the fairness index is always above 0.99.
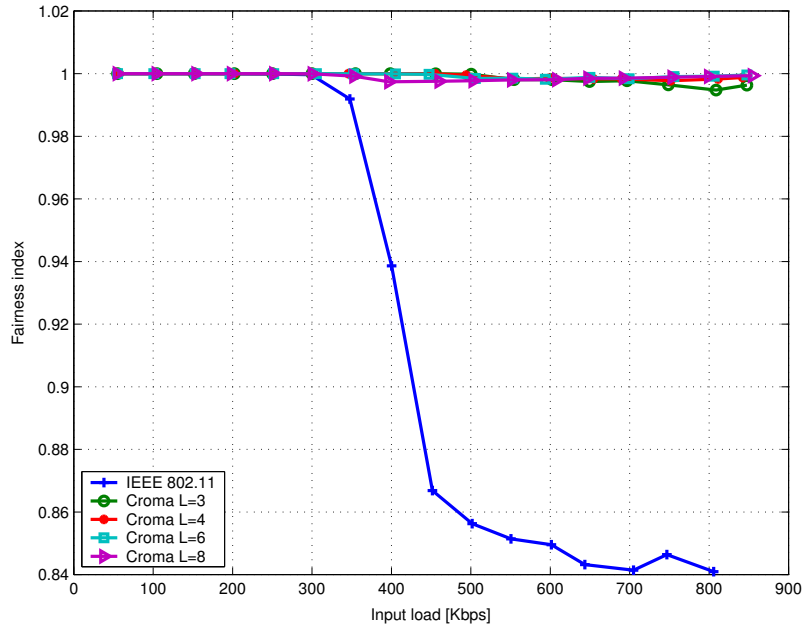
Figure 3.19: Fairness index vs. input load, with bit t, squares topology.

On the other hand, when the number of slots per frame is scarce, e.g., $L = 3$ and $L = 4$, the fairness strategy plays a decisive role. For the IEEE standard and for CROMA without the use of the bit t, the index is close to 1 for low to moderate input loads. After a threshold, the increase of input load leads to a drop of the index. This threshold is 350 Kbps for IEEE 802.11, and approximately 650 Kbps for CROMA. With the use of the bit t, the fairness index of CROMA remains always above 0.99 for both $L = 3$ and $L = 4$.

What is the price to pay for such a result? About 10% of maximum throughput and slightly higher delays at high input loads. In figure 3.20, the achievable throughput drops from 580 Kbps to 510 Kbps for $L = 3$, and from 540 Kbps to 485 Kbps for $L = 4$. The fairness mechanism brings however stability to the protocol by eliminating situations of blockage. Higher delays are also observable in figure 3.21.

How can we interpret the relative degradation of performance in presence of the bit t? The explanation can be found in figures 3.22 and 3.23, which provide the number of REQ collisions as a function of the input load. As expected, the fairness strategy has little impact on CROMA $L = 6$ and $L = 8$ (curves are superimposed for $L = 8$). We observe a first phase of increase due to the fact that more input load means also more reservations. Then, in a second phase, nodes take advantage of the reservation process, less REQ are needed, and packet trains are longer and longer.

146
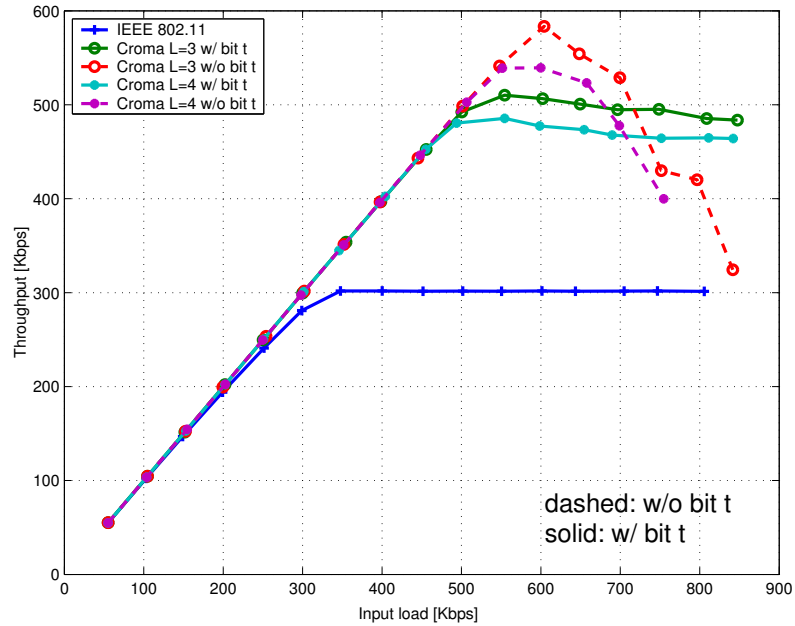
Figure 3.20: Throughput vs. input load, $L = 3$ and $L = 4$, influence of the bit t, squares topology.

On the contrary, with $L = 3$ and $L = 4$, the number of slots is small and not sufficient to allow each receiver to have the floor on a slot, especially at high input loads. In this case, the fairness strategy plays a decisive role to periodically redistribute the slots. This mechanism artificially keeps the average message length constant and implies new reservations.

Up to now, the parameter $MAX\_FULLFRAMES$ has always been set to 30 although it has also an influence on the performance of CROMA. If a receiver locally senses a frame fully utilized, it cannot keep the floor on its slot more than $MAX\_FULLFRAMES$ frames. After this number, the slot is released and new reservations are needed. Thus, the first drawback of decreasing this parameter is to increase the number of REQ collisions and decrease the average packet train length. This is illustrated on figure 3.24 for $L = 3$. Note that the fairness strategy has always the effect of keeping the number of collisions more or less constant at high input loads.

CROMA is more efficient for high average message length. Hence, increasing $MAX\_FULLFRAMES$ has a direct consequence on the throughput, delay and jitter performance (figures 3.25, 3.26, and 3.27). These phenomena are less pronounced but still observable for $L = 4$ and $L = 6$. Simulation results show that the choice of $MAX\_FULLFRAMES$ has no influence on $L = 8$.
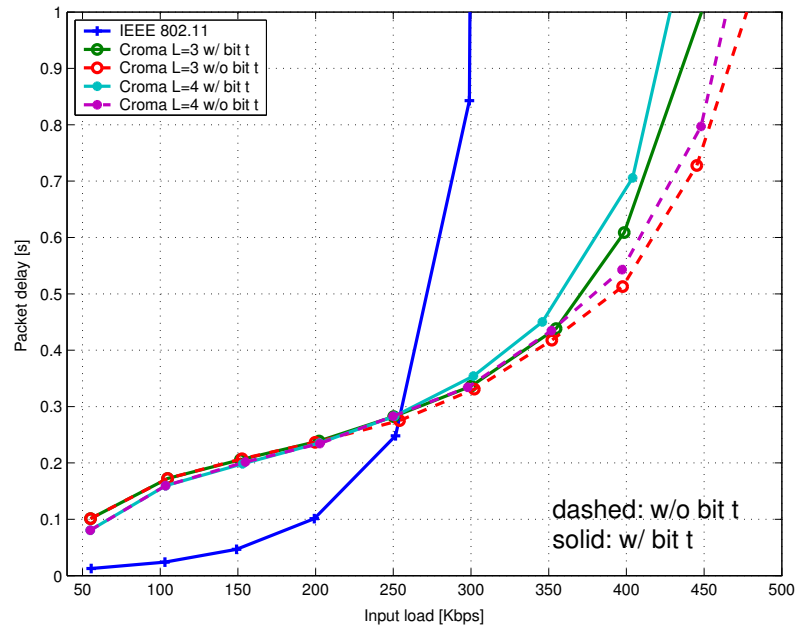
Figure 3.21: Packet delay vs. input load, $L = 3$ and $L = 4$, influence of the bit t, squares topology.

What is the drawback of increasing $MAX\_FULLFRAMES$? This parameter in fact controls the short term fairness of the protocol. On long simulation runs of 200 s or more, the degradation of the fairness index is not particularly visible (figure 3.28). However, if fairness is required on a short period of time, $MAX\_FULLFRAMES$ should be reduced. As an example, we consider two extreme values for $MAX\_FULLFRAMES$: 1 and 500. In figure 3.29, the fairness index is observed all along a simulation run, i.e., it is computed and reset every 5 s. While the index remains always above 0.82 with the parameter set to 1, it drops sometimes to 0.56 with 500.

Thus, $MAX\_FULLFRAMES$ has no influence on the long term fairness but allows a fine control of the short term fairness. It should be designed according to the type of transported traffic, e.g., voice or data traffic. As we don't focus on a particular application in this study, we arbitrary set $MAX\_FULLFRAMES$ to 30 to ensure fairness on a simulation run of 200 s. Moreover, it can be noted that the performance bounds for CROMA throughput and delay are attained in case where the bit t is not used (figures 3.20 and 3.21).

Let us summarize the conclusions of the fairness analysis:

> Long frame versions of CROMA ($L = 6$ and 8) exhibit a very good behavior
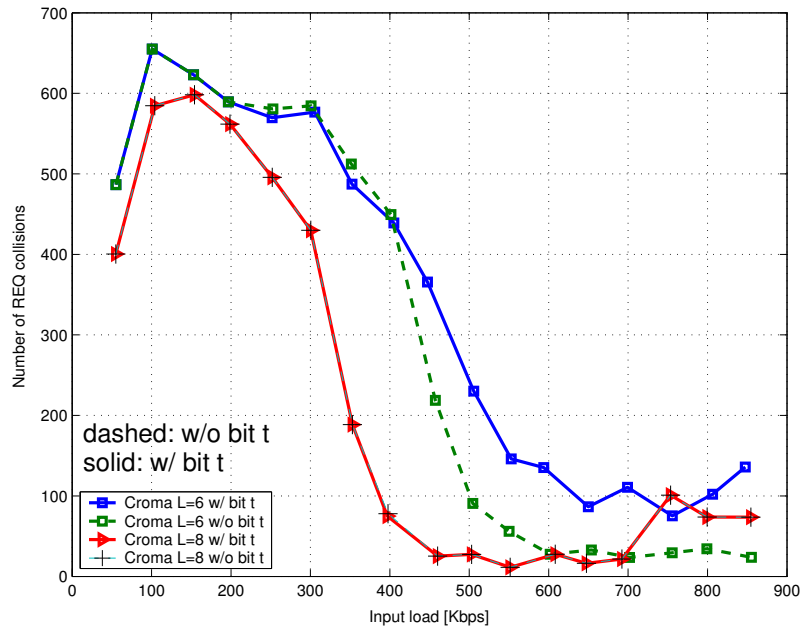
Figure 3.22: Number of collisions of REQ vs. input load, $L = 6$ and $L = 8$, influence of the bit t, squares topology.

in terms of fairness and outperform IEEE 802.11.

> Short frame versions of CROMA ($L = 3$ and 4) also outperform IEEE 802.11 thanks to the use of bit t.

> The use of bit t has the effect of stabilizing the protocol at high input loads at the expense of the saturation throughput and packet delays.

> Increasing $MAX\_FULLFRAMES$ improves the performance of CROMA but decreases the short term fairness of the protocol.

The main characteristics of CROMA have now been shown through the study in a challenging environment. The following sections show that these trends are confirmed in a random network.
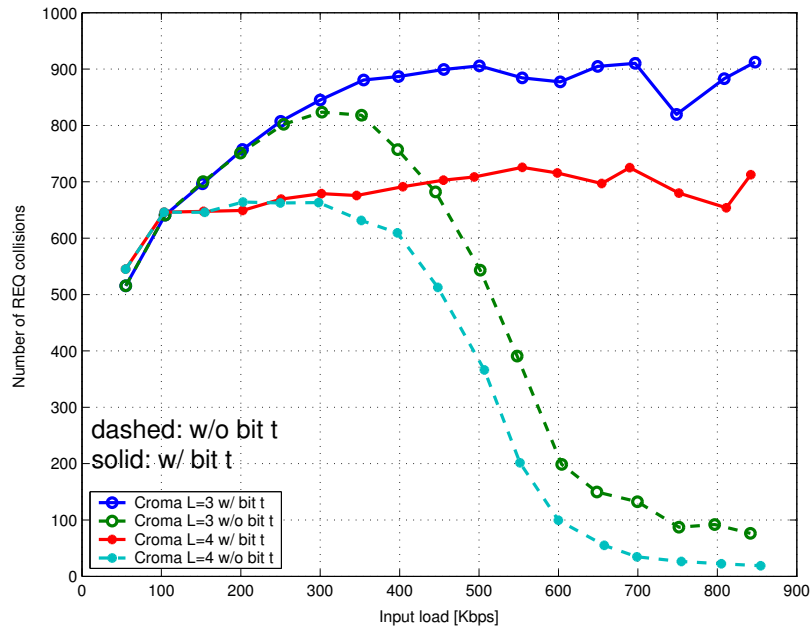
Figure 3.23: Number of collisions of REQ vs. input load, $L = 3$ and $L = 4$, influence of the bit t, squares topology.
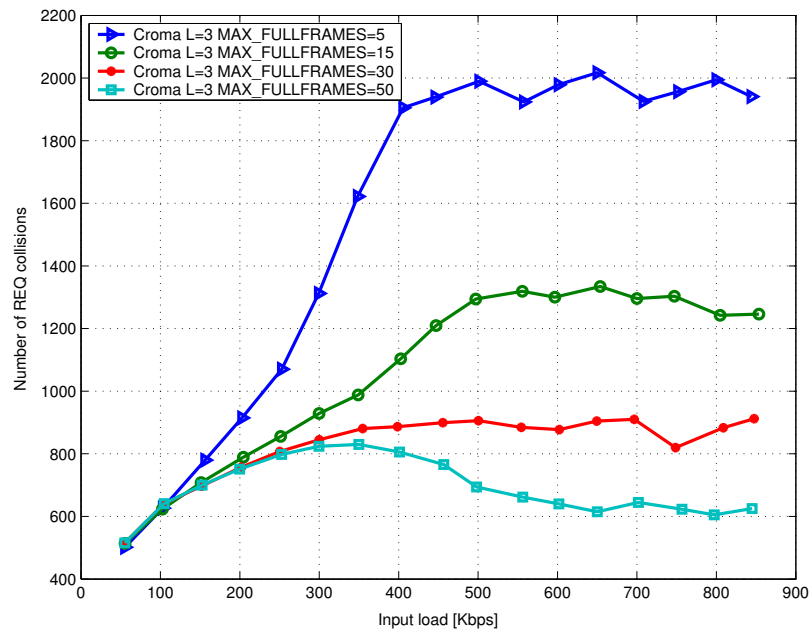
Figure 3.24: Number of collisions of REQ vs. input load, $L = 3$, influence of $MAX\_FULLFRAMES$, squares topology.

150
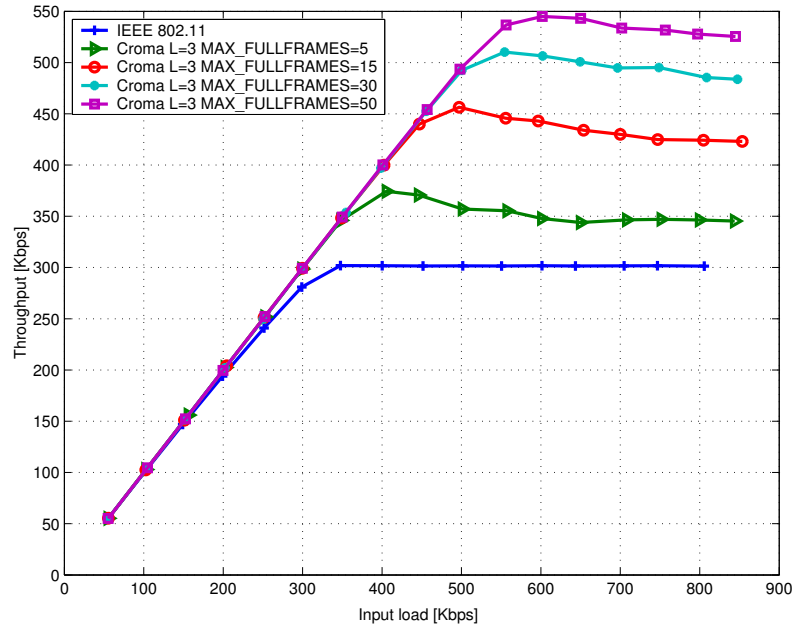
Figure 3.25: Throughput vs. input load, $L = 3$, influence of MAX_FULL-$FRAMES$, squares topology.
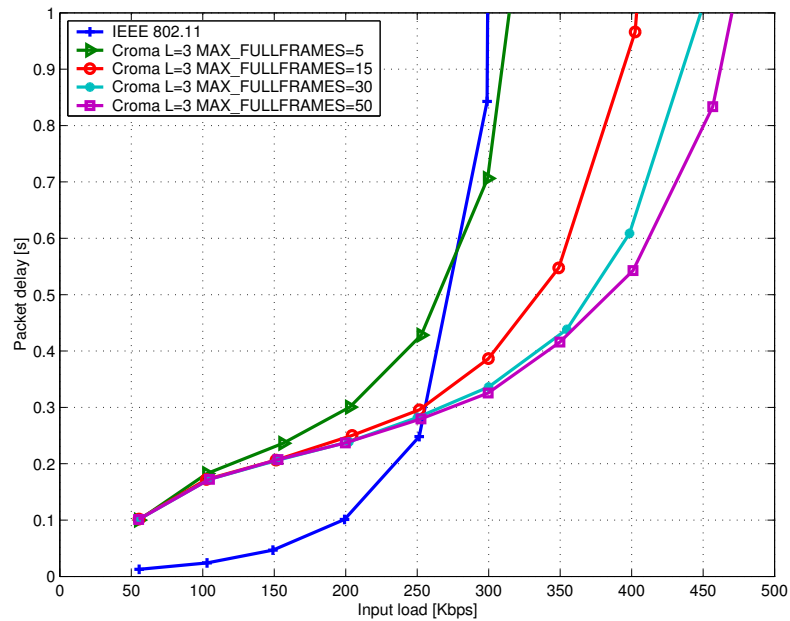
Figure 3.26: End-to-end delay vs. input load, $L = 3$, influence of MAX_FULL-$FRAMES$, squares topology.
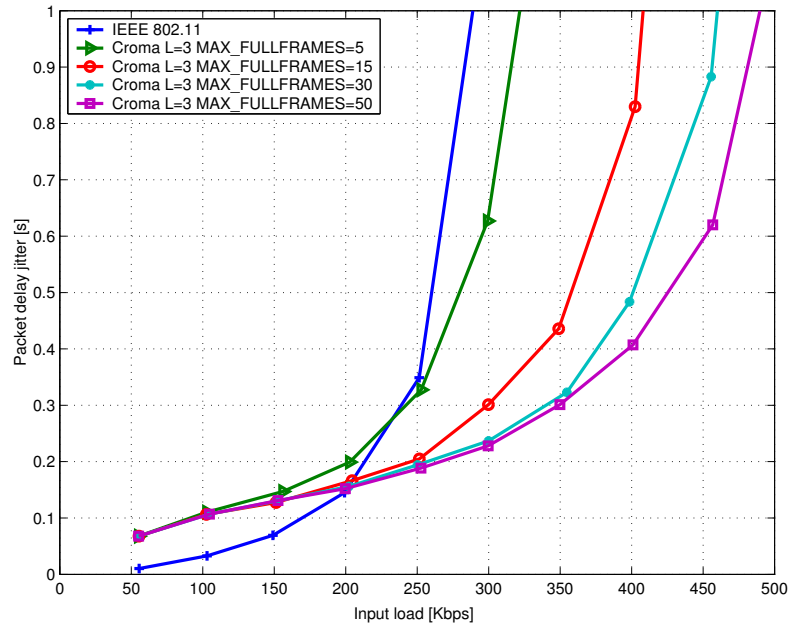
Figure 3.27: End-to-end delay jitter vs. input load, $L = 3$, influence of $MAX\_FULLFRAMES$, squares topology.
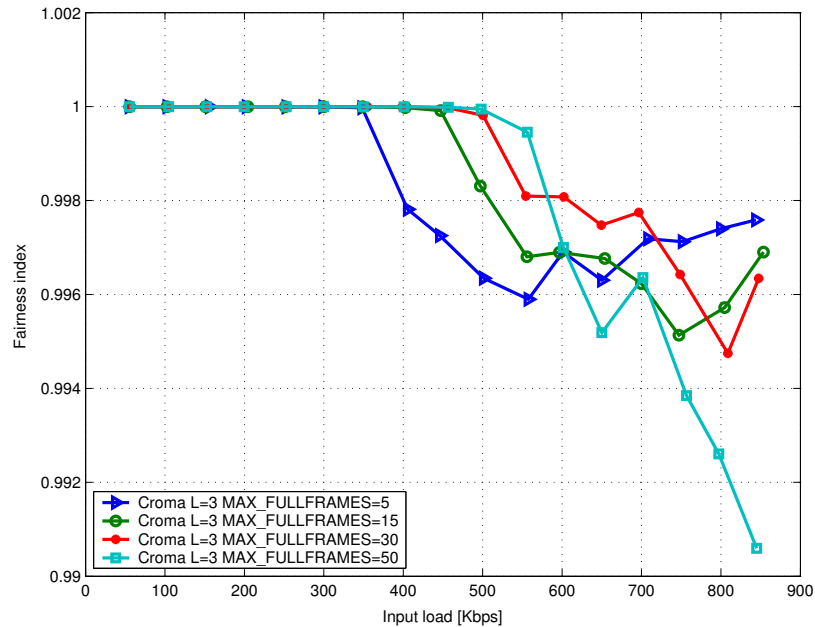
Figure 3.28: Fairness index vs. input load, $L = 3$, influence of $MAX\_FULL\text{-}FRAMES$, squares topology.
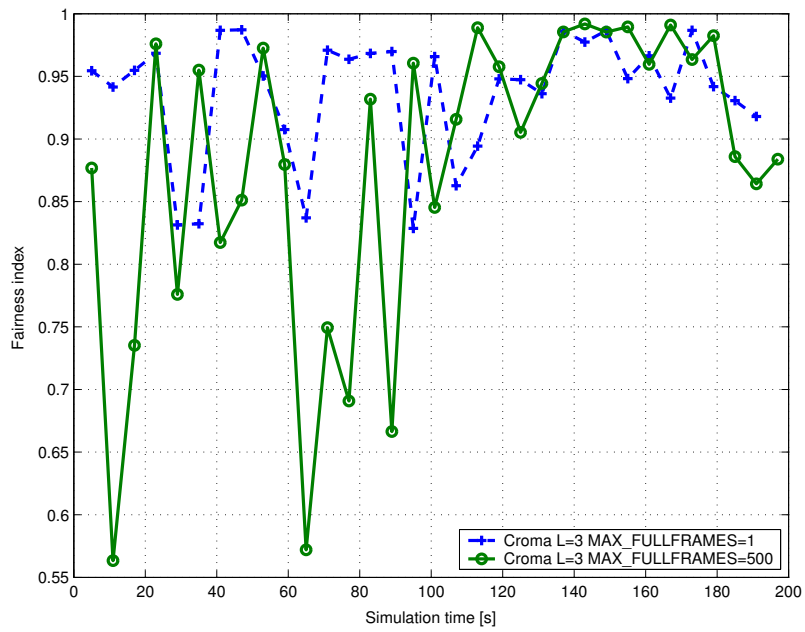
Figure 3.29: Fairness index vs. simulation time, $L = 3$, influence of $MAX\_FULL\text{-}FRAMES$, squares topology.

3.4.3 A Random Network

In the previous section, we compared IEEE 802.11 and CROMA over a simple and pre-defined multi-hop topology. In this section, we consider a random connected network. 30 nodes are drawn at random in a 1000 mx1000 m square area, each node having a transmission range of 250 m. This network is shown in figure 3.30. 10 communications are established between 10 pairs of nodes, 0-1, 2-3, ..., 18-19. Routes are computed by the Dijkstra algorithm. The mean number of hops is 3 and the degree of the network, i.e., the mean number of neighbors of a given node is approximately 5.5, instead of 4 for the previous topology. The traffic is assumed to be ON/OFF with exponential distributions for both ON and OFF periods and with the same parameters as in the previous section. Moreover, we now always consider that the fairness strategy is implemented.
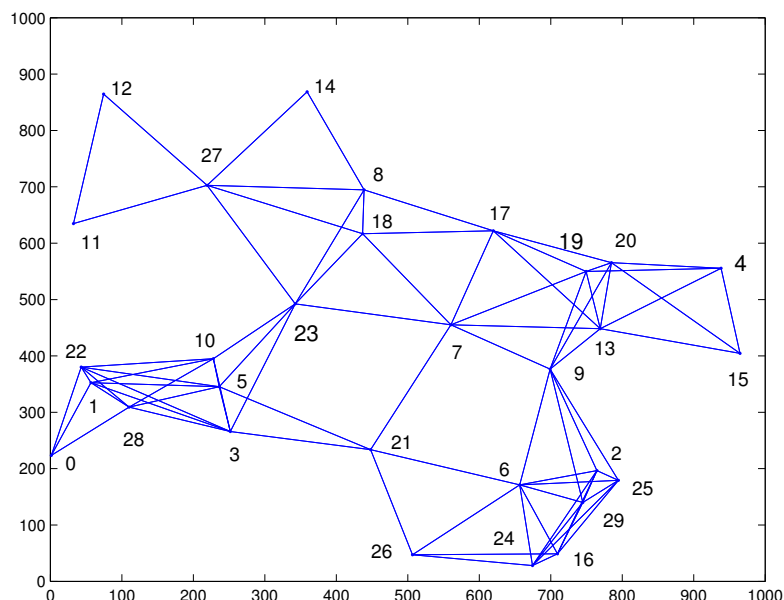
Figure 3.30: Random topology with 30 nodes in a 1000mx1000m area.

Figure 3.31 shows the aggregate throughput of the network as a function of the input load. While IEEE 802.11 saturates at 425 Kbps, CROMA reaches 500, 550, 638, and 685 Kbps respectively for $L = 8$, $L = 6$, $L = 4$, and $L = 3$.

Figure 3.32 shows the mean end-to-end delay of the data packets as a function of the input load. It is clear that, at low input loads, better performance of CROMA in terms of throughput is obtained at the expense of packet delays and jitters (see figure 3.33). In this case, IEEE 802.11 outperforms CROMA. However, CROMA allows to extend the range of acceptable delay and jitter by one third. Note also
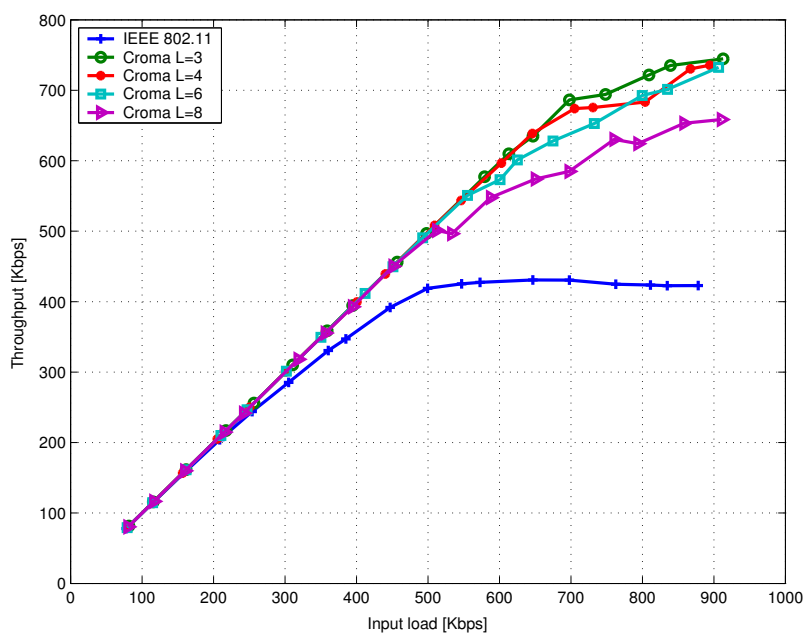
Figure 3.31: Throughput vs. input load, random topology.

that at low input loads, the frame length of CROMA has little influence on the end-to-end delay.

In terms of fairness, CROMA still outperforms IEEE 802.11 in a random topology as shown in figure 3.34, but its fairness index is adversely affected by increasing input load. This reflects the fact that the proposed strategy ensures a local fairness at the MAC sub-layer. In particular, the end-to-end conditions of a flow are not taken into account. As an example, the communication between nodes 0 and 1 is one-hop long, while there are four hops between nodes 4 and 5.

Let us enumerate the conclusions of this section:

> The throughput and delay analysis in this environment confirms the trends observed in the challenging environment.

> CROMA is fairer than IEEE 802.11, but simulations in a random network highlight the fact that only local fairness is ensured.

Despite the fact that simulating random topologies can be cumbersome, ten different random connected networks (not shown here) were simulated and similar results were observed. This is also the case for mobile networks. In this case, the multi-slot communications feature of CROMA presented in the next section is
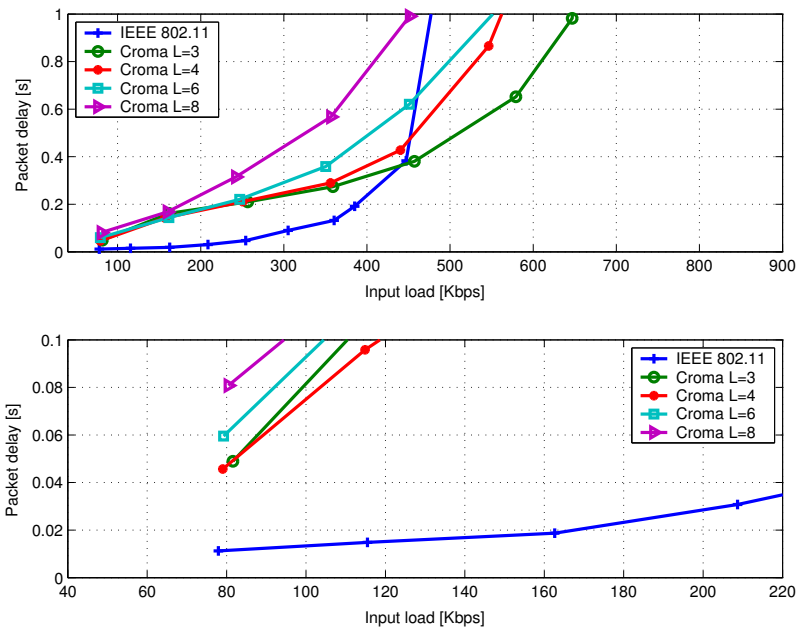
Figure 3.32: End-to-end delay vs. input load, random topology.

particularly adapted. That is the reason why simulation results for a mobile network are not presented here.
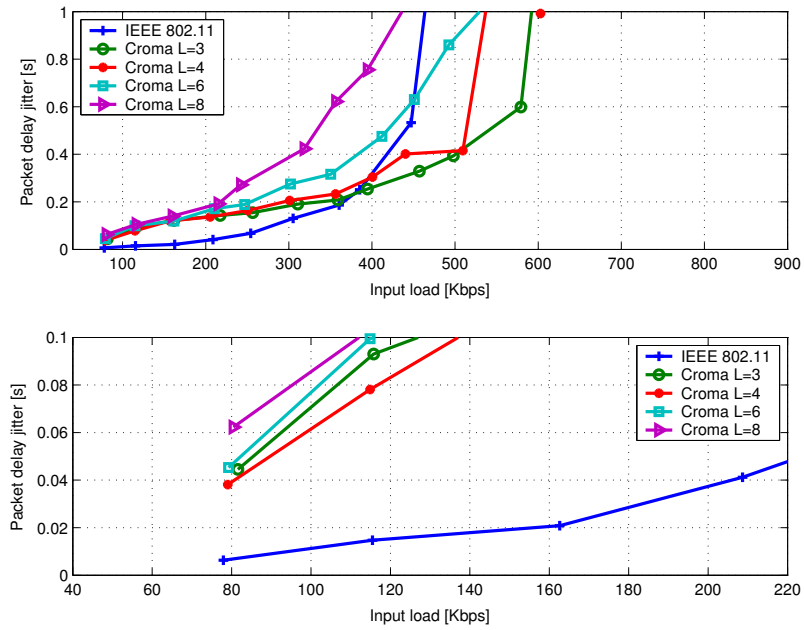
Figure 3.33: End-to-end delay jitter vs. input load, random topology.
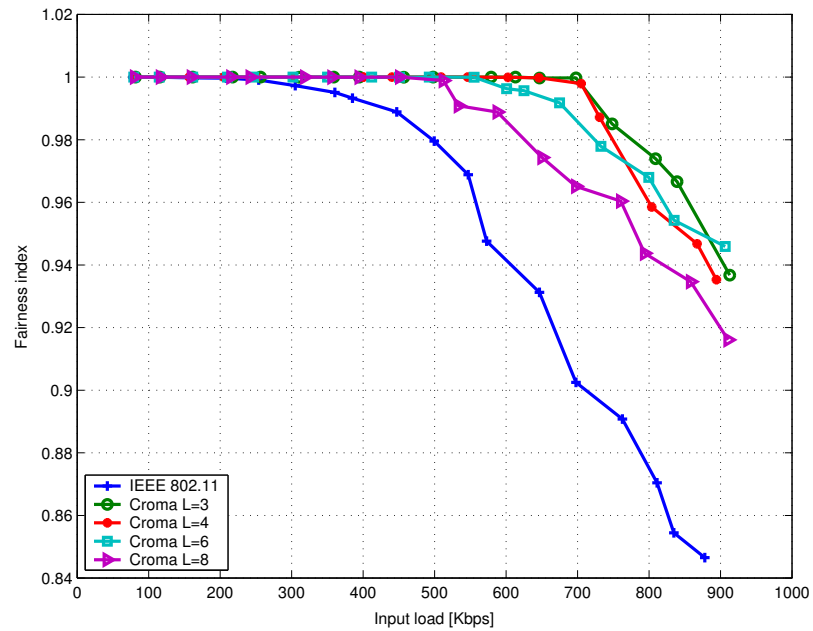
Figure 3.34: Fairness index vs. input load, random topology.

3.5 Multi-slot Communications in CROMA

3.5.1 Feature Description

Up to now a communication between a source-destination pair was associated to a single slot. With this rule, we have seen in the previous section that the performance of CROMA is dependent on the pre-determined frame length, L (see e.g. figure 3.14). As a consequence a multi-slot communication feature has been added to CROMA for optimizing its performance. When it is activated, a communication can be split over several slots. This allows a better utilization of all the slots of the frame.

Each data packet includes a *buffer status* (*bs*) field that indicates whether the sender buffer exceeds a pre-defined value, MS_THRESH (figures 3.35 and 3.3). If it is the case, the receiver is requested to find a free slot in the frame in order to split the communication with its sender. Thus, two or several slots in the frame can be attributed to a single sender-receiver pair.

For a new slot, the receiver has not priority. Indeed, if it has chosen a free slot and receives or senses a packet during the REQ phase of this slot, it refrains from sending a RTR. With this algorithm, new communications that are initiated by REQ packets have priority on already running communications for the first access to the slot.

Figure 3.36 shows an example of multi-slot communication (*bs* is shown next to data packets). On the left hand side, a reservation is done by the sender on slot i, the buffer status field is set to 0. On the right hand side, the buffer exceeds the threshold, *buffer status* is set to 1. The receiver chooses slot j to split the connection on two slots. The RTR is correctly received by the sender. Slot j is attributed to the receiver until the end of the communication on this slot.
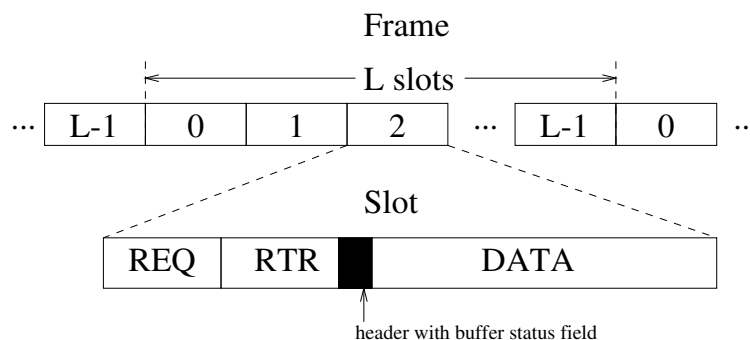
Figure 3.35: Frame structure of CROMA with buffer status field.
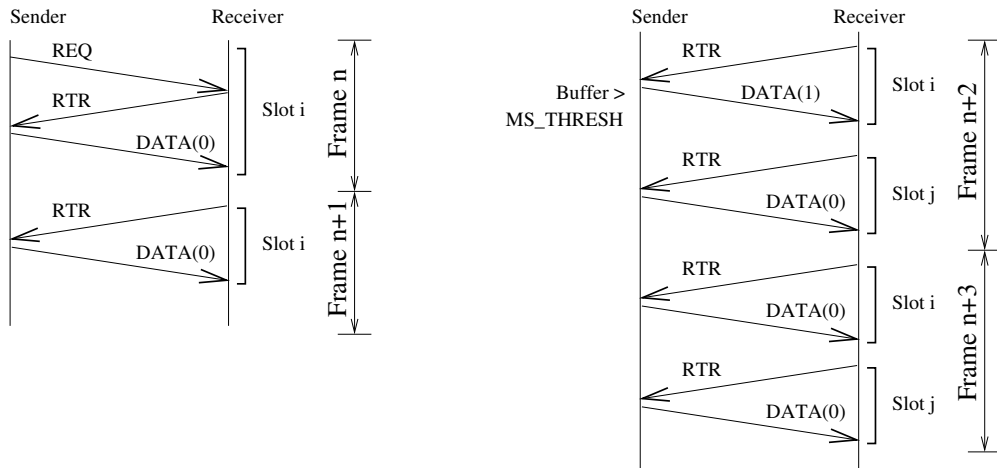
Figure 3.36: Example of multi-slot communication.

3.5.2 A Challenging Topology

Throughput and Delay Analysis

In this section, we study the multi-slot communications feature of CROMA over the challenging topology (figure 3.13) and with the simulation parameters given in table 3.2.
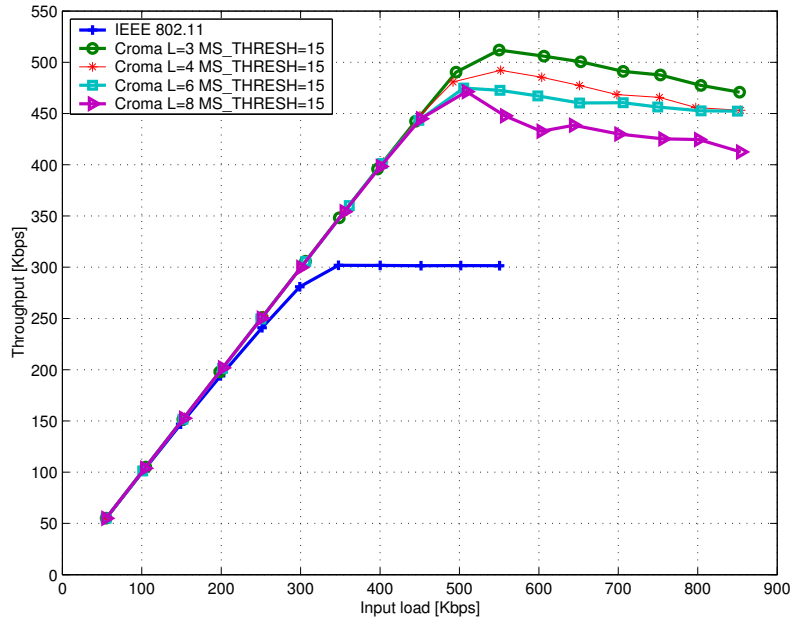
Figure 3.37: Throughput vs. input load, $MS\_THRESH = 15$, squares topology.

The performance of CROMA in terms of throughput, packet delay, and jit-
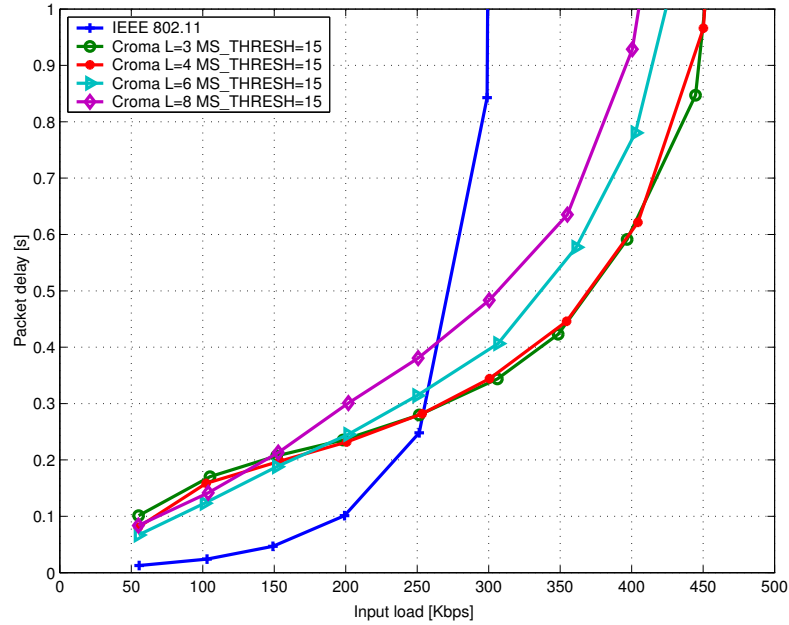
Figure 3.38: Packet delay vs. input load, $MS\_THRESH = 15$, squares topology.

ter are presented respectively on figures 3.37, 3.38, and 3.39 with the parameter $MS\_THRESH = 15$. Comparing to the basic case (figures 3.14, 3.15, and 3.17), it is clear that the performance gap between CROMA $L = 3$ and CROMA $L = 8$ has been reduced. And no change in performance for CROMA $L = 3$ can be noticed. Figure 3.40 presents channel utilization and MAC throughput on the same set of curves.

> Multi-slot communication improves the throughput for long and very long frame length, whereas no difference in performance for small L is observed.

This is further illustrated in figure 3.41 for the delay, where the performance of CROMA with ($MS\_THRESH = 15$) and without the feature ($MS\_THRESH = \infty$) are compared. Similar behavior is seen for delay and jitter. The study of the impact of MS_THRESH on CROMA performance shows that:

> The saturation level throughput is independent on this threshold value (figure 3.42). This is explained by the fact that at such high input loads the system is not stable and the buffers grow indefinitely.

> The influence is slightly greater for packet delay and jitter (figures 3.43 and 3.44). Small values of MS_THRESH, e.g. 1, imply worse delays and jitter
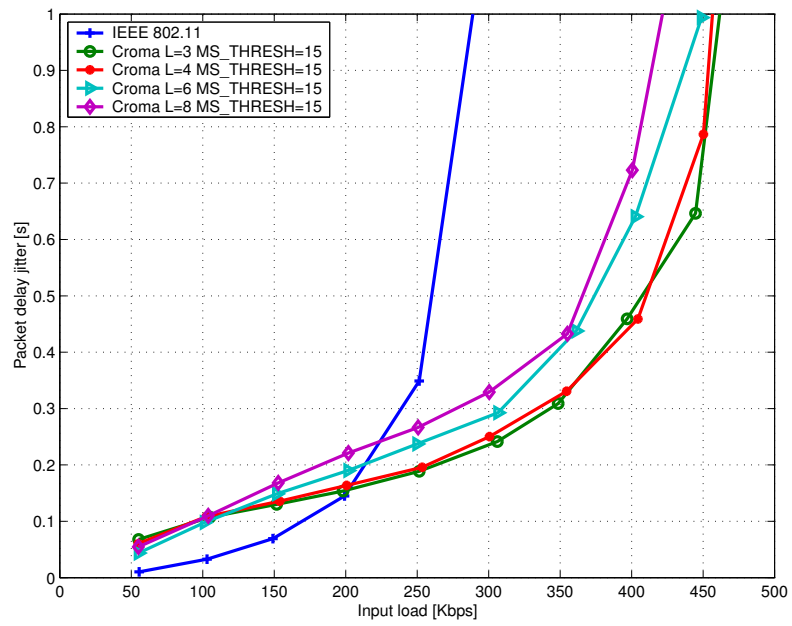
Figure 3.39: Packet delay jitter vs. input load, $MS\_THRESH = 15$, squares topology.

because of an increased contention for slots. This leads to unused resources. At the opposite, high values of MS_THRESH, e.g. 15, means that only a few communications can take advantage of free slots.
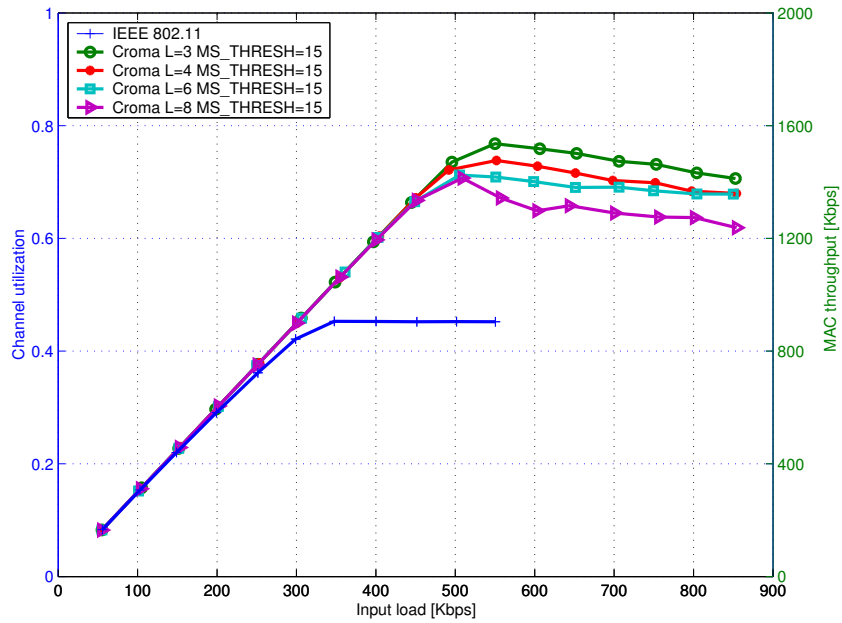
Figure 3.40: Channel utilization and MAC throughput vs. input load, $MS\_THRESH = 15$, squares topology.
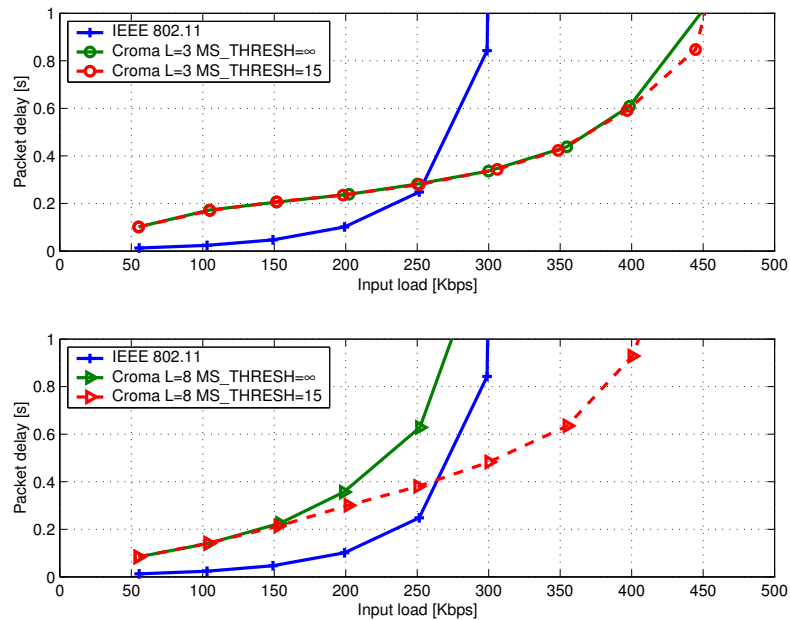
Figure 3.41: Packet delay vs. input load, $MS\_THRESH = 15$, $L = 3$ and $L = 8$, squares topology.

162
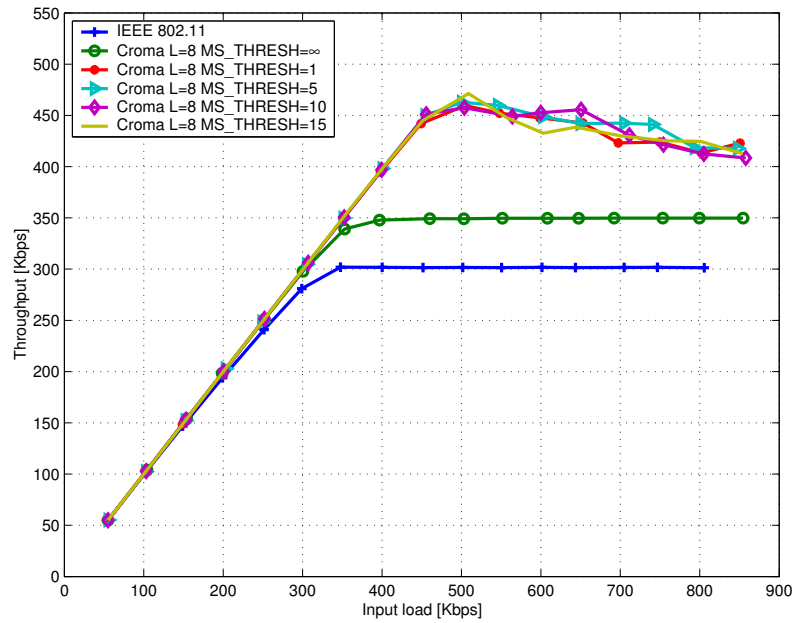
Figure 3.42: Throughput vs. input load, $L = 8$ with multi-slot communication feature, squares topology.
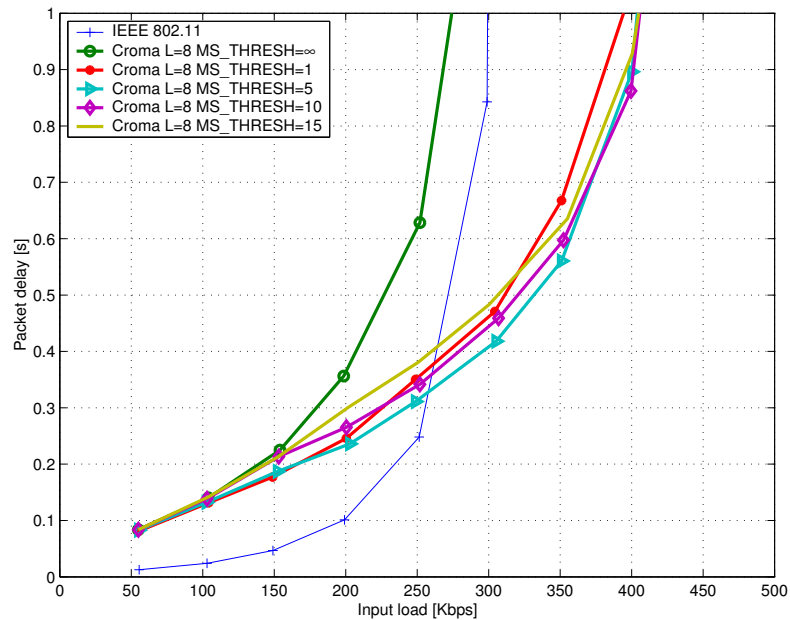
Figure 3.43: Packet delay vs. input load, $L = 8$ with multi-slot communication feature, squares topology.
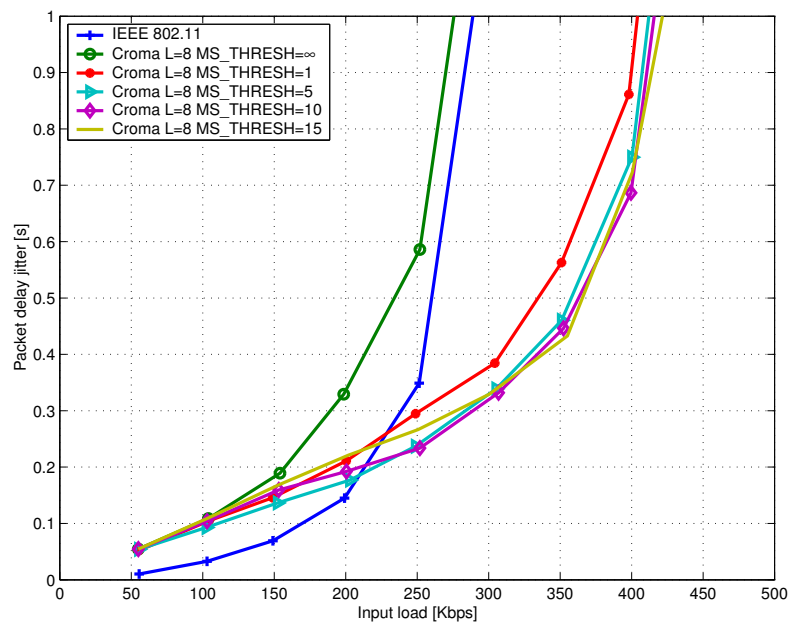
Figure 3.44: Packet delay jitter vs. input load, $L = 8$ with multi-slot communication feature, squares topology.

Fairness Analysis

From the preceding section, sensitivity of CROMA throughput and delay performance to frame length can be controlled by use of multi-slot communication. It however also implies a small degradation of the fairness index, as shown in figure 3.45 for $L = 8$. Since a communication is split on several slots on an opportunistic basis, links with less contention take advantage of their situation. On the other hand, this feature has no impact on fairness for smaller frame lengths (see e.g. figure 3.46 for $L = 6$)
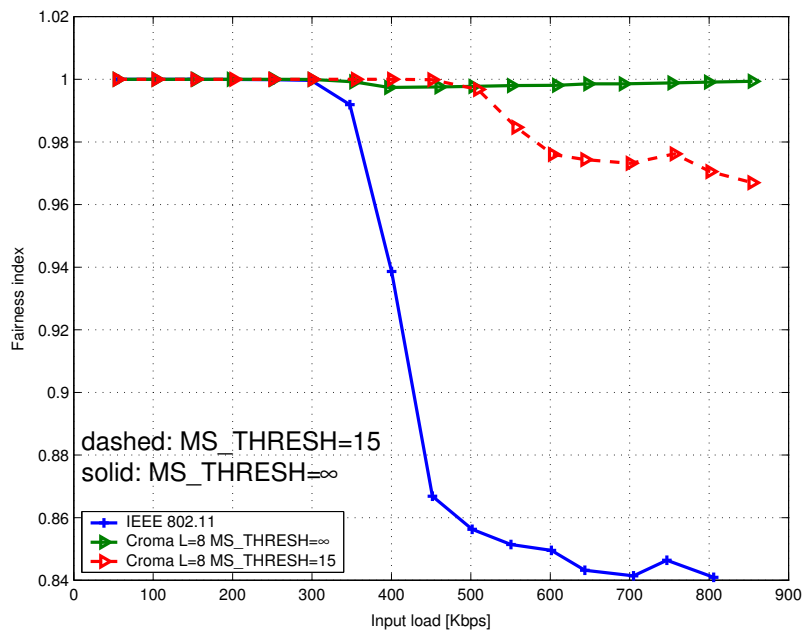
Figure 3.45: Fairness index vs. input load, $L = 8$, $MS\_THRESH = 15$, squares topology.
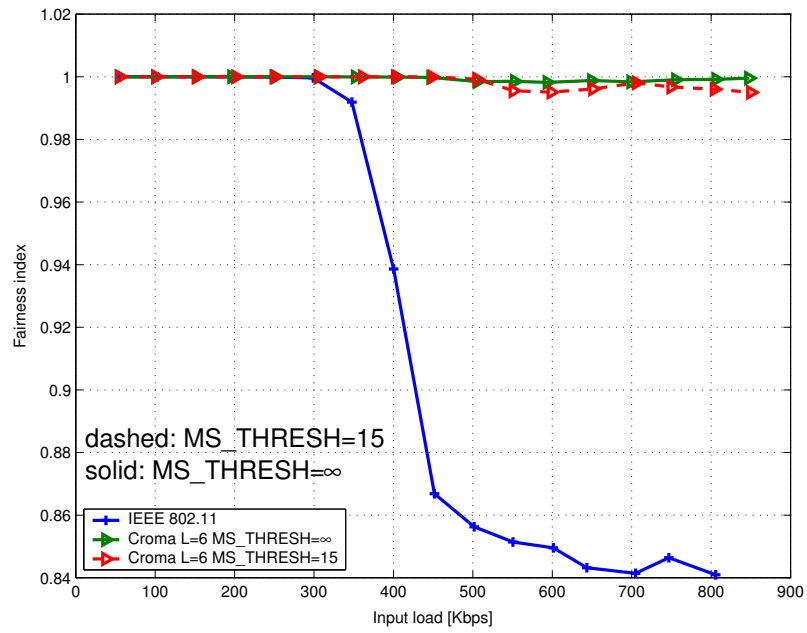
Figure 3.46: Fairness index vs. input load, $L = 6$, $MS\_THRESH = 15$, squares topology.

3.5.3 A Random Network

Similar to the challenging networks topology, the multi-slot communication feature helps to improve the performance of the over-dimensioned versions of CROMA in random networks, e.g. with $L = 8$. The influence of the frame length is decreased. Simulations are performed over the random topology of figure 3.30. The effect of the feature is shown in figures 3.47, 3.48, 3.49, and 3.50.

It can be noted that:

> All throughput curves now converge whatever the frame length of CROMA is.

> Delay and jitter performance are slightly improved at low input loads, the gain is however greater at moderate to high input loads.

> In terms of fairness, the feature doesn't help much: end-to-end conditions of the communications, e.g. the number of hops, are not taken into account, so that only a local fairness is ensured. The index is nevertheless still better than IEEE 802.11.
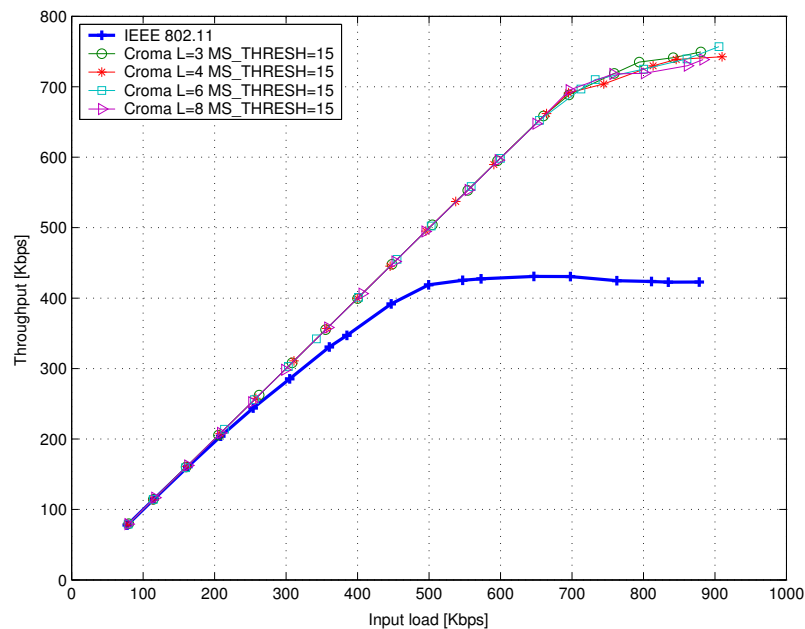
Figure 3.47: Throughput vs. input load, $MS\_THRESH = 15$, random topology.
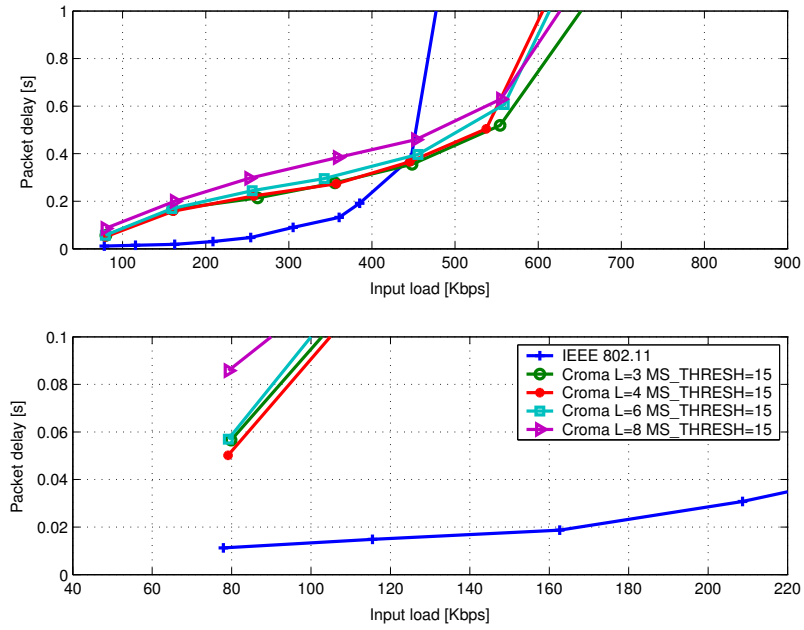
Figure 3.48: End-to-end delay vs. input load, $MS\_THRESH = 15$, random topology.
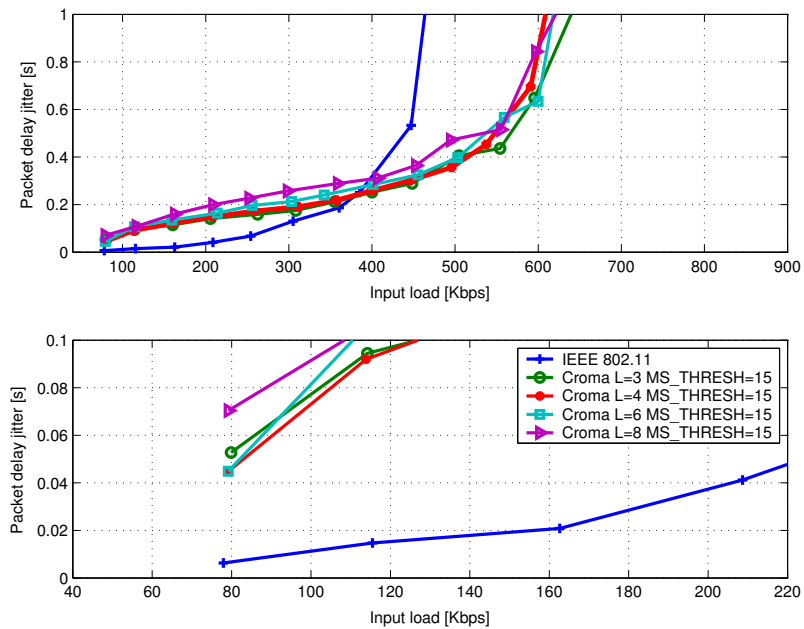
Figure 3.49: End-to-end delay jitter vs. input load, $MS\_THRESH = 15$, random topology.

168
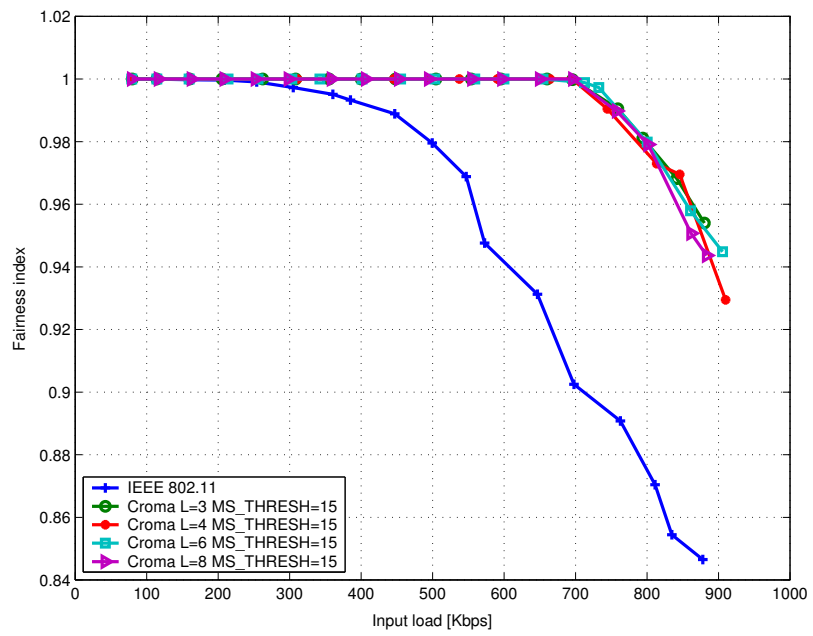
Figure 3.50: Fairness index vs. input load, $MS\_THRESH = 15$, random topology.

3.5.4 A Mobile Network

In this section, the performance of CROMA in a mobile network is evaluated. In order to study the behavior of ad hoc networks, the literature often relies on the *random waypoint* mobility model. This is mainly true for the routing protocol comparison or evaluation, e.g. in [57, 77, 117, 148, 170], more rarely for other layers, like transport [112].

The random waypoint mobility model is a simple stochastic model that describes the movement of a node in a two-dimensional area: A node randomly chooses a destination in the network area and moves with constant speed, picked uniformly in $[0; v_{max}]$, to this point. At the destination, it waits for a while before choosing a new destination, and so on. Thus, the mobility is characterized by two parameters: v_{max}, the maximum speed and *pause_time*, the waiting time at destination.

It is difficult to imagine an application that could match with this model, especially in a bounded area. It doesn't even provide a uniform point distribution [53]. Since it is a widely used mobility model for benchmarking, we adopt it in our study. The detailed parameters of the simulations are given in table 3.3. Note that two quite extreme cases have been chosen for v_{max}. The dependence on *pause_time* is not studied.

The routing layer behaves as follows: On receiving a packet, the routing protocol (Dijkstra algorithm, i.e. an ideal case) of a node computes the next hop towards the final destination of the packet. If the destination is not reachable because the network is momentarily not connected, the packet is discarded. These dropped packets explain why the throughput curve of CROMA $L = 6$ with $v_{max} = 20$ m/s slightly moves away from the first bisecting line in figure 3.51.

As in other environments, the advantage of CROMA over IEEE 802.11 is shown in terms of throughput (figure 3.51), delay (figure 3.52), jitter (figure 3.53), and fairness (figure 3.54).

How can we explain the degradation of performance with mobility, even with a "perfect" routing? Although based on the Dijkstra algorithm, the routing scheme is nevertheless partly responsible. There is indeed a time delay between the computation of the best route and the effective transmission of a packet. This delay is made of a queuing delay and of a channel access delay. In fact for high mobility and high input loads, a lot of packets may not find the next hop. In this case, the MAC layer retransmits several times the packet before discarding it. Hence, the throughput of

the network is reduced and delays increase.

There is also a reason related to CROMA. As nodes move, the communication between a source and a destination may be broken, either because nodes are not in communication range anymore, or because another communication on the same slot implies its end (see section 3.2.7). If the topology of the network changes slowly compared to the duration of a ON period (or message length), this phenomenon has a small effect on the performance. Otherwise, mobility artificially reduces the average message length leading to a saturation in CROMA efficiency.

Two conclusions may be drawn for the mobile network study:

> Performance trends of CROMA are confirmed with a mobile network. In particular, CROMA outperforms IEEE 802.11 in terms of throughput but exhibits higher delay and jitter at low input loads.

> Mobility artificially reduces the average message length, so that CROMA performance decreases with an increasing node speed.

Table 3.3: Main parameter values for simulations in a mobile network.

Parameter	Value
L	6
MS_THRESH	15 packets
v_{max}	1 m/s and 20 m/s
$pause\_time$	0 s
Network size	1500mx300m
Transmission range	250 m
Number of nodes	50
Number of source-destination pairs	10
Routing protocol	Dijkstra
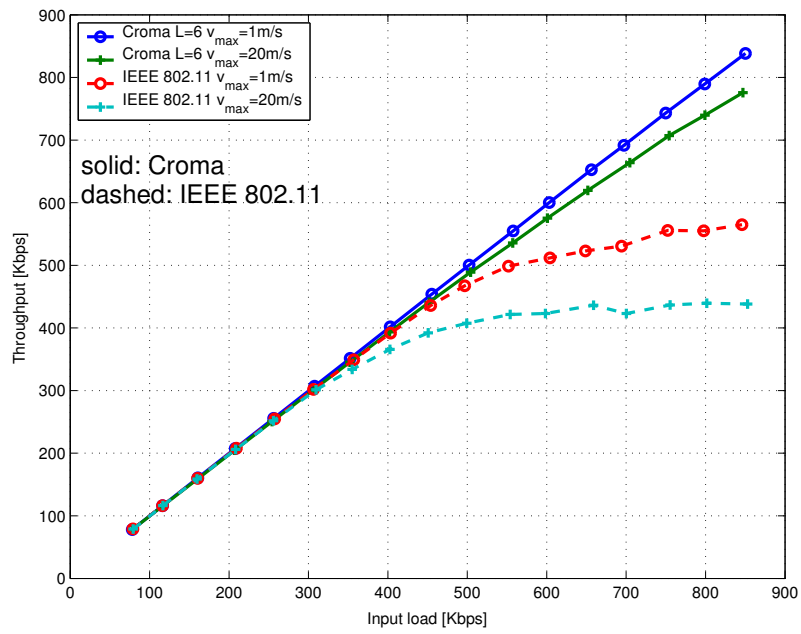
Figure 3.51: Throughput vs. input load, mobile network.
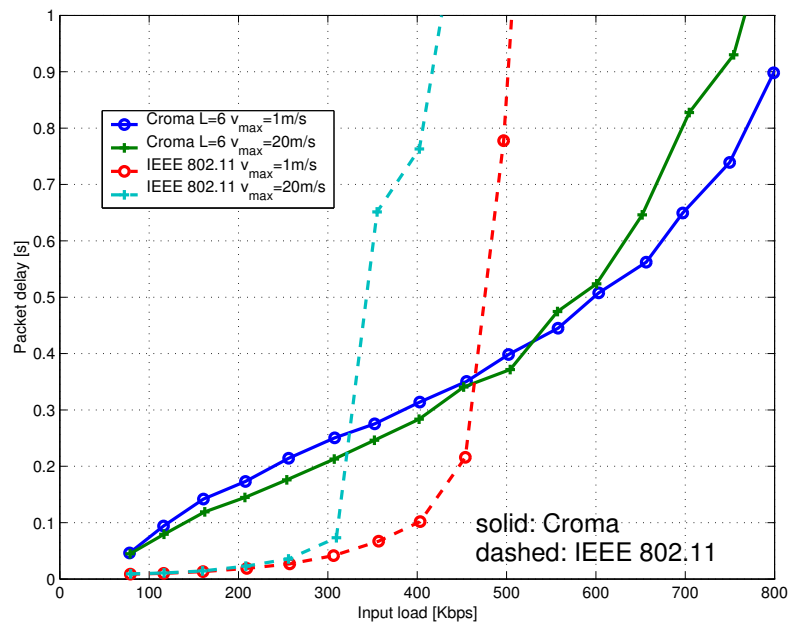
Figure 3.52: End-to-end delay vs. input load, mobile network.
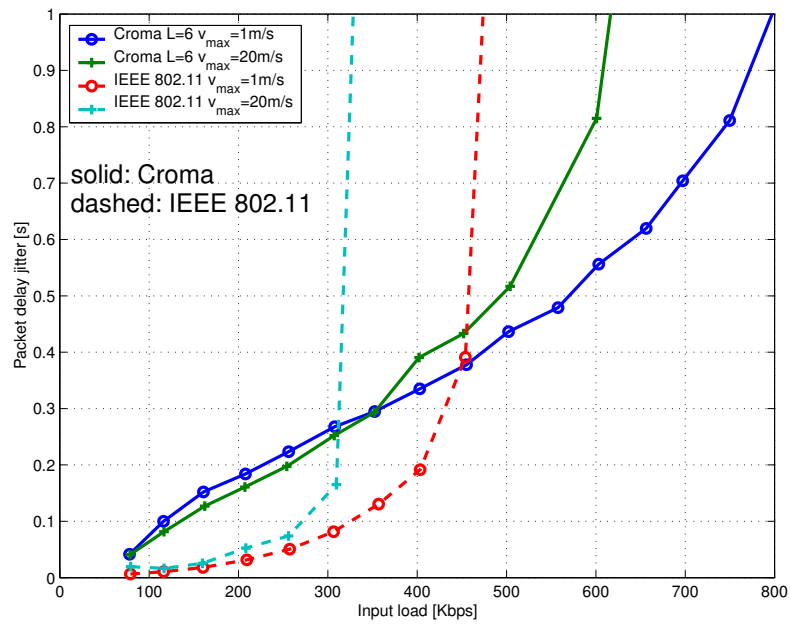
Figure 3.53: End-to-end delay jitter vs. input load, mobile network.
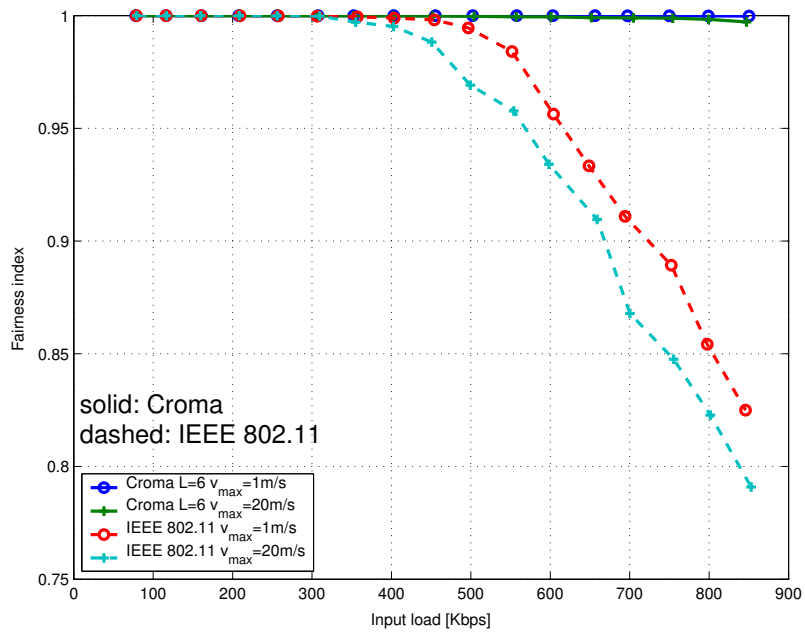
Figure 3.54: Fairness index vs. input load, mobile network.

3.6 Conclusion

In this chapter, a new MAC protocol, called CROMA has been proposed for mobile ad hoc networks. CROMA operates in a synchronized environment and divides time into frames and slots. It is collision-free and receiver-oriented. The reservation of the resources is made through a random access phase on each slot of the frame. The transmission is ensured by a polling mechanism controlled by the receivers. Thus, receivers of a connection act as local and temporary base-stations.

Theoretical analysis and extensive simulations allow us to draw the main conclusions about the performance of CROMA:

> CROMA can reach a very high throughput in both fully connected and multi-hop networks. It outperforms IEEE 802.11 thanks to a better channel utilization. CROMA exhibits however higher delays and jitters than the IEEE standard at low input loads.

> A network operating with CROMA experiences a better local fairness than with IEEE 802.11 in multi-hop networks.

> CROMA is more robust with respect to node mobility than IEEE 802.11. However, it suffers from a minor throughput degradation for mobile ad hoc networks.

Some possible avenues for further investigation on CROMA are:

- a synchronization scheme

- link layer functions, e.g. segmentation and/or aggregation of data packets can be studied and link adaptation algorithms

- a specific algorithm for broadcast packets that could take into account the flooding of broadcast packets

Glossary

π	Stationary probabilities of the DTMC.
a_i	Number of communications on slot i.
$BACKOFFWND$	Back-Off window.

BO	Back-Off timer.
$BOmin$	Minimum value for the back-off timer.
$BOmax$	Maximum value for the back-off timer.
buffer status	DATA header field that indicates wether or not a new slot has to be found for this communication.
dest.ad	MAC address of the destination of a packet.
EOT	End of Transmission, last sequence number of a communication.
fc	Frame control: generic information in the MAC header.
fcs	Frame check sequence.
FREE	Slot state "free".
K	Maximum number of communications per time-slot.
k (bits)	Indication in a RTR of the number of current communications on a slot.
L	Number of time-slots per frame.
M	Maximum number of retransmissions of DATA packets.
$MAX\_FULLFRAMES$	Fairness parameter. If a receiver detects $MAX\_FULLFRAMES$ successive full frames, it sets the bit t of its RTR to 1.
MS_THRESH	Multi-slot communication parameter. If the *buffer status* field of a DATA header indicates that the sender buffer exceeds MS_THRESH, the receiver is requested for finding a new slot for this communication.
N	Number of nodes.
n (byte)	Slot utilization information in RTR, consists of seven bits for k and one bit for t.
N_c	Average number of communications on a slot.
OCC-A-COL-k	Slot state "occupied, available, a collision occurred on the REQ-mini-slot, there are k communications on the slot".
OCC-A-NCOL-k	Slot state "occupied, available, a collision didn't occur on the REQ-mini-slot, there are k communications on the slot".

OCC-NA	Slot state "occupied, not available".
P	Transition matrix of the DTMC.
p	Parameter of the geometrical law describing OFF periods.
$pause\_time$	Waiting time at destination of the random waypoint mobility model.
$polled.ad$	MAC address of the polled sender.
q	Parameter of the geometrical law describing ON periods.
qs	Requested quality of service.
r	Response to a REQ included in a RTR (ACK, NACK, NOTRECVD, and COL).
REQ	Request control packet.
$req.ad$	MAC address of a requesting node to which a RTR is destined.
RTR	Ready-to-Receive control packet.
S	Number of occupied slots in a frame.
sn	Sequence number of DATA packets, used for acknowledgment in RTR.
$source.ad$	MAC address of the source of a packet.
t (bit)	Fairness bit included in RTR.
v_{max}	Maximum speed of the random waypoint mobility model.
W	Maximum number of pollings of a sender without response, after W pollings, the communication is released by the receiver.
W'	Maximum number of frames without polling from the receiver, after W' frames, the communication is considered to be broken by the sender.

Chapter 4

Additional Mechanisms for Capacity Enhancements

4.1 Introduction

Recent technologies in wireless networks illustrate the fundamental trade-off between coverage and offered data rates: Smaller is the communication range, higher is the proposed physical layer throughput. Could ad hoc networking be a disruptive concept against this trend? At a first glance, relaying seems to be a means to extend the coverage of high bit rate technologies. But this has to be realized by using a shared radio channel between hops, an as shown in chapter 2, this results in a throughput decrease.

This fundamental trade-off is probably one of the reasons why the evaluation of the capacity of ad hoc networks as well as the study of their scalability are very active research areas. In this field, Gupta and Kumar [104] provided a major result that is presented in the first section of this chapter. It is also a very pessimistic result because they stated that ad hoc networks are not scalable.

In this chapter, we investigate cross-layer interactions to enhance the performance of distributed MAC protocols. The presented solutions are also very good candidates to improve the performance of CROMA that relies on slotted ALOHA. To this end, we first scrutinize the assumptions made by Gupta and Kumar (section 4.2). Then, we try to relax these assumptions in order to overcome the limits given in [104] for fixed networks. We illustrate this approach through three distributed algorithms.

Theoretical results, presented in [102], leads to the counter-intuitive observation that mobility increases the capacity of ad hoc networks. It is observed that the relaying traffic over multiple hops is a source of capacity degradation. Authors make however the observation that if nodes are mobile in a bounded area, there is a non-zero probability that two nodes are neighbors at a given time instant. They thus reduce the allowed maximum number of hops by exploiting this characteristics of mobile networks. In section 4.3, we propose a distributed scheduling policy based on a simplified version of CROMA that takes advantage of node mobility and illustrates the results of [102]. The importance of the traffic model is highlighted. Theoretical results on the optimal transmission range are also provided.

In section 4.4, the assumption of a fixed channel capacity is challenged thanks to the concept of multi-user diversity. This notion is applied to random access with the evaluation of the capacity of the channel aware slotted ALOHA: The transmission of a packet is constrained to the channel quality experienced by the sender. Results show the increase of the protocol capacity with the number of users.

Finally, section 4.5 imposes a modification of the interference model by considering multi-packet receptions. In this context, the throughput of the traditional slotted ALOHA is studied and the advantage of using multi-user detection in ad hoc networks is shown. In these two latter cases, we found good ways of investigation to improve the reservation phase of CROMA based on slotted ALOHA.

Details and ideas of this chapter can be found in the following papers of the author: [7, 9, 12, 13].

4.2 The Capacity of Wireless Networks

In this section we give a short summary of the paper of Gupta and Kumar [104], which explains the fundamental limits on the capacity of wireless networks. The rather pessimistic implications of these results as well as the possibility of relaxing certain assumptions are discussed here.

Gupta and Kumar consider n nodes located in a 1 m^2 area, either a disk of radius $1/\sqrt{\pi}$ in the plane or a three-dimensional sphere. The latter assumption allows the authors to be independent on the edge effects. Each node can transmit at W bits/s on a shared channel. Then, two types of networks are considered for the evaluation of capacity: arbitrary networks and random networks.

4.2.1 Arbitrary Networks

In arbitrary networks, the node locations, the traffic pattern, and the transmission ranges are arbitrarily selected. Thus, the following paragraphs describe the best case for performance evaluation.

The transport capacity is defined as the sum of products of number of bits and the distances over which these bits are carried per second. This definition is somewhat similar to the MAC throughput, i.e., if we set the transmission range to a fixed value and express distances in number of hops. If one bit per second is transported k-hops away from its source, the MAC throughput is k bits/s. The transport capacity as defined by Gupta and Kumar is also k bit-hops/s.

Two models for the successful reception of a transmission over one hop are given by [104]: the *protocol model* and the *physical model*. We assume that a node X_i transmits to a node X_j over the shared channel. This transmission is successful according to the protocol model if

$$|X_k - X_j| \geq (1 + \Delta)|X_i - X_j| , \qquad (4.1)$$

for every node X_k simultaneously transmitting ($\Delta > 0$). This situation is illustrated in figure 4.1: No transmitting node should be in the disk of center X_j with radius $(1 + \Delta)|X_i - X_j|$.

With the physical model, the transmission is successful if the signal-to-interference plus noise ratio (SINR) is exceeding a given threshold γ_0:

$$\frac{\frac{P_i}{|X_i - X_j|^\alpha}}{\sigma^2 + \displaystyle\sum_{k \in T, k \neq i} \frac{P_k}{|X_k - X_j|^\alpha}} \geq \gamma_0 , \qquad (4.2)$$

where P_i is the output power of node i, σ^2 is the noise power level, $\alpha > 2$ is the path loss exponent, and T is the set of simultaneously transmitting nodes.

The two main results of the paper for arbitrary networks can now be stated:

(i) The transport capacity under the protocol model is $\Theta(W\sqrt{n})$ [1] bit-meters per second if the nodes are optimally placed, the traffic pattern is optimally chosen, and if the range of each transmission is chosen optimally. In particular, if the capacity

---

[1] We recall the notation of Knuth: $f(n) = \Theta(g(n))$ denotes that $f(n) = O(g(n))$ as well as $g(n) = O(f(n))$.
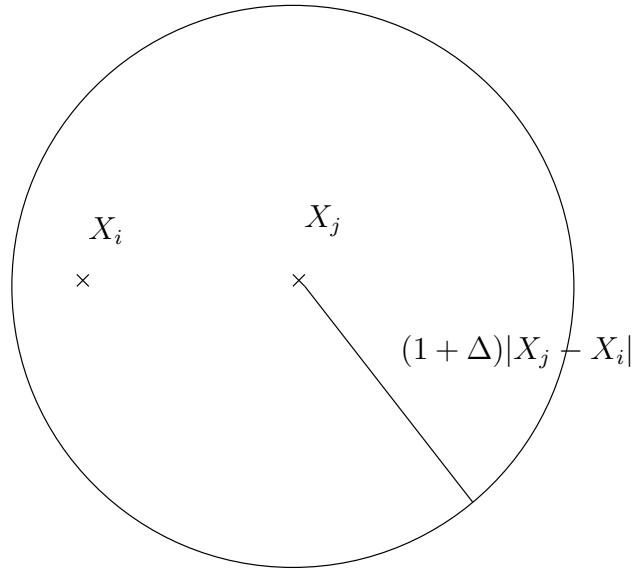
Figure 4.1: Protocol model range of successful reception.

is uniformly divided between the nodes, then each of them would obtain $\Theta(W/\sqrt{n})$ bit-meters per second.

(ii) For the physical model, $cW\sqrt{n}$ bit-meters per second is feasible, while $c'Wn^{\alpha-1/\alpha}$ bit-meters per second is not, for appropriate constants c and c'.

All necessary details are explained in [104].

4.2.2 Random Networks

In random networks, nodes are independently and uniformly located on the 1 m^2 area. Each node has a randomly chosen destination to which it wishes to send $\lambda(n)$ bits/s. All transmissions employ the same power P. For the protocol model, the transmission is successful if

$$|X_i - X_j| \leq r \quad \text{and} \tag{4.3}$$
$$|X_k - X_j| \geq (1+\Delta)r, \tag{4.4}$$

for every simultaneously transmitting node X_k, where r is the nominal transmission range.

The physical model is not modified, except that all powers are the same:

$$\frac{\frac{P}{|X_i - X_j|^\alpha}}{\sigma^2 + \displaystyle\sum_{k \in T, k \neq i} \frac{P}{|X_k - X_j|^\alpha}} \geq \gamma_0 \ . \tag{4.5}$$

The transport capacity is here defined in the usual manner as the number of bits per second that can be transmitted by every node to its destination. The two results for random networks are:

(i) The order of the throughput capacity is $\lambda(n) = \Theta(W/\sqrt{n \log n})$ bits per second for the protocol model.

(ii) For the physical model, a throughput of $\lambda(n) = cW/\sqrt{n \log n}$ bits per second is feasible, while $\lambda(n) = c'W/\sqrt{n}$ is not, for appropriate constants c and c'.

4.2.3   Implications

Which conclusions can be drawn from these results?

- The capacity of ad hoc networks is limited by interference on the one hand, by the amount of relaying traffic on the other hand. Increasing the transmission range reduces the number of hops to reach the destination and hence the relaying traffic. The spatial reuse of resources is however also reduced since the neighbors of a sender have to idle while it is transmitting. In fact, the loss of increasing r is quadratic while the gain is linear, so that the range should be reduced a much as possible. The limit is however given by the network connectivity constraint [103], the critical radius for connectivity being $\sqrt{\log n / \pi n}$.

- The presented results are very pessimistic because they imply that ad hoc networks are not scalable. Even if the traffic is limited to the close neighborhood of the sources and for a fair sharing of the resources, each source-destination pair sees its throughput decreasing as $1/\sqrt{n}$ in bit-meters per second.

In this chapter, we propose the design of some MAC enhancements by relaxing a part of rather stringent assumptions of [104]. It is demonstrated that throughput higher than the very conservative estimates presented above are achievable in certain specific scenarios.

4.3 Mobility as a Source of Diversity

4.3.1 Related Work

Grossglauser and Tse proved in [102] that the limitations presented in [104] can be overcome through node mobility by exploiting the concept of multi-user diversity. This notion is already known in a cellular environment [133]: At each time-slot the base-station sends data to the mobile station with the best channel conditions. [102] gives an analogy in mobile ad hoc networks: At each time-slot the only packets allowed to be sent are those that are one hop away from their final destination, i.e., with the best *route conditions*. This analogy leads to one-hop transmissions, i.e., packets are sent only when the destination is in the communication range of the source. The idea behind is to completely eliminate the relaying traffic.

As in [104], Grossglauser and Tse consider n nodes in a 1 m^2 disk on the plane, and each node is a source node for a session, and a destination node for another session. The source-destination association doesn't change with time. We will designate this model as the *single-destination traffic model* since each source has packets for a single destination. The main difference is that nodes are moving and successful sender-receiver exchanges are direct without any relaying.

So, according to [102], there is a constant c such that

$$Pr\left\{\lambda(n) = cn^{-\frac{1}{1+\alpha/2}}W \text{ is feasible}\right\} = 0 \tag{4.6}$$

for sufficiently large n. It means that without relaying the achievable throughput per source-destination tends towards zero at a rate following $n^{-\frac{1}{1+\alpha/2}}$. Thus, completely eliminating the relaying traffic does not bring a lot.

In fact it is claimed that mobility brings a substantial increase in system capacity of ad hoc networks, especially if no more than one relay node between each active source and destination pair is considered. For that purpose, [102] describes a centrally controlled scheduling policy based on a two phase transmission method, i.e., from source to a waiting queue in a relay node and then from the relay node to destination (see figure 4.2, taken from [102]). The basic idea is that in a dense network, the probability of finding adequately matched source and destination nodes as well as the same for finding relay nodes as and when required, increases with node mobility. With the two-phased algorithm, a source-destination throughput of $\Theta(1)$

can be achieved, i.e.,

$$\lim_{n\to\infty} Pr\left\{\lambda(n) = cW \text{ is feasible}\right\} = 1 .\tag{4.7}$$

An intermediate result is that at a given time-slot, the expected number of feasible sender-receiver pair is $\Theta(n)$. This proves that the limits given by Gupta and Kumar can be overcome. An extension of this work to one-dimensional mobility can be found in [78].

Since distributed scheduling policies are known to be more suitable for implementation in ad hoc networking applications, we demonstrate the usefulness of such a scheme. So, we don't focus on the asymptotic capacity but rather on the possibility to use mobility as a source of diversity to increase the network throughput. Moreover, the influence of the traffic model is highlighted in the analysis presented below.
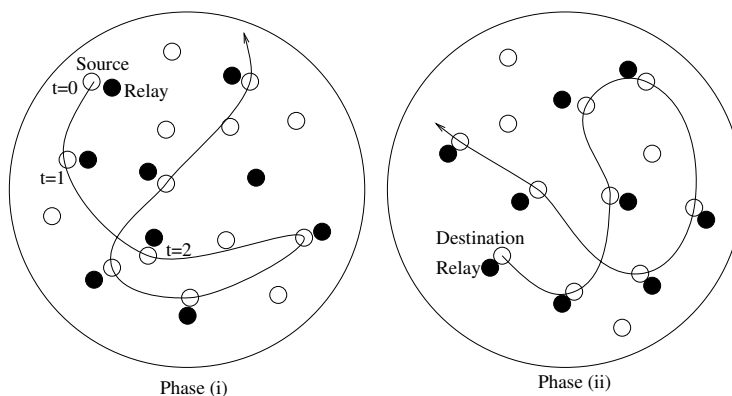
Figure 4.2: Phase (i): The source disseminates packets along its route in relays; Phase (ii): The destination retrieves packets from the relays.

4.3.2 Scheduling Policy and Access Scheme

In the proposed study of a scheduling policy, the network is assumed to be perfectly synchronized and the channel is supposed to be slotted. As for the previous chapter, the issue of synchronization is not addressed.

The MAC protocol is very similar to CROMA: It is a simplified version of the protocol without reservation mini-slot, nor frame structure. A sender that receives an RTR and that has a packet for the receiver can transmit data. Packets have a

fixed length, so that the two-way handshake is possible within a time-slot (see figure 4.3). This protocol is not reliable and there is no collision avoidance mechanism, thus some packets can be lost. We assume that higher layers are responsible for acknowledgment and retransmissions.
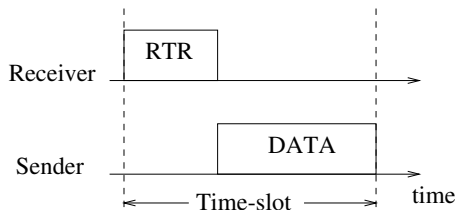
Figure 4.3: Two-way handshake within a time-slot.

We will compare two basic strategies. The first one is based on the analytical study of [102] and considers at most two hops between source and destination. The second one considers only one hop, i.e., a packet is directly sent from a source to a destination without any relay node, when they are in communication range of each other.

We will also compare two traffic models. The *single-destination traffic model* is the same as that assumed in [104] and [102]. In the *multi-destination traffic model*, the destination of each generated packet is randomly chosen among all nodes of the network, e.g. for messaging traffic between nodes.

At each time-slot, θN nodes among N are designated as senders, the remaining nodes are receivers, $\theta \in]0, 1[$. This is done in a distributed way by generating a uniform random variable in each node and comparing the result with the predefined parameter θ, called the sender density. All receivers send a RTR message as described above. The behavior of senders that receive a RTR depends on the predefined strategy.

In the *one-hop strategy*, senders transmit packets with destination address included in the received RTR. As a consequence, packets are transmitted only when the destination is in the transmission range of the source. Thus, only one hop is allowed.

In the *two-hop strategy*, each node manages two packet queues between the MAC layer and the packet generator. One of these, called the source queue, stores packets coming from its own packet generator. The other one, called the relay queue, stores the incoming packets that have to be relayed. A sender receiving a RTR looks in its queues for any packet destined for this receiver. Any such existing packet is

184

transmitted considering the fact that the source queue has priority over the relay queue. Otherwise, a packet is chosen in the source queue to be transmitted to the receiver/relay. This strategy is detailed in pseudo-Specification and Description Language in figure 4.4.
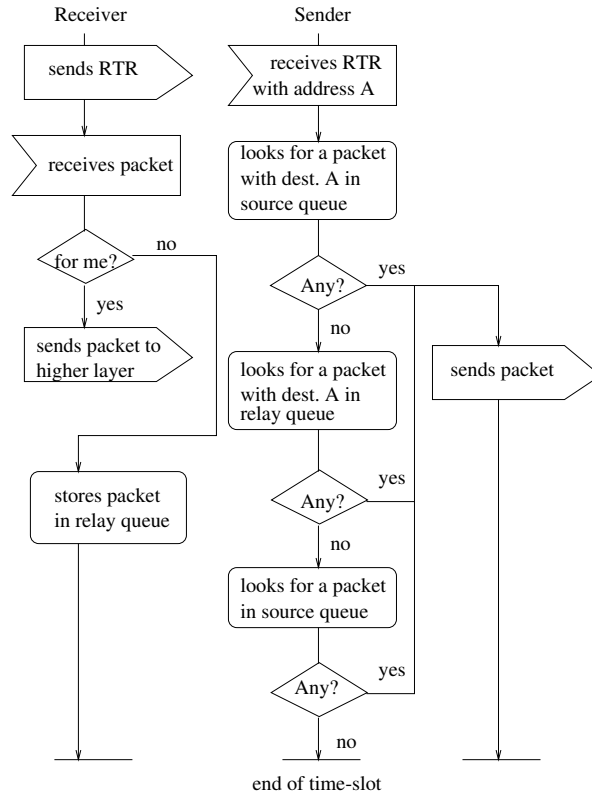
Figure 4.4: Two-hop strategy.

4.3.3 Numerical Results

Simulations have been performed using the event-driven network simulator ns2 [155], as in the previous chapter. N nodes (30 and then 50) with omni-directional antennas have been considered moving in a $1000u \times 1000u$ square field, where u is an adequate unit of distance. The sender density is set to $\theta = 1/2$.

The mobility model uses a simplified version of the random waypoint model (as in chapter 3), where the speed is fixed for all nodes for the entire simulation.

Each node generates traffic according to an exponential ON/OFF distribution. Packets are sent at a fixed rate during ON periods, and no packets are sent during OFF periods. Packets and RTR are of constant size (respectively 512 bytes and

44 bytes). The average ON-time and OFF-time are 0.5 s. The sending rate during ON-times is 64 Kbits/s. Simulations are run for 50 simulated seconds. We also assume that queues could reach an infinite length.

Propagation delay and receive-to-transmit transition time are assumed to be negligible. As in the previous chapter, it is assumed that transmissions are possible within a disk of fixed range and SINR is not taken into account. The only source of packet loss is collisions. Moreover, problems related to high mobility with respect to the channel model, e.g. Doppler effect, are not taken into account.

Figure 4.5 shows the benefit of mobility on the network throughput as a function of the transmission range for $N = 30$ nodes, running the one-hop strategy and with the multi-destination traffic. As explained before, in a multi-hop network, long range communications ensure a very good connectivity of the network and reduce the mean number of hops. However, network throughput is fundamentally limited because of the high level of interference induced by high transmitted power. The number of collisions is also high because of the number of nodes contending for the channel. As a consequence, this design choice increases significantly the MAC overhead and limits spatial reuse of the resources. On the other hand, communications between nearest neighbors increases the mean number of hops and thus routing overhead. In this case, most of the packets carried by the network are relayed packets.

The scheduling policy in [102] and its realization presented here operate with an adequately selected transmission range corresponding to a low transmitted power and a low number of hops. Figure 4.5 shows that an optimal range is achieved at about 150 u in the simulation conditions, and that this range is constant as mobility pattern varies.

The benefit of mobility is also shown for the two-hop strategy in figure 4.6. We also note from these figures that the relaying scheme (two-hop strategy) does not bring additional diversity. Instead, the relaying traffic degrades the performances of the system. This result seems to contradict the conclusion of [102] that claims that better performances are achieved with relaying. This is true as n, the node density, grows indefinitely: The throughput is in the order of $\Theta(1)$ with the two-hop strategy, while it decreases like $n^{-\frac{1}{1+\alpha/2}}$ without relaying. It can be also noted that in this section, only collisions are taken into account, whereas [102] allows reception according to the signal-to-interference ratio. [102] also considers that each sender node transmits packets to its nearest neighbor among all nodes. That is not necessarily the case with the proposed scheduling policy.
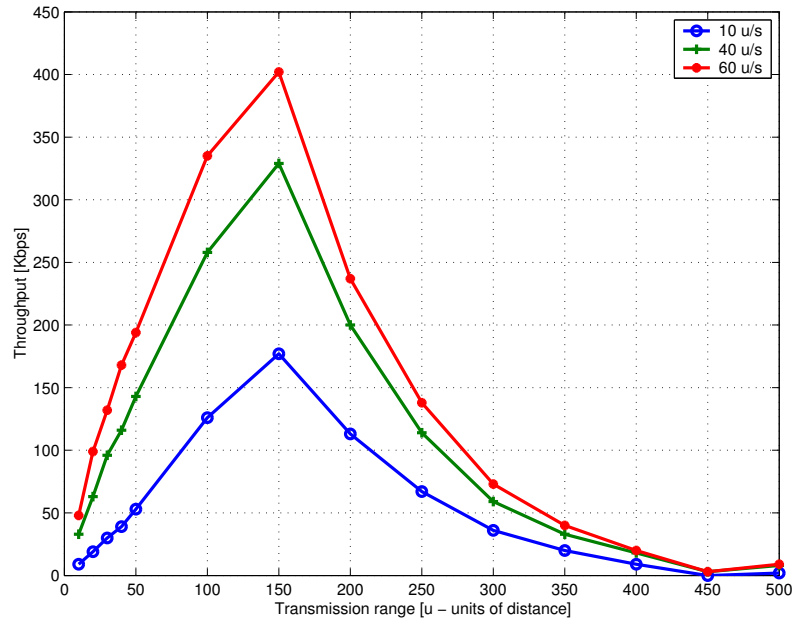
Figure 4.5: Aggregate throughput vs. transmission range - one-hop strategy, $N = 30$, multi-destination traffic model.
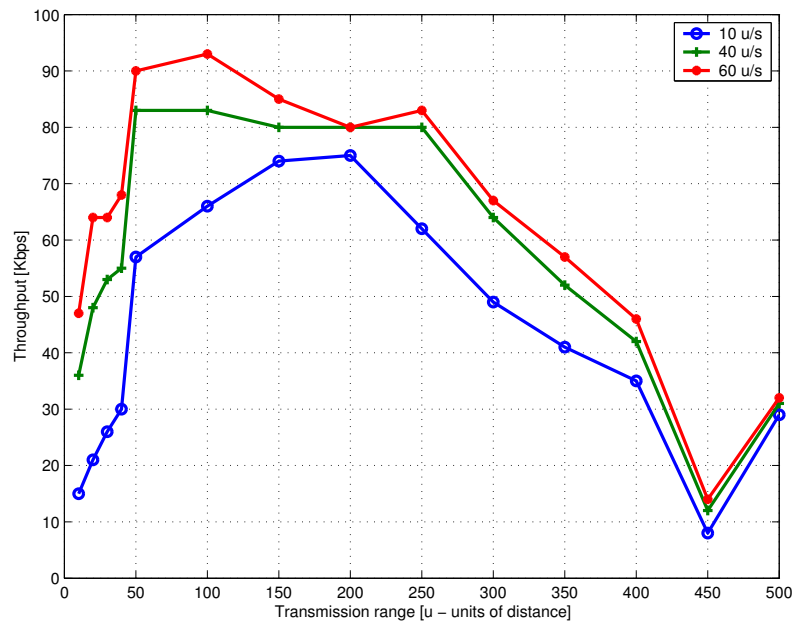
Figure 4.6: Aggregate throughput vs. transmission range - two-hop strategy, $N = 30$, multi-destination traffic model.

Figure 4.7 shows a clear influence of the traffic model on aggregate throughput. In the single destination traffic model, a given source generates packets for only one well determined destination, whereas in the multi-destination traffic model destinations are randomly chosen for each packet. With the latter traffic model, the system exhibits better performances. The distribution of the packets for a given destination among all node queues is indeed a key factor for the throughput of the network. The more these packets are disseminated in the network, the higher is the probability that the destination has a neighbor with a packet for it. Such a dissemination is automatically the case with the multi-destination traffic model.
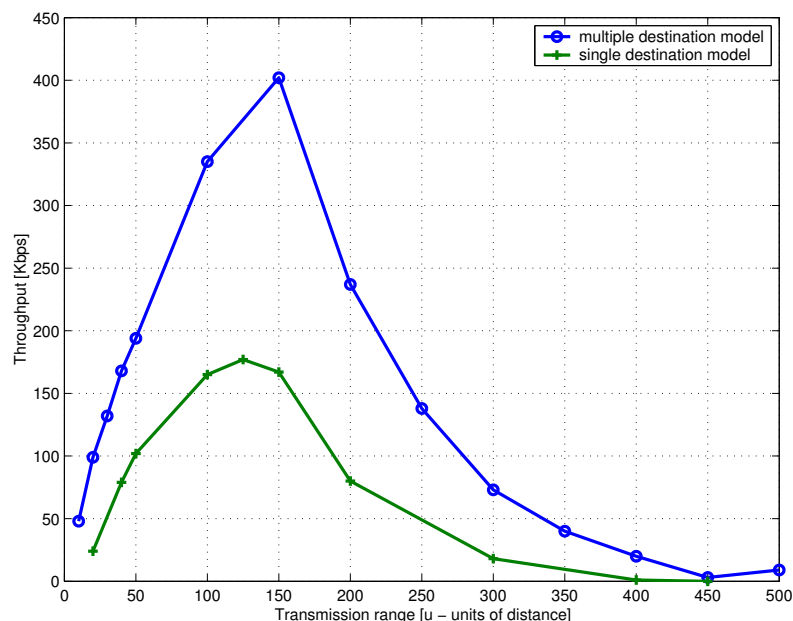
Figure 4.7: Aggregate throughput vs. transmission range - one-hop strategy, $N = 30$, 60 u/s.

With the multi-destination traffic model, a throughput of $\Theta(1)$ can be achieved without relaying. An intuitive explanation is given in figure 4.8 [2]. According to Grossglauser and Tse, the throughput over a direct route from source to destination is $\Theta(1/n)$. This has been proved in the case of the physical model with the assumption that a source has always some packet to send to the destination. With the multi-destination traffic model, there are $n-1$ routes to node D, so that the total average throughput is $\Theta(1)$.

Again figure 4.9 shows that for a higher number of nodes ($N = 50$) the one-hop strategy outperforms the two-hop strategy in the simulation conditions and with

[2]Conclusion of a discussion with D. Tse at the WiOpt'03 conference in Sophia-Antipolis.
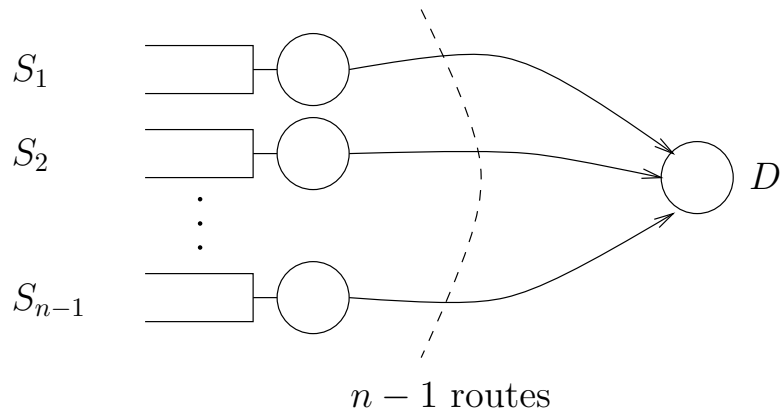
Figure 4.8: Queuing system of the multi-destination traffic model without relaying.

the single-destination traffic model.
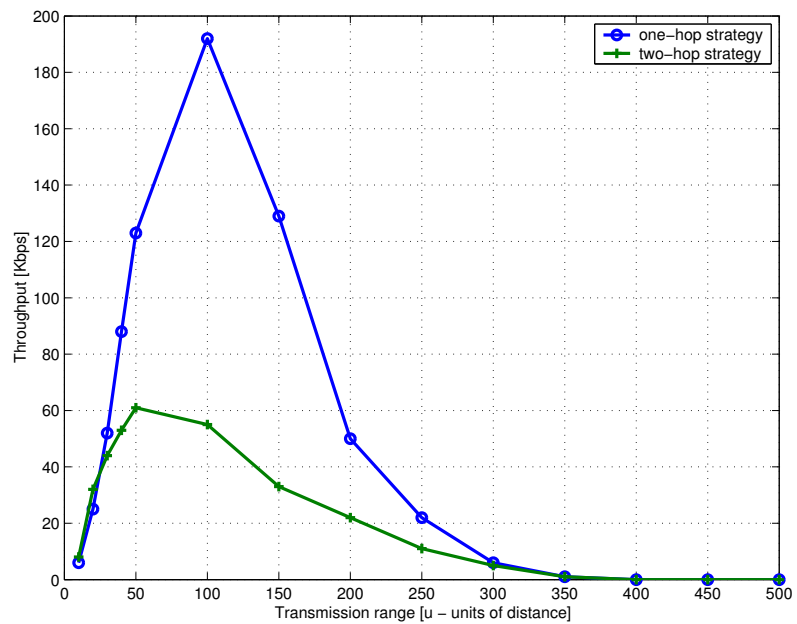
Figure 4.9: Aggregate throughput vs. transmission range - single destination traffic model, $N = 50$, 40 u/s.

4.3.4 Optimal Transmission Range

In this section, we try to derive the optimum transmission range applicable to the one-hop strategy for a given sender density. For that, we consider that during a given time-slot the positions of senders and receivers are two independent Poisson

point processes with density respectively $\theta\lambda$ and $(1-\theta)\lambda$ [3]. This is one snapshot of the simulation. In order to simplify the problem, we also assume that a sender has always something to transmit to the receiver from which it received an RTR. This assumption is not realistic with respect to the previous simulations. However, infinite queues combined with a multi-destination traffic generator make this assumption quite reasonable after some time of simulation[4]. If the edge effects are neglected, the probability of finding k senders in a region of area A is

$$Pr[k \text{ in } A] = \frac{(\theta\lambda A)^k}{k!}e^{-\theta\lambda A}. \tag{4.8}$$

The probability of finding k receivers in a region of area A is

$$Pr[k \text{ in } A] = \frac{((1-\theta)\lambda A)^k}{k!}e^{-(1-\theta)\lambda A}. \tag{4.9}$$

If interference and capture are not taken into account, a sender receives a RTR if and only if there is a single receiver in its transmission range r. Thus, according to equation 4.9 (with $k = 1$ and $A = \pi r^2$), the probability for a sender to receive a RTR is the following:

$$p_1 = (1-\theta)\lambda\pi r^2 e^{-(1-\theta)\lambda\pi r^2}. \tag{4.10}$$

Now, a receiver decodes a data packet if and only if there is a single sender that received a RTR in its transmission range r. Given k the number of senders in the communication disk, this probability is

$$kp_1(1-p_1)^{k-1}. \tag{4.11}$$

---

[3]The notation λ is here preferred to n because in our simulations the network area is not necessary 1 m^2.

[4]It is also assumed in [102].

Thus, according to equations 4.8 and 4.11, the probability for a receiver to receive a data packet is

$$
\begin{aligned}
P &= \sum_{k=1}^{\infty} Pr\big[_{1 \text{ RTR received}}\big|_{k \text{ senders}}\big] Pr\big[_{k \text{ senders}}\big] \qquad (4.12) \\
&= \sum_{k=1}^{\infty} k p_1 (1-p_1)^{k-1} \frac{(\theta \lambda \pi r^2)^k}{k!} e^{-\theta \lambda \pi r^2} \\
&= p_1 \theta \lambda \pi r^2 e^{-\theta \lambda \pi r^2} \sum_{k=1}^{\infty} (1-p_1)^{k-1} \frac{(\theta \lambda \pi r^2)^{k-1}}{(k-1)!} \\
&= p_1 \theta \lambda \pi r^2 e^{-\theta \lambda \pi r^2 p_1} \\
&= \theta(1-\theta)(\lambda \pi r^2)^2 exp\left[-(1-\theta)\lambda \pi r^2 (\theta \lambda \pi r^2 e^{-(1-\theta)\lambda \pi r^2} + 1) \right] \; .
\end{aligned}
$$

In figure 4.10, $P(r)$ is plotted with the parameters of the simulations, $\theta = 1/2$ and $\lambda = 3.10^{-5}$ nodes/u^2. The performance curves for one-hop strategy in figures 4.5 and 4.7 confirm our assumptions related to the optimal transmission range. The difference is due to edge effects, especially for long transmission ranges. Note that
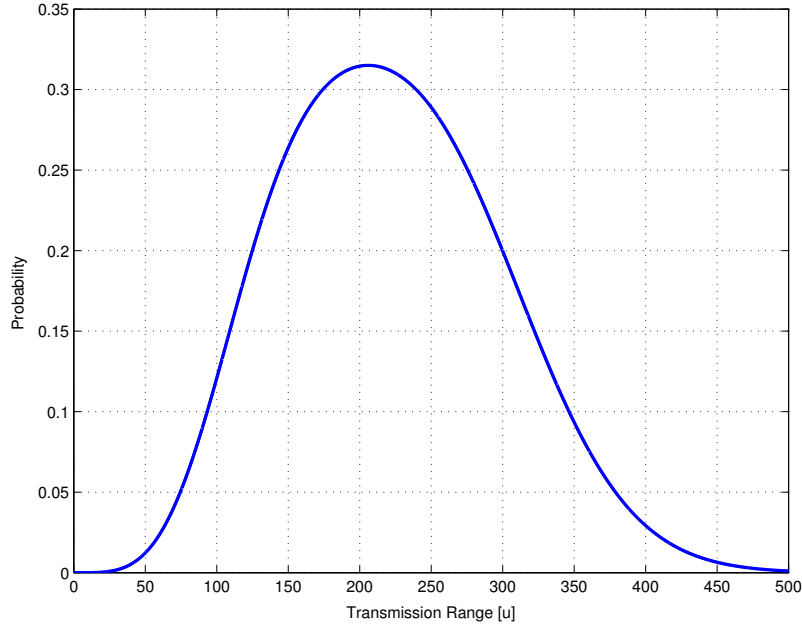
Figure 4.10: Probability for a receiver to receive a data packet, $\theta = 1/2$.

for $\theta = 1/2$, $P(r) = 1/2 p_1 \lambda \pi r^2 e^{-1/2 p_1 \lambda \pi r^2}$. It can be written as follows: $P(r) = y(r) e^{-y(r)}$ with $y(r) = 1/2 p_1 \lambda \pi r^2$. The function $y e^{-y}$ is increasing for $y \geq 0$ until a

maximum at $y = 1$. Now:

$$\forall r, \ y(r) = \left(\frac{1}{2}\lambda\pi r^2\right)^2 e^{-\frac{1}{2}\lambda\pi r^2} \leq 1 \qquad (4.13)$$

because $\forall x, \ x^2 e^{-x} \leq 1$. As a consequence, the optimal transmission range maximizes $y(r)$ and

$$r_{opt} = \frac{2}{\sqrt{\lambda\pi}} \ . \qquad (4.14)$$

For $\theta \in]0, 1[$,

$$y(r) = \theta(1-\theta)(\lambda\pi r^2)^2 e^{-(1-\theta)\lambda\pi r^2} \ . \qquad (4.15)$$

$y(r)$ reaches its maximum for $r_0 = \sqrt{2/((1-\theta)\lambda\pi)}$ and $y(r_0) = 4\theta e^{-2}/(1-\theta)$. Thus, for $\theta \leq 1/(4e^{-2} + 1)$, $y(r_0) \leq 1$ and

$$r_{opt} = \sqrt{\frac{2}{(1-\theta)\lambda\pi}} \ . \qquad (4.16)$$

For $\theta \geq 1/(4e^{-2} + 1)$, there are two optimum transmission ranges that are solutions of the following equation (correcting a typographical error in [9]):

$$\theta(1-\theta)(\lambda\pi r^2)^2 e^{-(1-\theta)\lambda\pi r^2} = 1 \ . \qquad (4.17)$$

This equation has two solutions (see appendix F):

$$r_{opt1} = \sqrt{\frac{2W_0\left(-\sqrt{\frac{1-\theta}{4\theta}}\right)}{\lambda\pi(\theta-1)}} \qquad (4.18)$$

$$r_{opt2} = \sqrt{\frac{2W_{-1}\left(-\sqrt{\frac{1-\theta}{4\theta}}\right)}{\lambda\pi(\theta-1)}} \ , \qquad (4.19)$$

where W_0 is the principal branch of the Lambert W-function, and W_{-1} is the second value [75]. For example with $\theta = 4/5$, $r_{opt1} \approx 194.7$ and $r_{opt2} \approx 478.0$ (see figure 4.11).

With the parameters values of the simulations, $r_{opt} \approx 206.0$ u, and $P(r_{opt}) \approx 0.31$ (see figure 4.10). If the optimum transmission range is chosen at the beginning of the simulation, the probability that a receiver receives a data packet is approximately 0.31 in a given time-slot. Thus, with 15 receivers in average at each time-slot, we get the average number of simultaneous transmissions during a given time-slot: $0.31 \times 15 = 4.65$. Unfortunately, this spatial reuse of the channel is not observed in
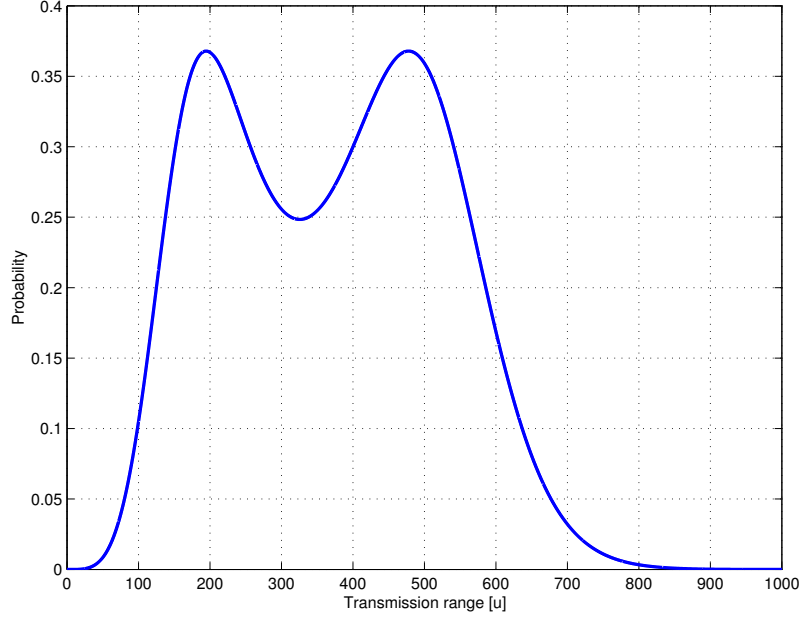
Figure 4.11: Probability for a receiver to receive a data packet, $\theta = 4/5$.

simulations because the senders do not always have packets to send to their nearest receivers.

In figure 4.12, $P(\theta)$ is plotted for $r = 206$ u. As for the optimal transmission range, the computation of the optimal sender density involves the study of the function $y(\theta) = \theta(1 - \theta)(\lambda \pi r^2)^2 e^{-(1-\theta)\lambda \pi r^2}$. If $y \leq 1$, the optimal sender density maximizes $y(\theta)$. In this case,

$$\theta_{opt} = \frac{x - 2 + \sqrt{x^2 + 4}}{2x} \; , \tag{4.20}$$

where $x = \lambda \pi r^2$. Otherwise, θ_{opt} is solution of the equation $y(\theta) = 1$. According to the variations of y, there are two such solutions (see e.g. figure 4.12). Closed-form formulas for them seem difficult to obtain. With the parameters values of the simulations, $\theta_{opt1} \approx 0.7$ and $\theta_{opt2} \approx 0.9$.

From this study, there is unfortunately no practical application, which comes easily in mind. Nevertheless, [102] brings the interesting idea that mobility should not be seen as a constraint but rather as an advantage that could be exploited. In that sense, this work can be related to the concepts of discontinuous coverage or info-stations (see e.g. [87]). In this kind of architecture, small and separated islands of coverage provide huge amounts of traffic for delay-insensitive applications. Mobile terminals retrieve information by moving through these high data rate but small and
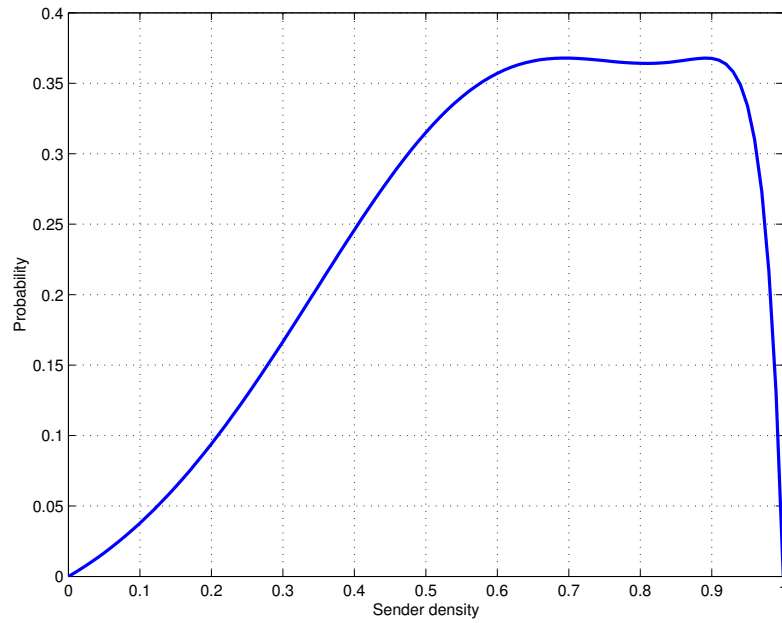
Figure 4.12: Probability for a receiver to receive a data packet as a function of the sender density, $r = 206$ u.

discontinuous cells. Note that a simple modification of CROMA allows a distributed implementation of the scheduling policies. We have also illustrated a variation of the notion of multi-user diversity seen from an original point of view. We now present another application of this concept with the aim to improve the random access.

4.4 The Channel Aware Slotted ALOHA

We saw in the previous section that one way to overcome the capacity results of Gupta and Kumar was to include mobility in the network model. Hereafter, we address the assumption of fixed channel capacity W. [104] didn't take into account the channel variations and the consequent possible opportunistic adaptation of transmission power and rate.

A new application of multi-user diversity for random access is described here. Qin and Berry have proposed in [165] a new medium access control protocol based on slotted ALOHA called channel aware slotted ALOHA. In this protocol, time is divided in equal time-slots where users are allowed to send packets to a central station, e.g. a base-station in a cellular network, an access point in a WLAN, or any node in an ad hoc network. Users transmit a packet on a given slot with a fixed probability p that is related to the quality of the channel.

Note that the studied protocol can be seen as an improvement of the reservation phase of CROMA. If a given receiver has the floor on a slot, it periodically sends out RTRs. These control packets can be used by the requesting nodes to estimate the channel quality. The decision to send a REQ on the first available mini-slot can then be based on this metric. Thus, the receiver plays the role of the central station described above.

One of the main results of [165] is the generic solution for the optimal throughput, in the case of optimal power and rate adaptation (OPRA) over Rayleigh fading channels. In this section, we derive from this solution a closed-form formula for the spectral efficiency as a function of p, we extend the result to Nakagami-m channels and we study the impact of maximum ratio combining (MRC) and selective combining (SC) on the capacity. Moreover, we provide a model for the optimal probability of transmission.

[99] details the general theory of capacity of single-user flat fading channels with channel side information, optimal power adaptation and average power constraint. [38] derived closed-form formulas for this capacity under different adaptive transmission and diversity-combining techniques. Hereafter, we use similar ideas and mathematical framework to obtain the spectral efficiency of the channel aware slotted ALOHA protocol.

4.4.1   System Model

In the considered system, there is a single-cell with a central station, e.g. a base-station or an access point, and users attempting to send information to this central point behave like requesting nodes in a CROMA network (see figure 4.13). In this latter case, a receiver having the floor on a slot plays the role of the central station.
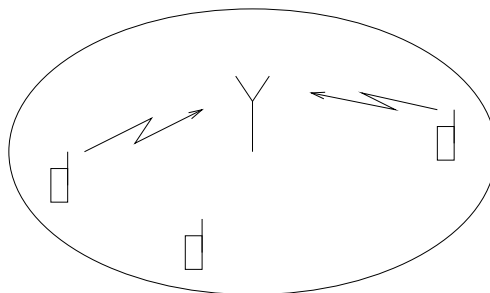
Figure 4.13: System model: Terminals send data packets to a central station.

Only uplink communications based on distributed and random access protocol are considered. This protocol is based on the traditional slotted ALOHA: Time is divided in equal time-slots, where terminals are allowed to transmit with a probability p. If several users take the same decision of transmitting during a given time-slot, a collision occurs at the central station and all packets are lost. In our study, we exclude the possibility of capture. Now, let T be the event of a successful transmission on a time-slot. We suppose that all users are backlogged, i.e., they have always something to send. If there are n users, we have:

$$P[T] = np(1-p)^{n-1} \ . \tag{4.21}$$

In the slotted ALOHA protocol, the decision to transmit is not correlated to the channel conditions experienced by the user. In the channel aware slotted ALOHA protocol, users are allowed to transmit only if their channel quality is good enough, i.e., if their signal-to-noise ratio (SNR), γ, is above a given threshold, γ_0 chosen in order to match the probability p of transmission. Let $F(\gamma)$ be the complementary distribution function (cdf) and $p_\gamma(\gamma)$ the probability distribution function (pdf) of

the SNR. We have for γ_0:

$$F(\gamma) = \int_{\gamma}^{+\infty} p_{\gamma}(\gamma)d\gamma \tag{4.22}$$

$$\gamma_0 = F^{-1}(p) . \tag{4.23}$$

The channel aware slotted ALOHA protocol assumes that each user is aware of its own channel conditions and that all users experience the same SNR distribution, including the same mean SNR. The first assumption is valid if we consider that the central station periodically sends a beacon frame (e.g. RTRs) on the same channel. The second assumption is realistic if users are in the same environment, at a similar distance from the central point.

4.4.2 Capacity with Optimal Power and Rate Adaptation (OPRA)

We now assume that the channel changes at a rate much slower than the data rate, so the uplink is a block fading channel. It is also assumed that there is no time-correlation of the channel between blocks. The general theory for the capacity of such fading channels has been developed in [99]. It has been shown that the fading channel capacity with channel side information at both the transmitter and receiver[5] is achieved when the transmitter adapts its power, data rate, and coding scheme to the channel variations. Given an average power constraint S, the time-varying channel capacity in bits/Hz/s is defined in [99] by:

$$C_{opra} = \max \int_{\gamma} \log_2(1 + \frac{S(\gamma)\gamma}{S})p_{\gamma}(\gamma)d\gamma \text{ s.t. } \int_{\gamma} S(\gamma)p_{\gamma}(\gamma)d\gamma \leq S . \tag{4.24}$$

In our case, the SNR takes values in $[\gamma_0; \infty[$ because of the *cut-off* value imposed by the protocol:

$$C_{opra} = \max \int_{\gamma \geq \gamma_0} \log_2(1 + \frac{S(\gamma)\gamma}{S})p_{\gamma}(\gamma)d\gamma \text{ s.t. } \int_{\gamma \geq \gamma_0} S(\gamma)p_{\gamma}(\gamma)d\gamma \leq S . \tag{4.25}$$

---

[5]This is a realistic model for a slowly varying channel with channel estimation and transmitter feedback [99].

We get by using the Lagrange multiplier (see appendix G and e.g. [174] for more details):

$$C_{opra} = \int_{\max(\gamma_0,\lambda)}^{+\infty} \log_2\left(\frac{\gamma}{\lambda}\right) p_\gamma(\gamma)d\gamma \, , \qquad (4.26)$$

where λ is a constant which must satisfy:

$$\int_{\max(\gamma_0,\lambda)}^{+\infty} \left(\frac{1}{\lambda} - \frac{1}{\gamma}\right) p_\gamma(\gamma)d\gamma = 1 \, . \qquad (4.27)$$

Now, let us look at the capacity of the protocol:

$$
\begin{aligned}
C &= E\left[\log_2(\gamma/\lambda)\right] \qquad\qquad\qquad\qquad\qquad\qquad (4.28)\\
&= P[T]E\left[\log_2(\gamma/\lambda)|T\right]\\
&= np(1-p)^{n-1}E\left[\log_2(\gamma/\lambda)|\gamma \geq \max(\lambda,\gamma_0)\right] \, .
\end{aligned}
$$

From this expression, we deduce the spectral efficiency of the channel aware slotted ALOHA protocol in case of optimal power and rate adaptation with average transmit power constraint in bits/Hz/s:

$$C = np(1-p)^{n-1}\frac{\int_{\max(\gamma_0,\lambda)}^{+\infty} \log_2\left(\frac{\gamma}{\lambda}\right) p_\gamma(\gamma)d\gamma}{\int_{\max(\gamma_0,\lambda)}^{+\infty} p_\gamma(\gamma)d\gamma} \, . \qquad (4.29)$$

Note that if $\gamma_0 \geq \lambda$, $\int_{\max(\gamma_0,\lambda)}^{+\infty} p_\gamma(\gamma)d\gamma = p$ and C can be written:

$$C = n(1-p)^{n-1}\int_{\gamma_0}^{+\infty} \log_2\left(\frac{\gamma}{\lambda}\right) p_\gamma(\gamma)d\gamma \, . \qquad (4.30)$$

4.4.3   OPRA Capacity in Rayleigh Channels

Hereafter, we assume that the uplink is a Rayleigh fading channel. So the pdf of the SNR is given by the exponential distribution:

$$p_\gamma(\gamma) = \frac{e^{-\gamma/\bar{\gamma}}}{\bar{\gamma}} \, , \qquad (4.31)$$

where $\bar{\gamma}$ is the average received SNR. If we now consider combining techniques, the pdf of the received SNR is modified. For the MRC technique with known branch amplitudes and phases, the pdf of the SNR output of an M-branch MRC combiner

is given by:

$$p_\gamma(\gamma) = \frac{\gamma^{M-1} e^{-\gamma/\bar{\gamma}}}{(M-1)! \bar{\gamma}^M} \, , \tag{4.32}$$

where $\bar{\gamma}$ is the average SNR on each branch, and branch signals are assumed to be independent and identically distributed (iid).

In the SC technique, only the best branch is chosen for processing. Again, assuming iid branch signals, the pdf of the SNR output can be written as:

$$p_\gamma(\gamma) = M \left(1 - e^{-\gamma/\bar{\gamma}}\right)^{M-1} \frac{e^{-\gamma/\bar{\gamma}}}{\bar{\gamma}} \, . \tag{4.33}$$

In the following, closed-form formulas for channel aware slotted ALOHA protocol spectral efficiency over Rayleigh fading channels are derived for reception without diversity scheme, with MRC and with SC diversity schemes.

Capacity Without Diversity Scheme

From equation 4.23, we get

$$\gamma_0 = -\bar{\gamma} \ln(p). \tag{4.34}$$

Assuming $\lambda \leq \gamma_0$ leads to $\max(\lambda, \gamma_0) = \gamma_0$. By substituting equation 4.31 in equation 4.27, the power constraint is:

$$\int_{\gamma_0}^{+\infty} \left(\frac{1}{\lambda} - \frac{1}{\gamma}\right) \frac{e^{-\gamma/\bar{\gamma}}}{\bar{\gamma}} d\gamma = 1 \, . \tag{4.35}$$

After some calculations λ can be written as:

$$\lambda = \frac{\gamma_0 E_0(\gamma_0/\bar{\gamma})}{\bar{\gamma} + E_1(\gamma_0/\bar{\gamma})} \tag{4.36}$$

$$\lambda = \frac{\bar{\gamma} p}{\bar{\gamma} + E_1(-\ln p)} \, , \tag{4.37}$$

where $E_n(x)$ is the exponential integral of order n defined by:

$$E_n(x) = \int_1^{+\infty} t^{-n} e^{-xt} dt \, , x \geq 0 \, . \tag{4.38}$$

If $\gamma_0 \leq \lambda$, the power constraint is:

$$\int_\lambda^{+\infty} \left(\frac{1}{\lambda} - \frac{1}{\gamma}\right) \frac{e^{-\gamma/\bar{\gamma}}}{\bar{\gamma}} d\gamma = 1 \, , \tag{4.39}$$

and λ is the solution of the following equation:

$$E_0\left(\frac{\lambda}{\bar{\gamma}}\right) - E_1\left(\frac{\lambda}{\bar{\gamma}}\right) = \bar{\gamma} \,. \tag{4.40}$$

We can easily show by derivation that this equation has a unique solution (see [38] for a similar demonstration).

To find the relationship between channel capacity and p, let $g(p)$ be the expression of λ when $\lambda \leq \gamma_0$, i.e., $g(p) = \frac{\bar{\gamma}p}{\bar{\gamma}+E_1(-\ln p)}$. Equation 4.36 is valid only for $g \leq \gamma_0$. So let us study the sign of $f(p) = p + \ln(p)(\bar{\gamma} + E_1(-\ln p))$, which has the same sign than $g(p) - \gamma_0(p)$. Taken into account that $E_1'(x) = -E_0(x)$, the derivative of f,

$$f'(p) = \frac{\bar{\gamma} + E_1(-\ln p)}{p} \,, \tag{4.41}$$

is clearly positive for p between 0 and 1. Moreover, $\lim_{p\to 0} f(p) = -\infty$.

Let us look at the behavior of f when p tends towards 1. Let $x = -\ln(p)$ and let us use an asymptotic expression of E_1 (see [101] p.927, $E_1(x) = -E_i(-x)$):

$$E_1(x) = -C - \ln(x) - \sum_{k=1}^{\infty} \frac{(-1)^k x^k}{k.k!} \tag{4.42}$$

$$xE_1(x) = -xC - x\ln(x) - \sum_{k=1}^{\infty} \frac{(-1)^k x^{k+1}}{k.k!} \tag{4.43}$$

$$\lim_{x\to 0} xE_1(x) = 0 \tag{4.44}$$

Now, it is straightforward that $\lim_{p\to 1} f(p) = 1$

So, let us define p^* as the unique solution of $f(p) = 0$. For $p \leq p^*$, λ is given by equation 4.36 and for $p \geq p^*$, λ verifies equation 4.40.

However, the case $p \geq p^*$ is not the most interesting one. Indeed, figure 4.14 shows p^* as a function of $\bar{\gamma}$ and figure 4.15 shows the capacity of the protocol for a small number of users, $n = 5$, as a function of p. It is clear that for $p \geq p^*$, the capacity is very low and the protocol parameter is not well dimensioned. We will see that the most interesting range of values for p is the neighborhood of $1/n$, where the capacity reaches its maximum. In the following steps, we will neglect this case and focus on situations where $\lambda \leq \gamma_0$.
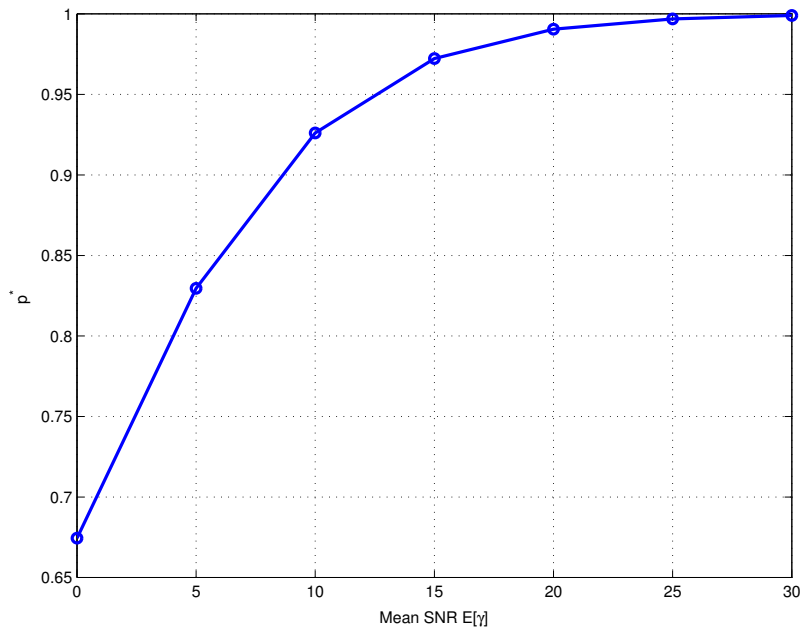
Figure 4.14: p^* as a function of the mean SNR $\bar{\gamma} = E[\gamma]$.
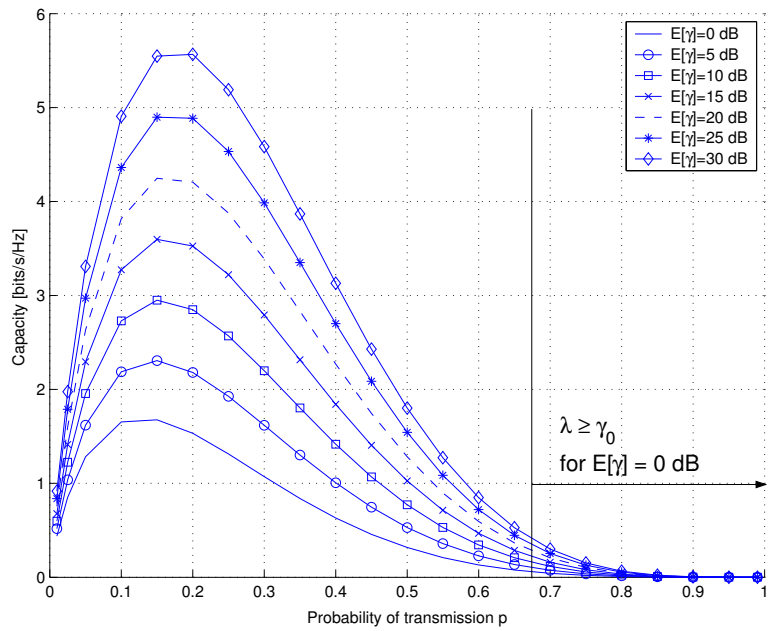
Figure 4.15: Capacity as a function of the mean SNR $\bar{\gamma} = E[\gamma]$ for $n = 5$ users.

By substituting equation 4.31 in equation 4.30:

$$C = n(1-p)^{n-1} \int_{\gamma_0}^{+\infty} \log_2\left(\frac{\gamma}{\lambda}\right) \frac{e^{-\gamma/\bar{\gamma}}}{\bar{\gamma}} d\gamma \ , \tag{4.45}$$

that can be reduced to:

$$C = n(1-p)^{n-1} \log_2(e)\left(\ln\left(\frac{\gamma_0}{\lambda}\right) e^{-\gamma_0/\bar{\gamma}} + \frac{\gamma_0}{\bar{\gamma}} J_1(\gamma_0/\bar{\gamma})\right) \ , \tag{4.46}$$

where the integral $J_n(\mu)$ are defined by:

$$J_n(\mu) = \int_1^{+\infty} t^{n-1} \ln(t) e^{-\mu t} dt, \ \mu > 0, \ n = 1, 2, \ldots \ . \tag{4.47}$$

The integration by parts of $J_1(\mu)$ yields: $J_1(\mu) = E_1(\mu)/\mu$. We obtain the capacity in bits/Hz/s:

$$C = n(1-p)^{n-1} \log_2(e) \times \left(\ln\left(\frac{\gamma_0}{\lambda}\right) e^{-\gamma_0/\bar{\gamma}} + E_1(\gamma_0/\bar{\gamma})\right) \ . \tag{4.48}$$

Capacity With Maximum Ratio Combining

From equations 4.23 and 4.32, we get

$$p = \frac{\Gamma(M, \gamma_0/\bar{\gamma})}{(M-1)!} \ , \tag{4.49}$$

where $\Gamma(\alpha, x)$ is the complementary incomplete gamma function:

$$\Gamma(\alpha, x) = \int_x^{+\infty} t^{\alpha-1} e^{-t} dt \ . \tag{4.50}$$

Again, we focus on the case $\max(\lambda, \gamma_0) = \gamma_0$, and we substitute equation 4.32 in the power constraint relation (equation 4.27). This gives us the expression of λ:

$$\lambda = \frac{\bar{\gamma}\Gamma(M, \gamma_0/\bar{\gamma})}{(M-1)!\bar{\gamma} + \Gamma(M-1, \gamma_0/\bar{\gamma})} \ . \tag{4.51}$$

If we substitute the pdf of the SNR (equation 4.32) in the capacity expression (equation 4.30), we get:

$$C = n(1-p)^{n-1} \int_{\gamma_0}^{+\infty} \log_2\left(\frac{\gamma}{\lambda}\right) \frac{\gamma^{M-1} e^{-\gamma/\bar{\gamma}}}{(M-1)!\bar{\gamma}^M} d\gamma \ . \tag{4.52}$$

We can simplify this expression by introducing Γ and J_M:

$$C = \frac{n(1-p)^{n-1}\log_2(e)\gamma_0^M}{(M-1)!\bar{\gamma}^M}\left(\ln\left(\frac{\gamma_0}{\lambda}\right)\left(\frac{\bar{\gamma}}{\gamma_0}\right)^M\Gamma(M,\gamma_0/\bar{\gamma}) + J_M(\gamma_0/\bar{\gamma})\right) . \quad (4.53)$$

J_M can be numerically evaluated using the expression proposed in the appendix of [38]:

$$J_n(\mu) = \frac{(n-1)!}{\mu^n}\sum_{k=0}^{n-1}\frac{\Gamma(k,\mu)}{k!} . \quad (4.54)$$

Capacity With Selective Combining

Let us determine γ_0 through equation 4.23 and the SNR pdf in the SC case (equation 4.33):

$$\gamma_0 = -\bar{\gamma}\ln\left(1 - (1-p)^{1/M}\right) . \quad (4.55)$$

As before, we focus on the case $\max(\lambda, \gamma_0) = \gamma_0$, and we find the expression of λ using the constraint relation equation 4.27:

$$\lambda = \frac{1 - \left(1 - e^{-\gamma_0/\bar{\gamma}}\right)^M}{1 + \frac{M}{\bar{\gamma}}\sum_{k=0}^{M-1}(-1)^k\binom{M-1}{k}E_1\left(\frac{(1+k)\gamma_0}{\bar{\gamma}}\right)} . \quad (4.56)$$

We now substitute equation 4.33 in equation 4.30 in order to obtain the capacity with a SC scheme:

$$C = n(1-p)^{n-1}\int_{\gamma_0}^{+\infty}\frac{M}{\bar{\gamma}}\log_2\left(\frac{\gamma}{\lambda}\right)\left(1 - e^{-\gamma/\bar{\gamma}}\right)^{M-1}e^{-\gamma/\bar{\gamma}}d\gamma . \quad (4.57)$$

This expression can be rewritten in a closed-form formula:

$$C = n(1-p)^{n-1}\log_2(e)\frac{M}{\bar{\gamma}} \quad (4.58)$$

$$\sum_{k=0}^{M-1}(-1)^k\binom{M-1}{k}\left(\frac{\bar{\gamma}\ln(\gamma_0/\lambda)}{1+k}e^{-\frac{(1+k)\gamma_0}{\bar{\gamma}}} + \gamma_0 J_1\left(\frac{(1+k)\gamma_0}{\bar{\gamma}}\right)\right) .$$

We can also write this expression using the traditional exponential integral:

$$C = n(1-p)^{n-1} \log_2(e) \frac{M}{\bar{\gamma}} \tag{4.59}$$

$$\sum_{k=0}^{M-1} (-1)^k \begin{pmatrix} M-1 \\ k \end{pmatrix} \left(\frac{\bar{\gamma} \ln(\gamma_0/\lambda)}{1+k} e^{-\frac{(1+k)\gamma_0}{\bar{\gamma}}} + \frac{\bar{\gamma}}{(1+k)\gamma_0} E_1 \left(\frac{(1+k)\gamma_0}{\bar{\gamma}} \right) \right) \ .$$

4.4.4 OPRA Capacity in Nakagami-m Channels

In Nakagami-m channels the pdf of the SNR is the following:

$$p_\gamma(\gamma) = \frac{m^m \gamma^{m-1}}{\bar{\gamma}^m \Gamma(m)} e^{-\frac{m\gamma}{\bar{\gamma}}} \ , \tag{4.60}$$

where m is the fading parameter and $\Gamma(m) = (m-1)!$. We can easily note that this expression is similar to equation 4.32 that gives the pdf of SNR at the output of the MRC. All results can be obtained from the previous study for MRC by replacing M by m and $\bar{\gamma}$ by $\frac{\bar{\gamma}}{m}$. For γ_0, λ, and C, we get:

$$p = \frac{\Gamma(m, m\gamma_0/\bar{\gamma})}{(m-1)!} \ , \tag{4.61}$$

$$\lambda = \frac{\frac{\bar{\gamma}}{m} \Gamma(m, m\gamma_0/\bar{\gamma})}{(m-1)! \frac{\bar{\gamma}}{m} + \Gamma(m-1, m\gamma_0/\bar{\gamma})} \ , \tag{4.62}$$

$$C = \frac{n(1-p)^{n-1} \log_2(e) m^m \gamma_0^m}{(m-1)! \bar{\gamma}^m} \left(\ln\left(\frac{\gamma_0}{\lambda}\right) \left(\frac{\bar{\gamma}}{m\gamma_0}\right)^m \Gamma(m, m\gamma_0/\bar{\gamma}) + J_m(m\gamma_0/\bar{\gamma}) \right) \ . \tag{4.63}$$

4.4.5 Numerical Results

In this section, we pay special attention to the influence of average SNR $\bar{\gamma}$ and of probability of transmission p on the capacity. Moreover, we give a model for the optimal value of p as a function of n and $\bar{\gamma}$ for Rayleigh channels without diversity scheme.

In this latter case, figures 4.15 and 4.16 show the capacity of the channel aware slotted ALOHA protocol as a function of p and for different values of $\bar{\gamma}$. The number of users is respectively $n = 5$ and $n = 10$. We see of course a clear increase in capacity with increasing $\bar{\gamma}$. Maximum capacity is obtained at a value of p that we call p_{opt} and it is around $1/n$. This can be seen in figure 4.17 for $n = 5, 7$, and 10 users.
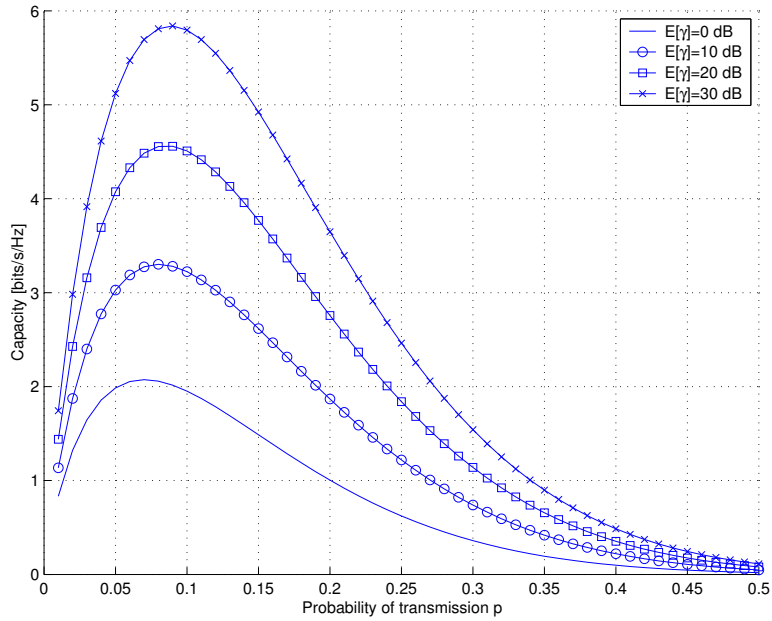
Figure 4.16: Capacity as a function of the mean SNR $\bar{\gamma} = E[\gamma]$ for n=10 users.

The main result is that p_{opt} is slightly different from the traditional value of $1/n$ but tends towards $1/n$ when $\bar{\gamma}$ tends towards infinity.

We are now interested in characterizing how far p_{opt} is from $1/n$ and what is the loss in capacity if users choose $p = 1/n$ instead of p_{opt}. For that, we restrict n to be between 3 and 20 and $\bar{\gamma}$ between 0 dB and 30 dB, thus covering a sufficiently large range of realistic scenarios. For higher values of n and $\bar{\gamma}$, the model is more difficult to obtain and $n = 2$ seems to be a singular point.

We consider the following family of functions to be fitted with the numerical computation of p_{opt} as a function of n and $\bar{\gamma}_{db}$, the average SNR in dB:

$$p_e(n, \bar{\gamma}_{db}) = \frac{1}{n} - \alpha_1(n)e^{-\alpha_2(n)\bar{\gamma}_{db}} \left(\alpha_3(n)\bar{\gamma}_{db}^2 + \alpha_4(n)\bar{\gamma}_{db} + \alpha_5(n) \right) . \qquad (4.64)$$

The least square method provides a very good approximation of p_{opt}, as shown on figure 4.18 for $n = 5, 7$, and 10 for example.

The same method is used in order to model the parameter α_i as functions of n. These functions are chosen to be either hyperbolic or polynomial (with maximum order 2), so that the expression of p_e is kept simple. We have deduced the following
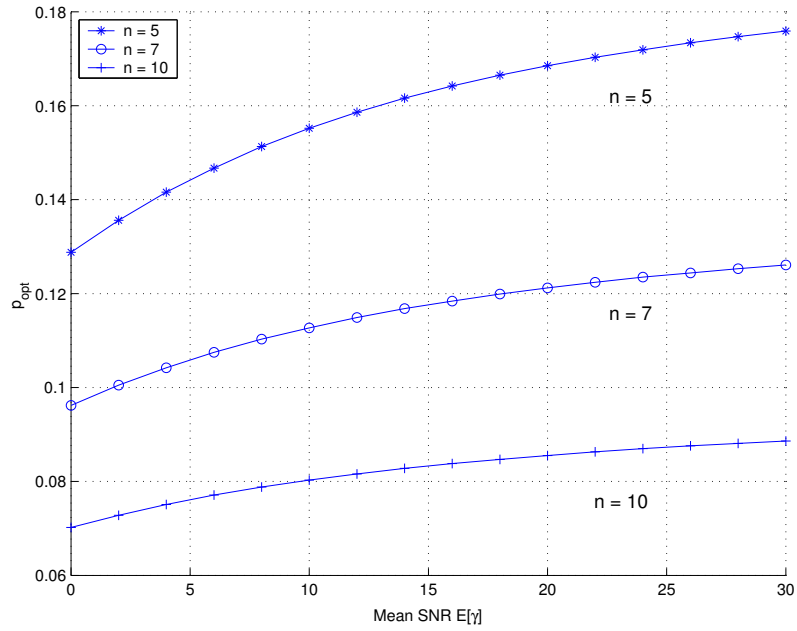
Figure 4.17: Optimal probability p_{opt} as a function of the mean SNR $\bar{\gamma} = E[\gamma]$ for n=5, 7, and 10 users.
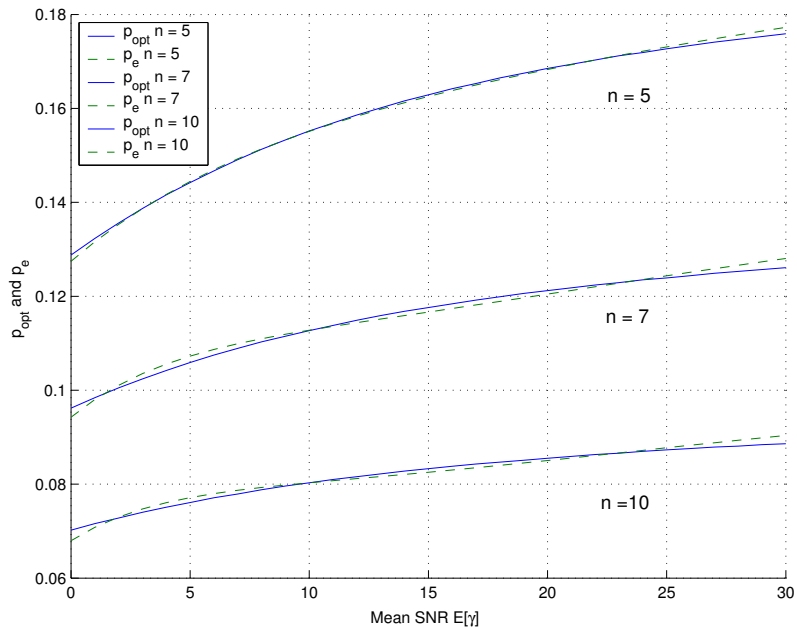
Figure 4.18: Optimal probability p_{opt} and estimate p_e as a function of the mean SNR $\bar{\gamma} = E[\gamma]$ for n=5, 7, and 10 users.

expressions:

$$\alpha_1 = \frac{0.1972}{n} - 0.0048 \tag{4.65}$$

$$\alpha_2 = -4.2148.10^{-4}n + 3.76.10^{-2} \tag{4.66}$$

$$\alpha_3 = -1.605.10^{-6}n^2 + 1.6756.10^{-5}n \tag{4.67}$$

$$+8.762.10^{-4}$$

$$\alpha_4 = 1.014.10^{-3}n - 3.672.10^{-2} \tag{4.68}$$

$$\alpha_5 = \frac{0.22}{n} + 2.0685 \tag{4.69}$$
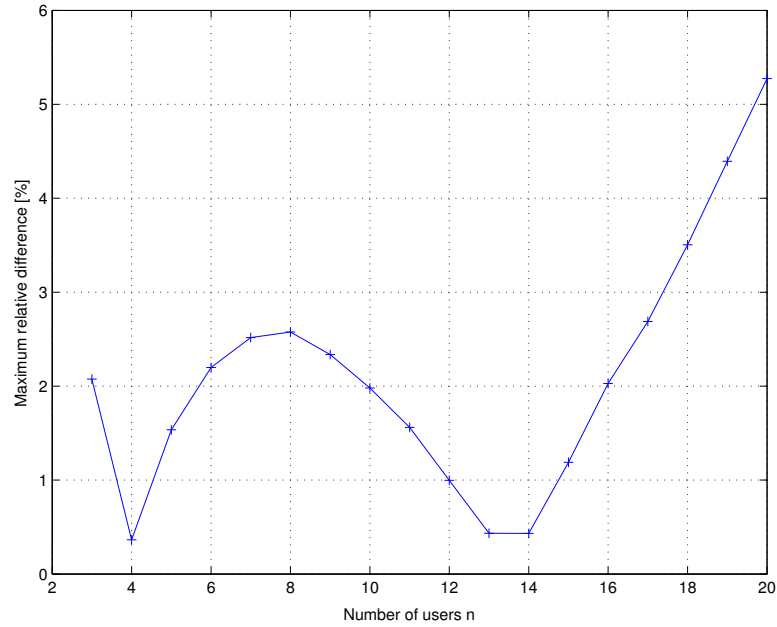
Figure 4.19: Maximum relative difference between the model (p_e) and the optimal probability p_{opt}.

Figure 4.19 shows the maximum error in the model of p_{opt}. The maximum difference always stays under 6% for $3 \leq n \leq 20$.

Let us look at the influence of multi-user diversity on capacity. Figure 4.20 shows the maximum spectral efficiency as a function of n for $\bar{\gamma} = 0, 15$, and 30 dB. The capacity at $p = 1/n$, which is a sub-optimal choice for p, is also shown. We can see that the loss in capacity for $p = 1/n$ is rather small. And $1/n$ is very close to the optimal choice when $\bar{\gamma}$ is high. As expected, the capacity increases with the number of users thanks to the multi-user diversity.
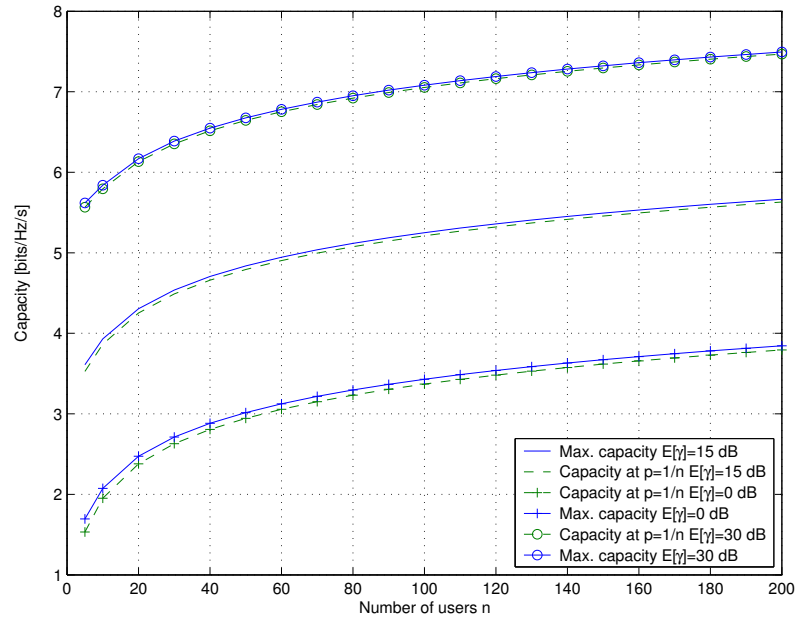
Figure 4.20: Influence of the number of users on the maximum capacity and on the capacity at $p = 1/n$ for $\bar{\gamma} = 0, 15$, and 30 dB.
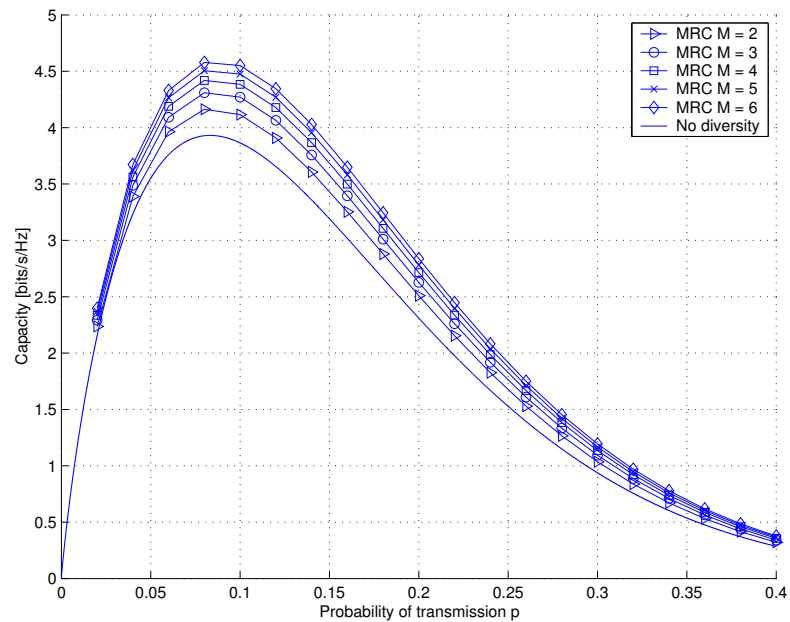
Figure 4.21: Capacity over a Rayleigh channel with MRC and without diversity scheme for $\bar{\gamma} = 15$ dB and $n = 10$.
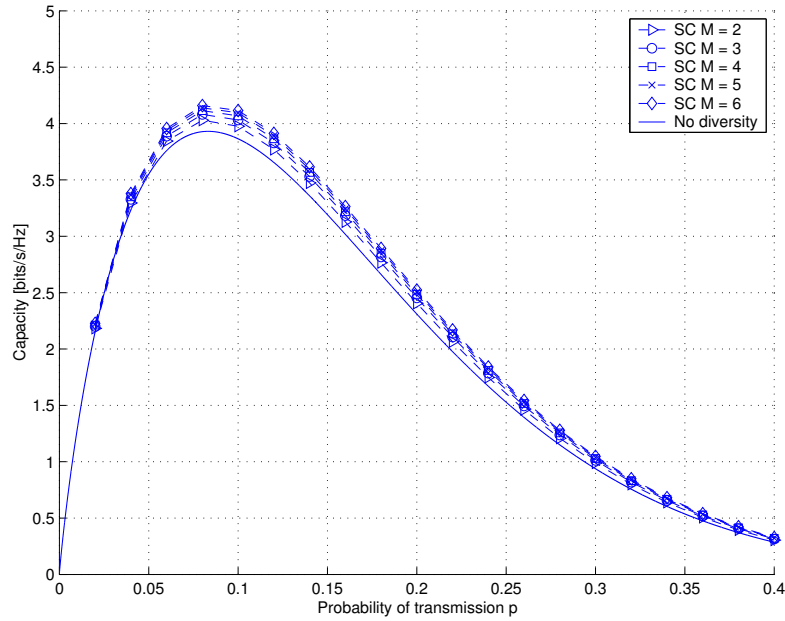
Figure 4.22: Capacity over a Rayleigh channel with SC and without diversity scheme for $\bar{\gamma} = 15$ dB and $n = 10$.

This maximum capacity can be improved by using diversity combining techniques like MRC or SC. In figures 4.21 and 4.22, the capacity of the protocol with MRC and SC over a Rayleigh channel is provided for $\bar{\gamma} = 15$ dB and $n = 10$ as a function of p and for several orders of diversity. We note that p_{opt} is independent on the order of diversity and that MRC is always more efficient than SC for a given order, which is a classical result [163].

The effect of multi-user diversity is also illustrated for the diversity combining schemes, as shown in figures 4.23 and 4.24: for $M = 6$, a gain of 0.5 bits/Hz/s is achieved with MRC and 0.13 bits/Hz/s with SC.

As a conclusion, the impact of multi-user diversity on network capacity is clearly visible. The channel aware slotted ALOHA is also a good candidate for the reservation scheme of CROMA, which relies on random access. Additional applications of MAC and physical layers interaction are given in the following section.

Glossary for section 4.4

$\Gamma(\alpha, x)$	Complementary incomplete gamma function.
γ	SNR.
γ_0	SNR threshold needed for transmission.

$\bar{\gamma}$	Average SNR.
λ	Lagrange multiplier.
C	Capacity in bits/s/Hz.
$E_n(x)$	Exponential integral.
M	Number of branches for MRC or SC diversity.
m	Parameter of the Nakagami-m channels.
n	Number of nodes.
p	Probability of transmission on a slot.
$p_\gamma(\gamma)$	Pdf of the SNR.
p_e	Approximation of the optimal probability of transmission p.
S	Average power constraint.
$S(\gamma)$	Transmit power.
T	Event of a successful transmission on a slot.
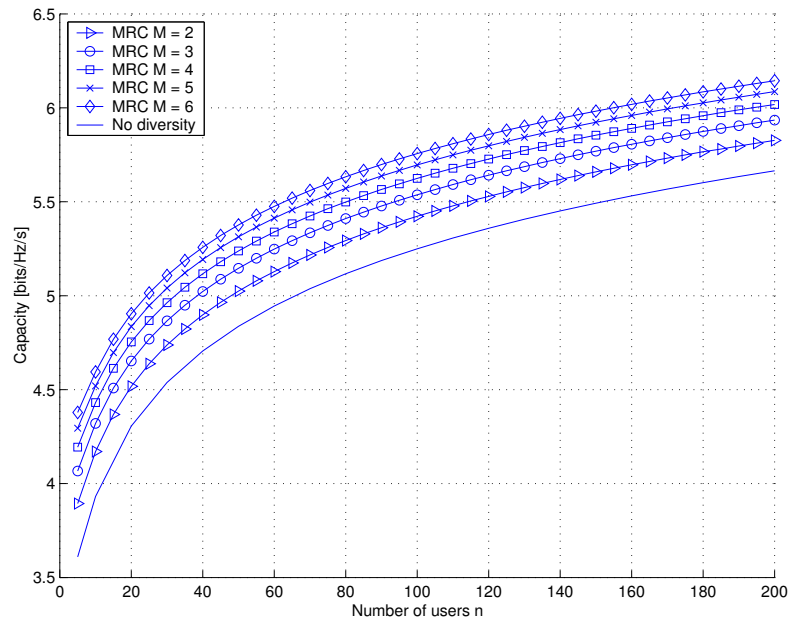
Figure 4.23: Influence of the number of users on capacity at $p = 1/n$ for $\bar{\gamma} = 15$ dB, with MRC and without diversity scheme.
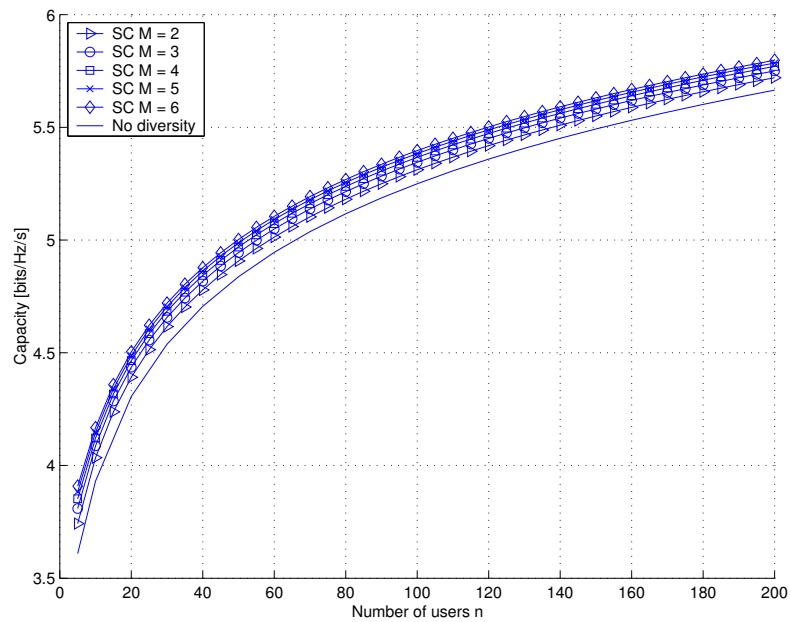
Figure 4.24: Influence of the number of users on capacity at $p = 1/n$ for $\bar{\gamma} = 15$ dB, with SC and without diversity scheme.

4.5 Multi-packet Reception

Further improvement of CROMA reservation scheme is feasible through implementation of multi-packet reception. In [104], a successful transmission is fundamentally seen as a spatial utilization of the medium in the bounded area of the network. For example, equation 4.1 implies that disks of radius $\Delta/2$ times the lengths of hops centered at the receivers in the same slot are disjoint. And according to equation 4.3, disks of radius $\Delta r/2$ centered at each receiver are disjoint. These considerations leads to upper bounds for capacity in [104]. In this section, a model assuming overlaps between disks of radius $\Delta r/2$ around the receiver nodes is an essential part of the study because multi-packet reception is assumed. The advantage of such a technique is shown in the case of the slotted ALOHA protocol.

The spatial capacity of the slotted ALOHA protocol with capture effect has been studied in [153]. This capacity has been obtained with the assumption that receivers are able to decode at most a single packet per slot. However, research performed since the early 1980's in the domain of multi-user detection in CDMA systems [195] suggests that this situation can be improved. Indeed, receivers using multi-user detection schemes can decode the packets from several transmitters simultaneously. In particular, the near-far resistance of the multi-user detectors [146] makes this technique very attractive for ad hoc networks, where power control schemes are much more difficult to implement than in traditional single-hop systems.

In this section, we extend the results of [153] for multi-packet reception and we provide a closed-form formula for the throughput of the slotted ALOHA as a function of the probability, $r_{n,k}$, for a receiver to decode k packets given that n have been transmitted in its neighborhood. Then, we detail three different models of multi-packet reception: a simple model often used in the literature, a bank of matched filters (MF), and a linear minimum mean-square error (MMSE) multi-user detector. Finally, we provide numerical results and highlight the near-far resistance of the MUD scheme.

4.5.1 Spatial Throughput of the Multi-hop Slotted ALOHA

Models

Throughout this section, we will consider an ad hoc network of nodes, spatially distributed in a plane according to a Poisson process with parameter λ. That means

that the probability to find k nodes in any region, A, of area $S(A)$ is:

$$P[k \text{ in } A] = \frac{(\lambda S(A))^k}{k!} e^{-\lambda S(A)} . \tag{4.70}$$

A large area network is assumed such that the edge effects can be neglected.

All nodes are assumed to operate with half-duplex radios. This means that a collision of the second order can occur if a node receives a packet, while it is itself transmitting during the same slot. In this case, the packet is lost. The transmit power is constant and equals P_0.

Equal sized time-slots are assumed for the implementation of slotted ALOHA protocol. During a given time-slot every node has a fixed probability p of transmitting a packet. Otherwise, it is able to receive one or several packets transmitted by its neighboring transmitters. A neighborhood distance R_0 related to receiver sensitivity as well as $r_{n,k}$ the probability of detection of k packets out of n transmitted in the range R_0 are defined.
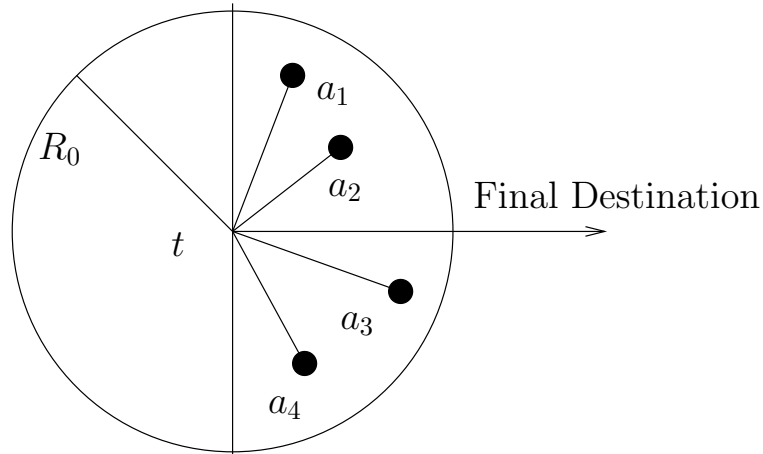
Figure 4.25: Routing model: The next hop after node t is chosen at random among neighbors (a_i, $i = 1...4$) that lie in the direction of the final destination.

The transmitted packets are equally likely to be routed to one of the neighboring nodes in the direction of the final destination [6] (see figure 4.25).

First of all, we are interested in the local throughput of the system, i.e., the expected number of packet received per slot. We will then evaluate the expected forward progress of a packet and conclude our study with the total throughput of

---

[6]The results obtained with simplifying assumption taken in [153] could be further refined on the basis of a more realistic model, e.g. the Markov path on the Poisson-Delaunay graph [46].

the network.

Preliminary Results

A particular node a is considered and we define the random variable X as the number of correctly decoded packets with destination a in a given slot and two events: (A) the event that a does not transmit; (T) the event that a particular sender t sends a packet to a, we have the two basic results (given in [153] and recalled in appendix H):

$$P[A] = 1 - p \, , \tag{4.71}$$

$$P[T] = \frac{1 - e^{-\lambda \pi R_0^2/2}}{\lambda \pi R_0^2} \, . \tag{4.72}$$

The variables p, λ, and R_0 are defined above and the senders at a given time-slot are spatially distributed according to a Poisson process with density λp [88].

Local Throughput

Let us define two more events: (T_n) the event that there are n senders in the neighborhood of a; (D_k) the event that a decodes exactly k packets in the given time-slot. The probability that a receives x packets given (A), (T_n), and (D_k) is:

$$P[X = x | A, T_n, D_k] = \binom{k}{x} P[T]^x (1 - P[T])^{k-x} \, , k \geq x \, , \tag{4.73}$$

because among the k packets decoded, x have destination a. This probability is zero if $k < x$. Considering equation 4.73 and the fact that (D_k) and (A) are independent, we can obtain:

$$P[X = x | A, T_n] = \sum_{k=0}^{n} P[X = x | A, T_n, D_k] P[D_k | A, T_n] \tag{4.74}$$

$$= \sum_{k=0}^{n} P[X = x | A, T_n, D_k] r_{n,k} \tag{4.75}$$

$$= \sum_{k=x}^{n} \binom{k}{x} P[T]^x (1 - P[T])^{k-x} r_{n,k} \, . \tag{4.76}$$

Now, assuming that the considered node a does not transmit and also that (T_n) and (A) are independent:

$$
\begin{aligned}
P[X = x | A] &= \sum_{n=0}^{\infty} P[X = x | A, T_n] P[T_n | A] \qquad\qquad (4.77) \\
&= \sum_{n=0}^{\infty} P[X = x | A, T_n] \frac{(\lambda p \pi R_0^2)^n}{n!} e^{-\lambda p \pi R_0^2} \qquad (4.78) \\
&= \sum_{n=0}^{\infty} \sum_{k=x}^{n} \binom{k}{x} P[T]^x (1 - P[T])^{k-x} r_{n,k} \frac{(\lambda p \pi R_0^2)^n}{n!} e^{-\lambda p \pi R_0^2} (4.79)
\end{aligned}
$$

Here, the senders density λp and the half-duplex nature of the nodes are taken into account. So, according to equation 4.71 and for $x \neq 0$:

$$
\begin{aligned}
P[X = x] &= P[X = x | A] P[A] \qquad\qquad\qquad (4.80) \\
&= P[X = x | A](1 - p) \ .
\end{aligned}
$$

The pdf of X, the number of packets received by a becomes:

$$
P[X = x] = \sum_{n=0}^{\infty} \sum_{k=x}^{n} \binom{k}{x} P[T]^x (1 - P[T])^{k-x} r_{n,k} \frac{(\lambda p \pi R_0^2)^n}{n!} e^{-\lambda p \pi R_0^2} (1 - p) \ . \quad (4.81)
$$

The throughput is immediately obtained by taking the expectation of X:

$$
E[X] = \sum_{x=1}^{\infty} x P[X = x] \ . \qquad\qquad\qquad (4.82)
$$

If there are N nodes in the network, the local throughput, S, of the network, i.e., the throughput at the MAC layer is:

$$
S = N E[X] \ . \qquad\qquad\qquad\qquad (4.83)
$$

Note that the single-packet detection without capture is a special case of the above formulas. Indeed, by taking $r_{1,1} = 1$, $r_{n,0} = 1$ for $n \neq 1$, and $r_{n,k} = 0$ otherwise, we get:

$$
\begin{aligned}
E[X] &= P[X = 1] \qquad\qquad\qquad\qquad (4.84) \\
&= P[T](\lambda p \pi R_0^2) e^{-\lambda p \pi R_0^2} (1 - p) \\
&= p(1 - p)(1 - e^{-\lambda \pi R_0^2/2}) e^{-\lambda p \pi R_0^2} \ ,
\end{aligned}
$$

215

which is in accordance with the results of [153].

Expected Forward Progress

The forward progress, z, of a successful packet is the distance covered over a single-hop in the direction of the final destination (figure 4.26). It has been proven in [153]
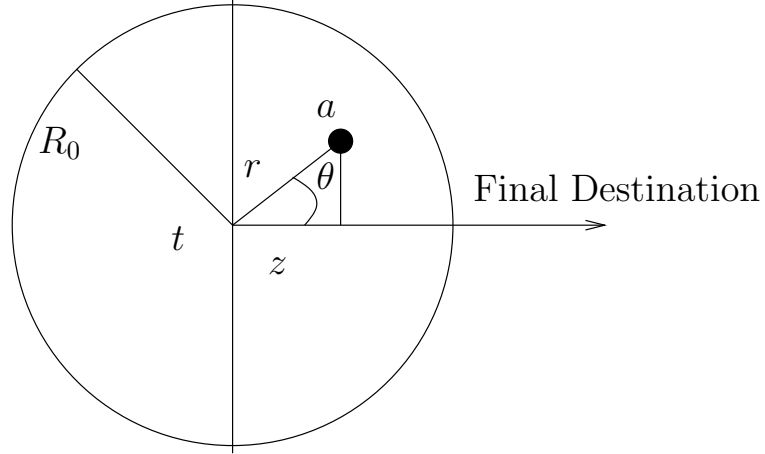
Figure 4.26: Expected forward progress.

that:

$$E[z] = \int_0^{R_0} \frac{2rd(r)}{\pi} dr \ , \tag{4.85}$$

where $d(r)$ is the pdf of the distance d between a sender and a receiver for a successful transmission (event that we denote (R)). $d(r)$ in multi-packet reception is given by:

$$P[r \leq d \leq r + dr | R, T_n, D_k] = \frac{P[R | r \leq d \leq r + dr, T_n, D_k] P[r \leq d \leq r + dr | T_n, D_k]}{P[R | T_n, D_k]} \tag{4.86}$$

We now make the realistic assumption that if k packets are decoded among n, the successful senders are the k closest senders to the receiver. Under this assumption, for $n > 1$ and $1 \leq k \leq n$:

$$P[R | T_n, D_k] = \frac{k}{n}, \tag{4.87}$$

$$P[R | r \leq d \leq r + dr, T_n, D_k] = \sum_{i=0}^{k-1} \binom{n-1}{i} \left(\frac{r^2}{R_0^2} \right)^i \left(1 - \frac{r^2}{R_0^2} \right)^{n-1-i} \tag{4.88}$$

$$P[r \leq d \leq r + dr | T_n, D_k] = P[r \leq d \leq r + dr] = \frac{2r}{R_0^2} dr \ . \tag{4.89}$$

Equation 4.87 is the proportion of successful transmissions during the considered slot. Equation 4.88 is justified by the fact that a transmission is successful at distance r from the receiver if and only if there are at most $k - 1$ senders in the disk of radius r. Moreover, the probability for a sender to be in this disk is r^2/R_0^2. We verify in appendix H that the integration of equation 4.88 over the disk of radius R_0 results in equation 4.87. The last equation is the pdf of the distance between any nodes in the disk of radius R_0^2 and the receiver. By taking into account the Poisson distribution and the $r_{n,k}$ probabilities, step by step integration of equation 4.86 provides:

$$P[r \leq d \leq r + dr | R, T_n] = \sum_{k=0}^{n} P[r \leq d \leq r + dr | R, T_n, D_k] r_{n,k} , \qquad (4.90)$$

$$
\begin{aligned}
P[r \leq d \leq r + dr | R] &= \sum_{n=1}^{\infty} \sum_{k=0}^{n} P[r \leq d \leq r + dr | R, T_n, D_k] r_{n,k} P[T_n] \qquad (4.91) \\
&= \sum_{n=1}^{\infty} \sum_{k=0}^{n} P[r \leq d \leq r + dr | R, T_n, D_k] r_{n,k} \frac{(\lambda p \pi R_0^2)^n}{n!} e^{-\lambda p \pi R_0^2}
\end{aligned}
$$

$d(r)$ becomes:

$$d(r) = \sum_{n=1}^{\infty} \sum_{k=1}^{n} \sum_{i=0}^{k-1} \binom{n-1}{i} \left(\frac{r^2}{R_0^2}\right)^i \left(1 - \frac{r^2}{R_0^2}\right)^{n-1-i} \frac{2nr}{kR_0^2} \frac{(\lambda p \pi R_0^2)^n}{n!} e^{-\lambda p \pi R_0^2} r_{n,k} . $$
$$(4.92)$$

The numerical evaluation of equation 4.85 implies the following integration:

$$
\begin{aligned}
\int_0^{R_0} r^{2(i+1)} \left(1 - \frac{r^2}{R_0^2}\right)^{n-1-i} dr &= \sum_{j=0}^{n-1-i} \binom{n-1-i}{j} \frac{(-1)^j}{R_0^{2j}} \int_0^{R_0} r^{2(i+j+1)} dr \\
&= \sum_{j=0}^{n-1-i} \binom{n-1-i}{j} \frac{(-1)^j R_0^{2i+3}}{1 + 2(i+j+1)} . \qquad (4.93)
\end{aligned}
$$

As a consequence:

$$E[z] = \sum_{n=1}^{\infty} \sum_{k=1}^{n} \sum_{i=0}^{k-1} \sum_{j=0}^{n-1-i} \binom{n-1}{i} \binom{n-1-i}{j} \frac{4(-1)^j R_0 r_{n,k} e^{-\lambda p \pi R_0^2} (\lambda p \pi R_0^2)^n}{k\pi(3 + 2i + 2j)(n-1)!}$$
$$(4.94)$$

Further simplification with the gamma function Γ gives:

$$E[z] = \frac{4R_0 e^{-\lambda p \pi R_0^2}}{\pi} \sum_{n=1}^{\infty} \sum_{k=1}^{n} \frac{\Gamma(k + \frac{3}{2})(\lambda p \pi R_0^2)^n r_{n,k}}{3\Gamma(n + \frac{3}{2})k!} \ . \qquad (4.95)$$

End-to-end Throughput

According to [153], for any randomly selected terminal, the expected path length between it and another selected terminal is given as $D = (128/45\pi)\sqrt{N/\lambda\pi}$, where N is the number of nodes in the network. Thus, the mean number of hops for a packet is $D/E[z]$ and the end-to-end throughput of the network per slot is:

$$t = \frac{SE[z]}{D} \ . \qquad (4.96)$$

4.5.2 Models for Multi-packet Reception

Now, we consider the spread slotted ALOHA protocol. At a given time-slot, all senders are supposed to choose at random a pseudo-noise (PN) code among a large book of low cross-correlated PN codes with spreading factor L, large. All potential receivers, i.e., all nodes have the knowledge of this book and are able to perform multi-packet reception. We neglect the probability that two neighboring senders choose the same code in order to simplify the calculations. From the presented models, we derive values for the $r_{n,k}$.

Simple Model

The first model is a very simple one, often used in the literature, e.g., in [144]. It states that all of the simultaneous transmissions can be successfully received if no more than K users are transmitting at the same time. If there are more than K users transmitting at the same time, the multi-user receiver is overwhelmed and a collision occurs. Thus:

$$r_{n,k} = \begin{cases} 1, & \text{if } k = n \text{ and } n \leq K \\ 1, & \text{if } k = 0 \text{ and } n > K \\ 0, & \text{otherwise} \end{cases} \qquad (4.97)$$

In the following two models, a packet is assumed to be decoded by an idle node if its SINR reaches a SINR target at the output of the detector.

4.5.3   Receiver with a Bank of Matched Filters

In this model, we suppose that radio receivers are made of MF banks of MF capable of decoding each spreading code individually. If P_0 is the transmit power, the received power at a distance r is assumed to be $P(r) = P_0/r^\alpha$, where $\alpha > 2$ is the path loss exponent. This expression is a far-field approximation that doesn't hold for small values of r. A packet is considered to be decoded if the SINR, γ, of a signal at the output of the MF reaches a SINR target γ_0, i.e., if:

$$\gamma = \frac{P}{\sigma^2 + \frac{1}{L}\sum_{i=0}^{n-1}\frac{P_0}{r_i^\alpha}} \geq \gamma_0 \; , \tag{4.98}$$

where σ^2 is the noise power, n is the number of interferers, and L is the spreading factor. This is the physical model.

In order to analytically evaluate the $r_{n,k}$ parameters, the cdf of the SINR is needed in the case of a Poisson field of interferers. This problem has been considered in [181] and in [179], where the characteristic function of the interference $Y = \sum_{i=0}^{n-1} P_0/r_i^\alpha$ has been obtained:

$$\phi_Y(\omega) = exp\left(-\pi\lambda p\Gamma(1 - 2/\alpha)e^{-i\pi/\alpha}\omega^{2/\alpha}\right), \; \omega \geq 0 \text{ and } \alpha > 2 \; , \tag{4.99}$$

where Γ is the gamma function and p is the probability of transmission. This expression leads to the exact cdf of γ and thus to the $r_{n,k}$ in the MF case.

Receiver with MMSE Multi-user Detector

Hereafter, we assume that receivers are able to perform multi-user detection thanks to a MMSE detector. While the traditional MF or Rake receiver treats interference from other users as noise, the MUD scheme jointly decodes all the signals.

The condition of decoding of a packet is still based on the SINR at the output of the signal detector. Successful decoding of packet with target SINR γ_0 can be determined on the basis of the following condition [189]:

$$\frac{P}{\sigma^2 + \frac{1}{L}\sum_{i=0}^{n-1} I(P_i, P, \gamma_0)} \geq \gamma_0 \; , \tag{4.100}$$

where $P = P_0/r^\alpha$ is the received power of the given sender, P_i is the received power from the interferer i and $I(P_i, P, \gamma_0)$ is the effective interference of sender i on reception for the target SINR γ_0:

$$I(P_i, P, \gamma_0) = \frac{PP_i}{P + P_i\gamma_0} \, .$$
(4.101)

Equation 4.100, also used in [172] in the context of call admission control, is an approximation since it is true for large systems, when $L \to \infty$, $n \to \infty$ and $L/n = \alpha$, and for random spreading sequences.

We show in appendix H that the characteristic function of the interference for a given sender and a given SINR target, γ_0 is:

$$\phi_Y(\omega) = exp\left(i\lambda p\pi\omega \int_0^{P/\gamma_0} \left(\frac{P_0}{t} - \frac{P_0\gamma_0}{P} \right)^{2/\alpha} e^{i\omega t}dt \right) \, .$$
(4.102)

While equation 4.99 is seen as the characteristic function of a stable law, equation 4.102 seems to be un-tractable for further computations.

That is the reason why we evaluate the $r_{n,k}$ probabilities thanks to Monte Carlo simulations for both MF and MMSE detectors to allow a fair comparison between schemes.

A Poisson field of interferers with density λp is generated on a two dimensional squared network $[-Xmax; Xmax] \times [-Ymax; Ymax]$. The considered receiver, a, is placed at $(0; 0)$. R_0 is fixed as the maximum distance from which can come packets for the receiver. In the absence of interferer, R_0 verifies the following expression: $\gamma_0 = P_0/(R_0^\alpha\sigma^2)$. n is the number of senders inside the disk of radius R_0 with center a. For each of these senders, the SINR is computed after summing the interference from the whole network. If the decoding condition is met (equation 4.98 for the MF case and equation 4.100 for the MMSE), the packet from this sender is assumed to be decoded. A snapshot of the simulation is shown on figure 4.27. Table 4.2 shows the parameter values used for our simulations.

Figure 4.28 shows the plot of the matrix $r_{n,k}$ in the MF case for $n \leq 14$ and $p = 0.2$. The mean number of senders in the disk of radius R_0 is $\lambda p\pi R_0^2 \simeq 5$, so the probability that $n > 14$ is very low. This figure shows that for small values of n, all packets are decoded. Then, when n increases, the number of decoded packets decreases.

Figure 4.29 shows the graph of the matrix $r_{n,k}$ for the MMSE case for $n \leq 14$.
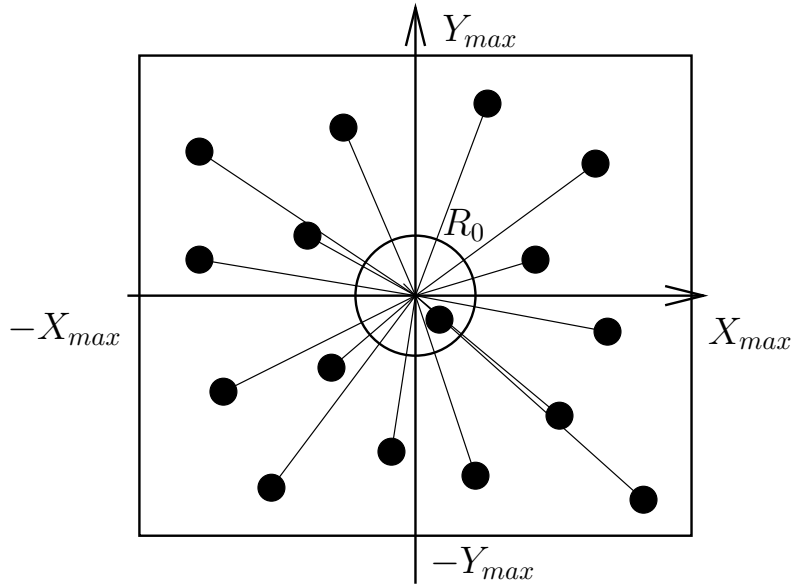
Figure 4.27: Snapshot of the Monte Carlo simulation: The power of all the interferers are summed at the receiver.

Table 4.2: Parameter values used for the Monte Carlo simulation in the case of MF receivers.

Parameter	Value
$Xmax$	50
$Ymax$	50
λ	0.25
p	0.2
L	32
P_0	5
γ_0	0.025
σ^2	0.2
α	4
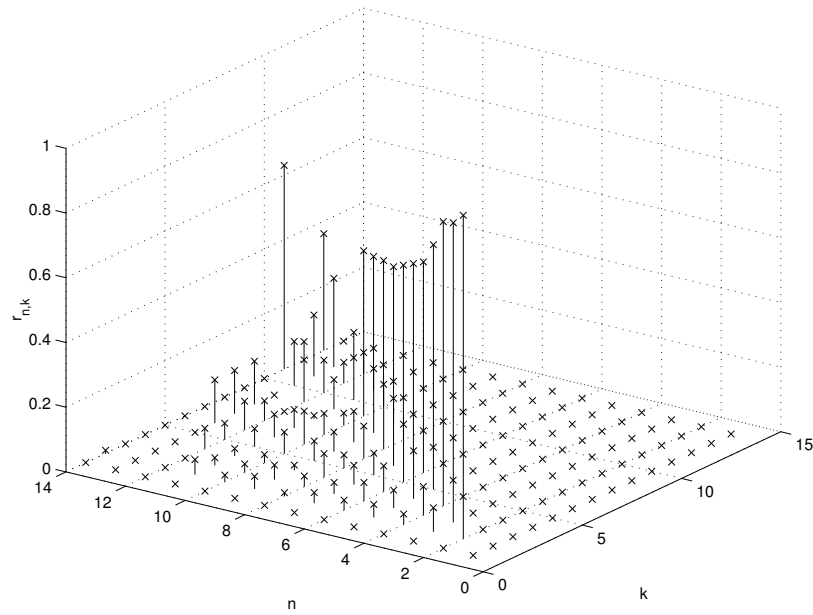
Figure 4.28: Probabilities, $r_{n,k}$, for a receiver to decode k packets given that n have been sent, MF bank detection.

It is clear that the MUD scheme offers much better performances than the MF decoding. This can be seen on the diagonal of the matrix: The higher are the probabilities, the better is the performance.

4.5.4 Numerical Results

In this section, we provide numerical results for the three models of multi-packet reception. We focus our attention on the local throughput and on the end-to-end throughput of the network.

Figure 4.30 presents the local throughput for the first simple model with different values of K. We observe in all cases the characteristic shape of the throughput of the ALOHA protocol as a function of the input load. As expected, the multi-packet reception feature improves the maximum achievable throughput.

Figure 4.31 shows the end-to-end throughput for the first simple model with different values of K. Here also, we see the advantage of multi-packet reception. Note that the optimum probability of transmission depends on K. For $K = 1$, we observe the classical result that p is optimum for $p = 1/(\lambda \pi R_0^2)$, which here is approximately 0.05. As K increases, p also increases because more packets can be handled by the receiver.

222
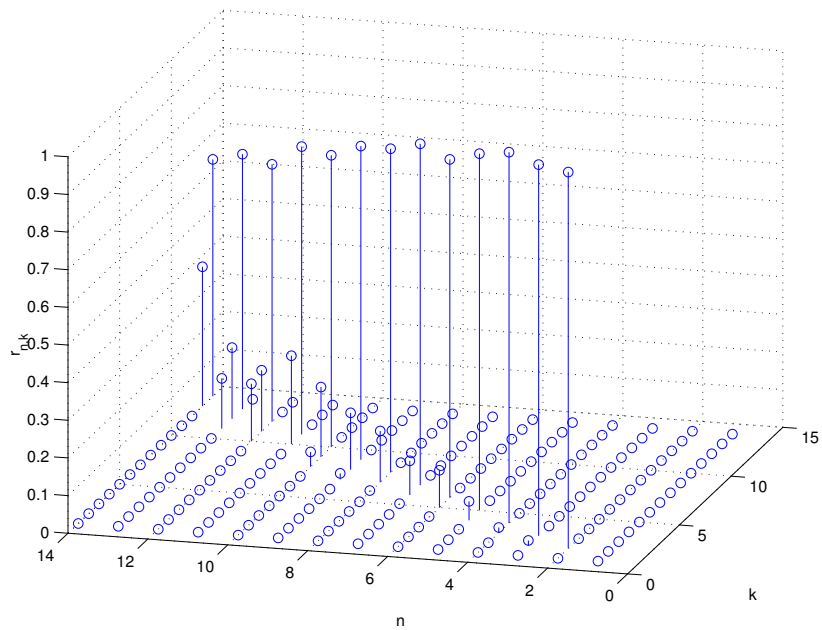
Figure 4.29: Probabilities, $r_{n,k}$, for a receiver to decode k packets given that n have been sent, MMSE multi-user detection scheme.

Local throughputs with MF and MMSE receivers are compared in figure 4.32. We observe the great advantage of the MUD over the conventional receiver (approximately 30% in our scenario). This advantage can also be seen in figure 4.33, that shows the end-to-end throughput. Indeed, the joint detection of all users makes the MUD very robust to near-far problems. This near-far resistance is of great interest in ad hoc networks because power control schemes are difficult to implement in such decentralized networks.
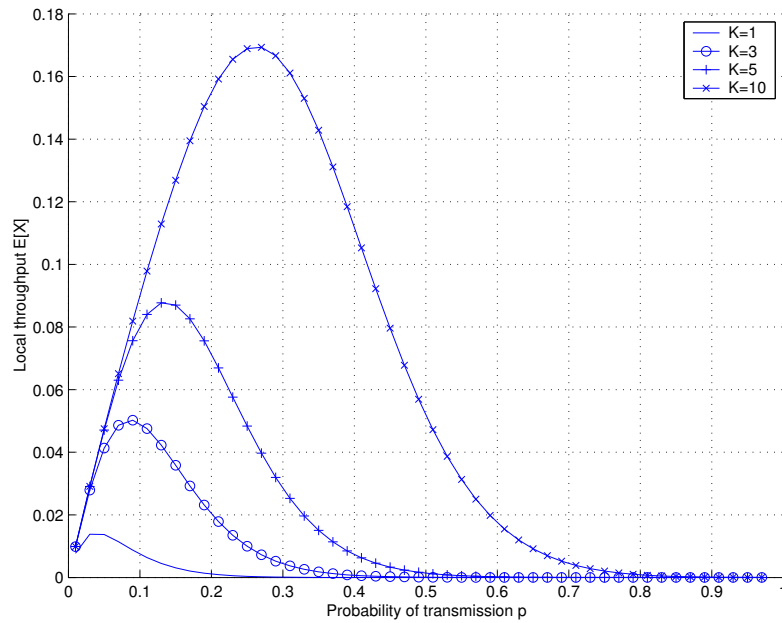
Figure 4.30: Local throughput in packets/time-slot for the simple model of multi-packet reception for different values of K.
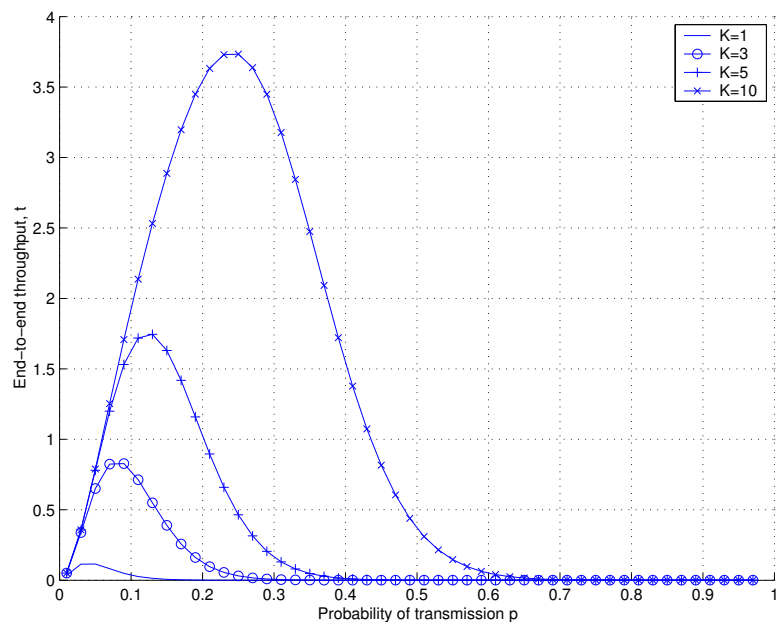
Figure 4.31: End-to-end throughput in packets/time-slot for the simple model of multi-packet reception for different values of K, and $N = 100$ nodes.
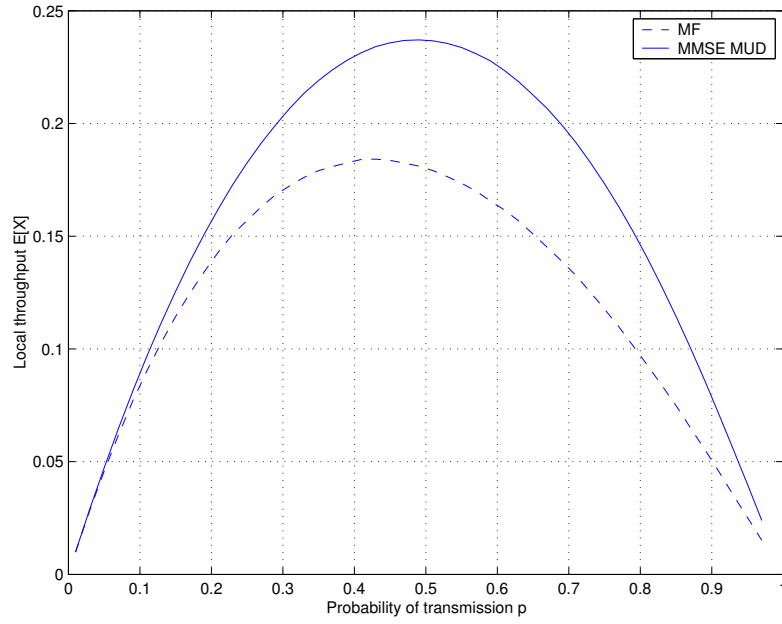
Figure 4.32: Local throughput in packets/time-slot for the MF receiver and the MMSE receiver.
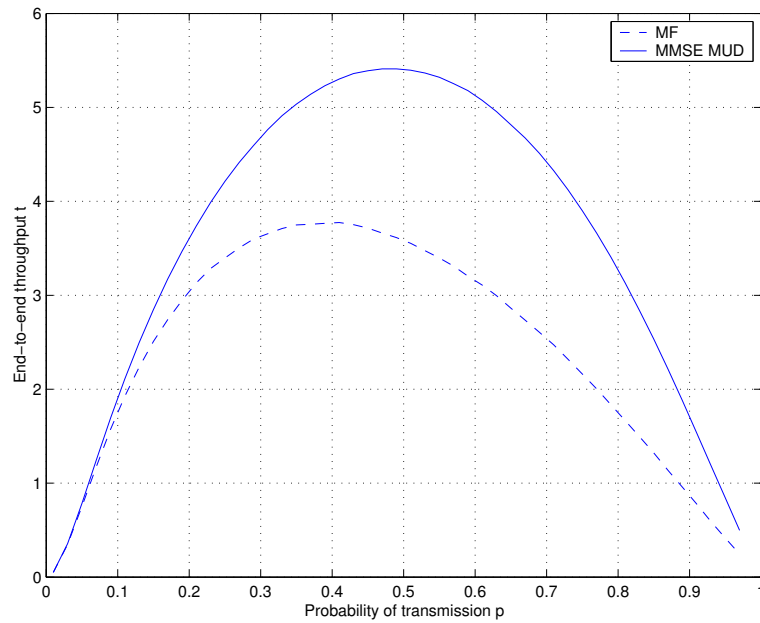
Figure 4.33: End-to-end throughput in packets/time-slot for the MF receiver and the MMSE receiver for $N = 100$ nodes.

Glossary for section 4.5

α	Path loss exponent.
Γ	Gamma function.
λ	Node density.
σ^2	Noise power.
A	Event that a node does not transmit.
D_k	Given a node a, event that a decodes exactly k packets in a given slot.
$d(r)$	Pdf of the distance d between a sender and a receiver for a successful transmission.
I	Effective interference.
K	Maximum number of simultaneously decoded packets in the simple model.
L	Spreading factor.
P	Received power.
p	Probability of transmission.
P_0	Transmit power.
R	Event of a successful transmission.
R_0	Reception radius of a receiver.
$r_{n,k}$	Probability for a receiver to decode k packets given that n have been sent in its neighborhood.
S	Local throughput.
T	Given a node a, event that a particular sender t sends a packet to a.
t	End-to-end throughput.
T_n	Given a node a, event that there are n senders in the neighborhood of a.
X	Given a node a, random variable of the number of correctly decoded packets with destination a in a given slot.
Y	Interference power.
z	Forward progress.

4.6 Conclusion and Further Work

Three parts of the study detailed in this chapter show that:

> Well designed scheduling policy associated to a simplified version of CROMA can help boost the throughput of ad hoc networks by taking advantage of node mobility.

> Cross-layer protocol mechanisms like multi-user diversity and multi-packet reception provide means for further improvement of CROMA reservation scheme as well as better network throughputs.

Multiple theoretical results (closed-form expressions) in this chapter are:

> the optimal transmission range for the one-hop scheduling policy

> the capacity of the channel aware slotted ALOHA over Rayleigh channels, with and without diversity techniques, and over Nakagami-m channels

> the local and end-to-end throughputs of slotted ALOHA with multi-packet reception

Some recommendation for further work are:

- study adaptations of multi-user diversity schemes to fulfill QoS requirements, especially in terms of packet delay

- study the impact of channel feed-back errors on multi-user diversity schemes

- study the influence of different SNR distributions at senders on the channel aware slotted ALOHA and the resulting fairness-throughput trade-off

- study the performance of CROMA with a reservation scheme based on multi-packet reception.

Conclusion

This thesis aimed to suggest MAC solution alternatives to IEEE 802.11 for wireless ad hoc networks suitable for better throughput and fairness performance in heavy loaded networks.

IEEE 802.11 results mainly from research on contention-based protocols where nodes compete for channel at each packet transmission. At high input loads, conflict-free protocols are more appropriate because resources can be reserved for a certain amount of data or time. Our thesis shows that only adaptive slot allocation protocols with random access reservation can adequately address varying topologies and traffic patterns.

From a detailed performance study of IEEE 802.11 in both single and multi-hop AP centric networks, it has been found that the IEEE standard has a high capacity for single-hop networks, especially for TCP based traffic. For VoIP traffic, specific parameter tuning or adaptation of the standard is required to cope with the high overhead of the MAC protocol. Multi-hop networking can be useful to extend the coverage of APs, or to mitigate the near-far effect. However, IEEE 802.11 leaves ample room for improvements in multi-hop networks, especially at high input loads.

In this context, two major accomplishments of this work are:

(i) A TDMA based receiver-oriented protocol has been proposed. CROMA achieves higher throughput and better local fairness than IEEE 802.11. This has been shown through extensive simulations considering a challenging network topology, a random network, and a mobile network. An important multi-slot extension of the protocol makes CROMA performance independent of the frame length. CROMA illustrates the advantage of using conflict-free protocols at high input loads.

(ii) The design of cross-layer mechanisms for CROMA enhancement is the other contribution from this study. The thesis presented three schemes to overcome the limitation suggested by Gupta and Kumar and demonstrated their effectiveness to

improve the reservation scheme of CROMA. First of all, simulations and theoretical results confirmed that mobility can increase the throughput of ad hoc networks thanks to a specific scheduling policy based on a simplified version of CROMA. The importance of the traffic model for capacity evaluation was duly stressed. Then, closed-form formulas have been obtained for the capacity of the channel aware slotted ALOHA. This is an interesting application of multi-user diversity that could replace the reservation scheme of CROMA. Finally, we have highlighted the benefit of using multi-packet reception in multi-hop networks. In particular, we demonstrated the advantage of multi-user detection against match filter reception. Moreover, the near-far resistance of this technique is of particular interest for ad hoc networking where power control algorithms are difficult to implement.

Multiple areas of further investigation have been suggested in the concluding paragraphs of the previous chapter.

For CROMA, further simulations on its performance evaluation in AP centric networks with TCP and UDP traffic are required. Possibilities of CROMA deployment in a real world scenario are foreseen. The additional results from these tasks would support an eventual CROMA proposal as a complete alternative to IEEE 802.11.

Besides the design of suitable link adaptation mechanisms for CROMA, cross-layer interactions seem to be the most important research area of immediate interest. Three examples in this regard are:

- ALOHA and multi-user diversity mechanisms for optimized performance considering QoS constraints, channel error feedback, multi-hop networks, and uneven SNR distributions experienced by different users should be studied.

- An exhaustive study of cross-layer interactions for CROMA and other slotted protocols like slotted ALOHA and slotted CSMA is needed.

- The closed-form formula of the end-to-end throughput of slotted ALOHA with MMSE based multi-user detection can be of interest for the sake of completeness.

Appendix A

Link Adaptation Models for IEEE 802.11b

In this appendix, we describe the model used for link adaptation in ns2 simulations of chapter 2. Approximations and assumptions for the computation of the bit error rate (BER) and the packet error rate (PER) are given in the next section. Then, PER curves are approximated with simple functions that are implemented in the simulator. Finally, other link adaptation strategies are presented.

A.1 Packet and Bit Error Rates Computation

Four physical modes are standardized by IEEE 802.11b: 1, 2, 5.5, and 11 Mbps. In our simulations, the basic set rate consists in the first two physical modes. Moreover, if compatibility is required with IEEE 802.11 and if all physical modes are allowed, control packets (RTS, CTS, and ACK) as well as broadcast packets have to be sent at 1 Mbps. In nowadays implementations, we can note by sniffing the radio channel that ACK packets are sent at the same physical mode as the preceding data packet. This is not in strict sense in accordance with the standard, it is reasonable since there is no utility to decode the ACK packet. In our simulations, ACK is sent at 1 Mbps: This is compatible with the standard and provides a worst case in terms of capacity.

The physical modes use spread spectrum in conjunction with DBPSK, DQPSK, and CCK modulations [36]. In order to simplify as much as possible the ns2 simulations, we have made two assumptions. The first one is that errors on bits are

independent, so that the PER can be deduced from the BER with the following formula: $PER = 1 - (1 - BER)^N$, where N is the number of bits in the considered packet.

The second approximation is that transmissions are done over an AWGN channel. This a very optimistic approximation since indoor channels are mostly characterized by Rayleigh channels and multi-path propagation. This approximation has been done for the sake of simplicity since BER can be analytically computed over AWGN [163].

[163] p. 270 gives the following formula for the BER of DBPSK (E_b is the bit energy and N_0 is the noise level):

$$BER_{DBPSK} = \frac{1}{2} \exp\left(-\frac{-E_b}{N_0}\right) . \tag{A.1}$$

The 11-chip direct sequence spreading increases the processing gain at the receiver by: $G_{DBPSK} = 10 \log(11) = 10.4$ dB. [163] p. 271 provides the following expression for DQPSK:

$$BER_{DQPSK} = Q(a, b) - \frac{1}{2} I_0(ab) \exp\left(-\frac{a^2 + b^2}{2}\right) , \tag{A.2}$$

where Q is the Marcum function, I_0 is the modified Bessel function of order zero, and:

$$a = \sqrt{2\frac{E_b}{N_0}\left(1 - \frac{1}{\sqrt{2}}\right)} \tag{A.3}$$

$$b = \sqrt{2\frac{E_b}{N_0}\left(1 + \frac{1}{\sqrt{2}}\right)} . \tag{A.4}$$

The processing gain is $G_{DQPSK} = 7.4$ dB.

Following [86], we now consider CCK as a variation of the M-ary Bi-Orthogonal Keying modulation (MBOK). In this case, [163] (p. 256):

$$BER_{MBOK} = \frac{2^{k-1}}{2^k - 1}\left\{1 - \frac{1}{\sqrt{2\pi}}\int_{-\sqrt{2\gamma}}^{+\infty} \exp\left(-\frac{v^2}{2}\right) dv \times \right. \tag{A.5}$$
$$\left. \left[\frac{1}{\sqrt{2\pi}}\int_{-(v+\sqrt{2\gamma})}^{v+\sqrt{2\gamma}} \exp\left(-\frac{x^2}{2}\right) dx\right]^{\frac{M}{2}-1}\right\},$$

where γ is the received signal to noise ratio (SNR) per k-bits symbol and M is the

size of the constellation. $M = 4$ for 5.5 Mbps and $M = 8$ for 11 Mbps. This formula can be numerically evaluated thanks to a mathematical software. The processing gains are the following: $G_{CCK5.5} = 3$ dB and $G_{CCK11} = 0$ dB.
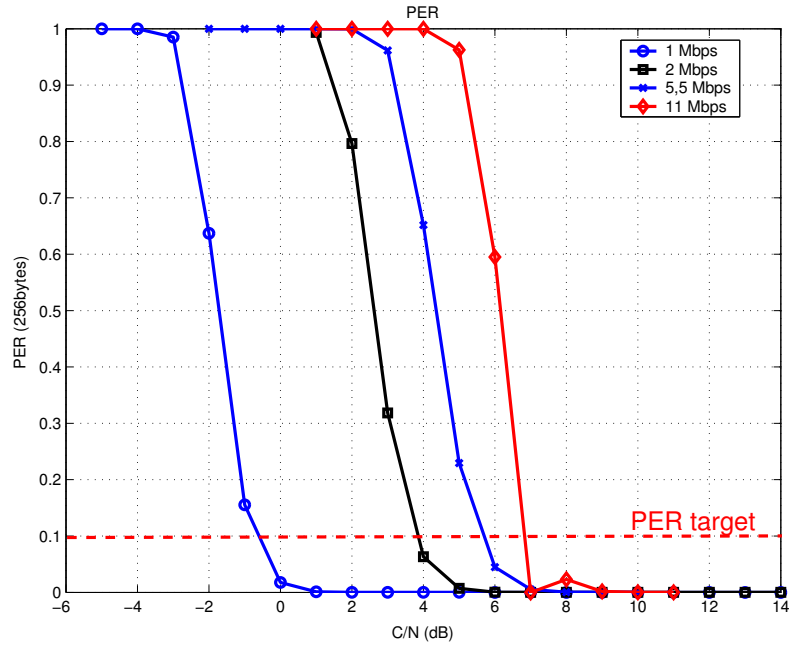
Figure A.1: PER vs. C/N for the four physical modes of IEEE 802.11b, 256 byte packets.

Figure A.1 depicts the PER as a function of the carrier to noise ratio (C/N) for 256 byte packets. We also obtain the throughput curve (figure A.2): $R(1 - PER)$, where R is the transmission data rate. This formula is true under the assumption of a perfect ARQ scheme.

A.2 Model

Since the closed-form formula described in the previous section cannot be evaluated during the simulations in a easy and fast way, a set of computed values have to be extrapolated. Following [137], the modeling of the throughput curves is processed via hyperbolic functions. By means of least square optimization, we obtain the parameters of these functions. The generic function that has to be fit with the previous results is the following:

$$f(x) = \alpha_1 Tanh[\alpha_2(x - \alpha_3)] + \alpha_4 .$$

$(A.6)$
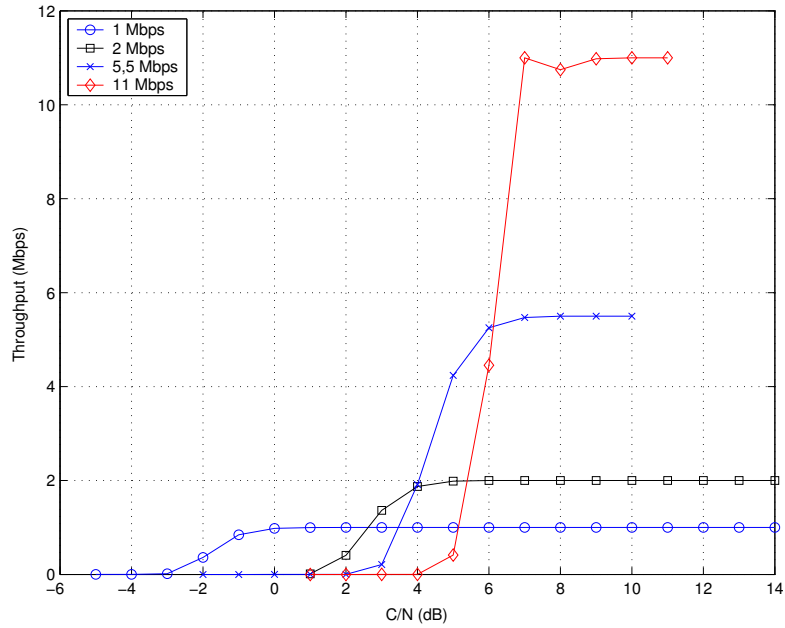
Figure A.2: Throughput vs. C/N for the four physical modes of IEEE 802.11b, 256 byte packets.

For 256 byte packets, the parameters are given in table A.1. The least square

Table A.1: Parameters of the hyperbolic functions for 256 byte packets.

Physical mode	α_1	α_2	α_3	α_4
1 Mbps	0.5058	1.1747	-1.7571	0.4938
2 Mbps	1.0345	1.0008	2.5995	0.9654
5.5 Mbps	2.7608	0.9708	4.3452	2.7304
11 Mbps	5.4465	2.5175	6.0778	5.5187

optimization provides the curves shown in figure A.3.

Note that there are two main strategies. The first one aims at maximizing the throughput. The second one, used in our simulations, aims at keeping the PER below a PER target. In our example, for a PER target of 0.1, we obtained the thresholds given in table A.2.

The choice of the physical mode is also conditioned to the sensitivity levels of the IEEE 802.11b card. The minimum sensitivity thresholds are given by the standard and are recalled in table A.3.
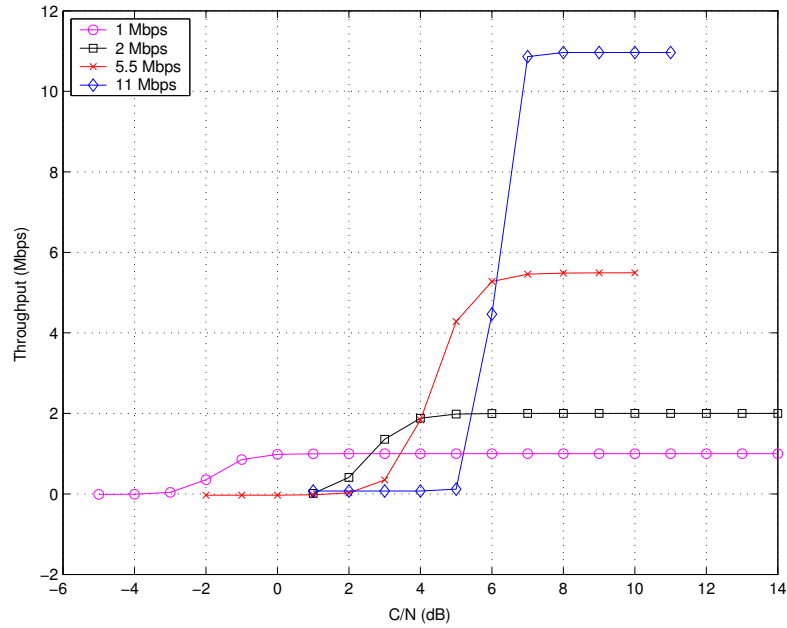
Figure A.3: Throughput model, 256 byte packets.

Table A.2: C/N thresholds for 256 byte packets.

C/N intervals	Physical modes
Up to 3.8 dB	1 Mbps
Between 3.8 and 5.7 dB	2 Mbps
Between 5.7 and 6.7 dB	5.5 Mbps
Above 6.7 dB	11 Mbps

A.3   Other Link Adaptation Algorithms

Other link adaptation mechanisms have been proposed in the literature and are summarized in [138] and [44]. We give hereafter a brief overview of some algorithms.

Auto Rate Fallback

With the auto rate fallback mechanism [126], the sender selects the best rate based on informations collected from previous data transmissions. The data rate is incrementally increased or decreased according to the number of consecutive received or missed ACK. This algorithm is very simple, requires no modification of the standard, and is a sender based approach.

Table A.3: Sensitivity thresholds.

Physical mode	Minimum sensitivity theshold
1 Mbps	-94 dBm
2 Mbps	-91 dBm
5.5 Mbps	-87 dBm
11 Mbps	-82 dBm

Receiver Based Auto Rate

Here [111], the receiver estimates the channel quality while receiving the RTS frame. Then, the chosen rate must be carried back to the sender by the CTS. A modification of the standard frame formats is needed.

Opportunistic Auto Rate

Opportunistic auto rate is another receiver based approach. Authors of [171] observed that the coherence time of the channel is typically at least multiple frames transmission times. Taking advantage of that, the algorithm allows the transmission of bursts of high data rate packets, when the channel is good. Modifications of the standard are needed.

Appendix B

Additional Results for Outdoor Coverage with IEEE 802.11 DCF

In this appendix, we investigate the influence of outdoor propagation on the MAC protocol of IEEE 802.11 DCF. In particular, we focus on the vulnerability period. We defined the vulnerability period as the time of a frame handshake, during which the transmission of a data packet can cause a collision. We study how this vulnerability period is increased in an outdoor environment, where propagation times are much longer than in the indoor case.

B.1   Exhaustive Analysis

In this section, we review all basic situations of possible collisions. Let us consider an AP and two nodes 0 and 1, in the cell of this AP. There are only two possibilities for their respective location: Either node 0 and 1 are in visibility (i.e., in transmission range from each other), or not (see upper part of figure B.1). If we now consider the status of the AP and of the nodes, there are also only two cases: either they are transmitting (T) or receiving (R). This leads to the few basic situations described in table B.1 (nodes 0 and 1 not in visibility) and table B.2 (nodes 0 and 1 in visibility). All sub-cases are depicted in figure B.1.

After an exhaustive analysis of each, the eleven sub-cases can be reduced to two basic ones, shown in figure B.2. All other sub-cases are superposition of these two last ones. We recognize the cases described in the literature as collision of the first order (two packets interfere at the receiver) and of the second order (a node cannot

Table B.1: Sub-cases when nodes 0 and 1 are not in visibility.

Node 0	AP	Node 1	Sub-case
T	R	R	No collision
T	R	T	Sub-case 1
T	T	R	Sub-case 3
T	T	T	Sub-case 2
R	R	R	No collision
R	R	T	No collision
R	T	R	No collision
R	T	T	Sub-cases 3 and 4

Table B.2: Sub-cases when nodes 0 and 1 are in visibility.

Node 0	Node 1	AP	Sub-case
T	R	R	No collision
T	R	T	Sub-cases 5 and 6
T	T	R	Sub-case 7
T	T	T	Sub-cases 8 and 9
R	R	R	No collision
R	R	T	No collision
R	T	R	No collision
R	T	T	Sub-cases 10 and 11
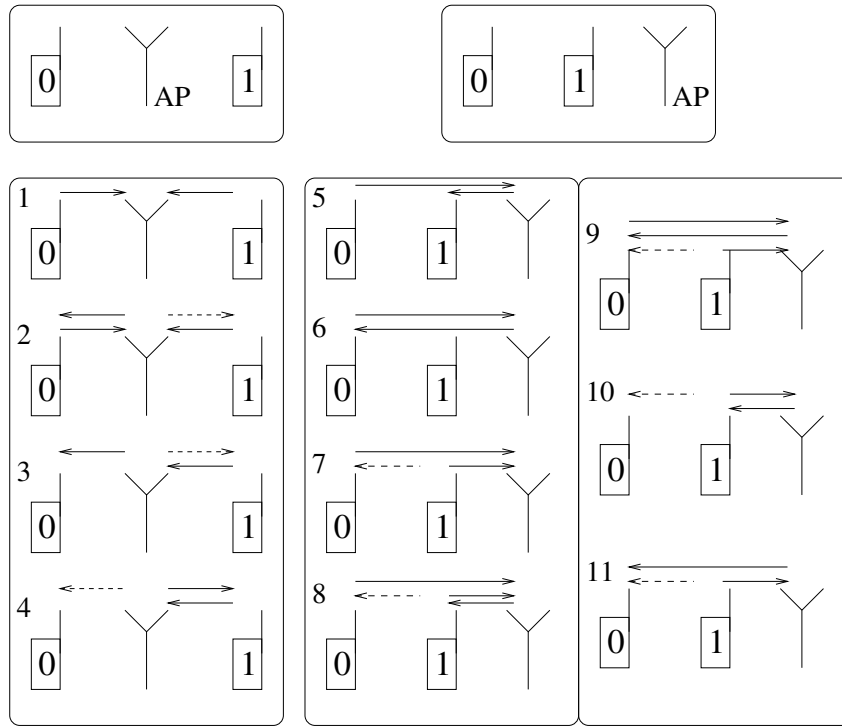
Figure B.1: Exhaustive list of sub-cases in case of packet collision (solid arrows represent packet transmissions).

Figure B.2: Two basic situations for the study of the vulnerability period.

receive a packet because it is transmitting). Let us now study in detail these two collision types.

B.2 Collision of the First Order

Figure B.3 shows the collision of the first order. P_0 and P_1 are the propagation times between the AP and respectively nodes 0 and 1. t_1 (respectively t_0) is the time at which node 1 (respectively node 0) sends an RTS. We look for the interval for t_0, during which a collision is possible, i.e., the vulnerability period.
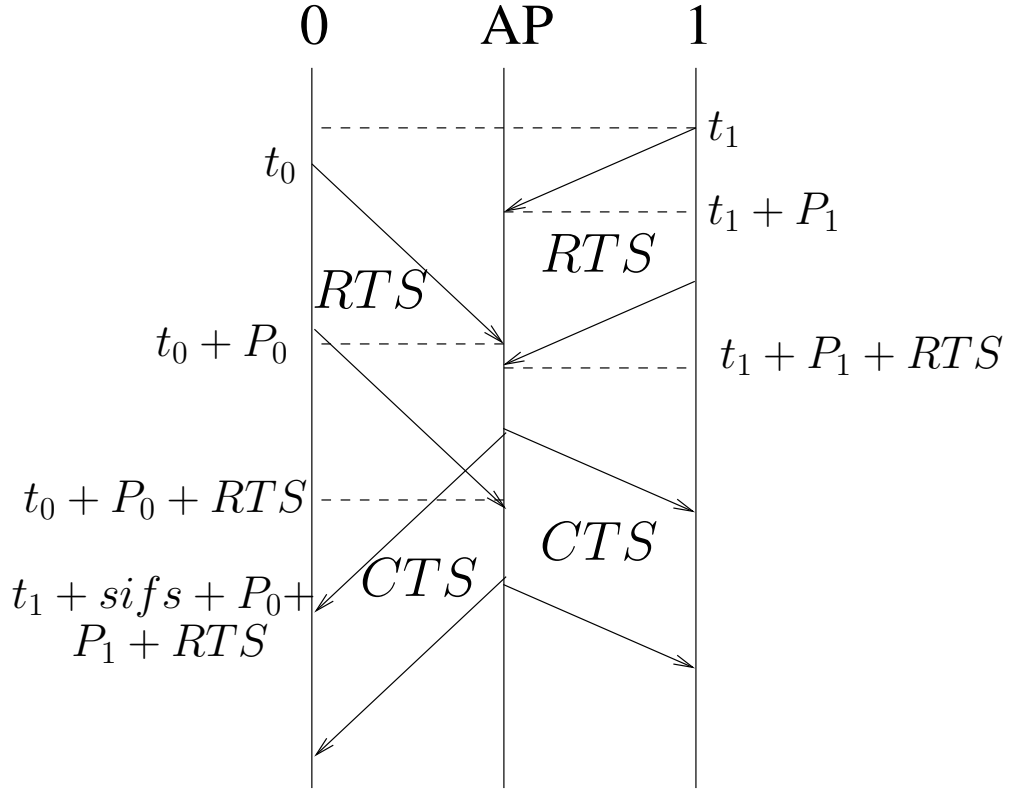
Figure B.3: Collision of the first order.

- t_0 cannot be greater than $t_1 + sifs + P_0 + P_1 + RTS$ since after this time the physical carrier sensing avoids the transmission of an RTS.

- For $t_1 + P_1 - RTS - P_0 \leq t_1 \leq t_1 + sifs + P_0 + P_1 + RTS$, the transmission of an RTS by node 0 either implies a collision at the AP, or comes during the the RTS / CTS handshake, which is an anomaly.

- For $t_0 \leq t_1 + P_1 - RTS - P_0$, the AP starts the handshake with node 0 and we have exactly the symmetric situation of the previous one. Thus, t_1 cannot be greater than $t_0 + sifs + P_0 + P_1 + RTS$, which is equivalent to: t_0 cannot be smaller than $t_1 - RTS - P_0 - P_1 - sifs$.

As a conclusion, $t_1 - RTS - P_0 - P_1 - sifs \leq t_0 \leq t_1 + sifs + P_0 + P_1 + RTS$. And the length of the vulnerability period is:

$$L = 2(RTS + P_0 + P_1 + sifs) \,. \tag{B.1}$$

It is now clear that a longer propagation delay implies a longer vulnerability period and hence, more collisions. However, the propagation term in the previous formula

is small (26 μs in the worst case in rural areas, see section 2.4) compared to the duration of a RTS (176 μs at 2 Mbps with short preamble).

B.3   Collision of the Second Order
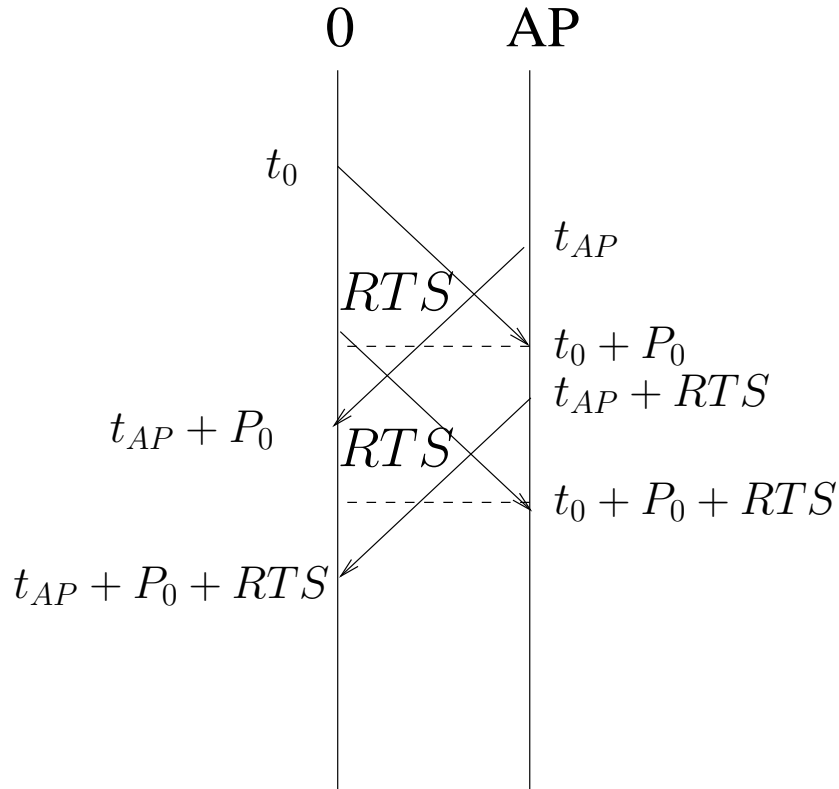
Figure B.4: Collision of the second order.

Figure B.4 shows a collision of the second order. t_0 and t_{AP} are the time instants at which respectively node 0 and the AP send an RTS. P_0 is the propagation time between the two entities. We look for the time interval for t_0, during which a collision is possible. The scenario is symmetrical.

- For a collision, t_0 must be such that $t_0 + P + 0 > t_{AP}$. Otherwise, the AP carrier sensing mechanism detects a signal and the AP does not transmit.

- For a collision, t_0 must be such that $t_0 < t_{AP} + P_0$. Otherwise, node 0 detects the RTS of the AP and doesn't transmit.

As a conclusion, there is a collision for $t_{AP} - P_0 < t_0 < t_{AP} + P_0$. And the vulnera-

bility period is:

$$L = 2P_0 \ . \tag{B.2}$$

Here also, the vulnerability period depends on the propagation time. With an increase of this time, more collisions are expected. In an indoor environment, L is no more than 2 μs, while in outdoor L can reach 26 μs.

B.4 Influence on the NAV

We now investigate the influence of the propagation delay on the NAV computation. A data transfer takes place between node 0 and the AP. Of course, all nodes in the cell are in visibility of the AP. The influence of the propagation delay is studied for the computation of the NAV of node 1. Three cases have to be considered: node 1 is hidden from the sender; node 1 is hidden from the receiver; node 1 can hear all packets of the handshake. We will see that this latter case is a sub-case of the first one. It is assumed that the NAV is updated each time a packet is received.

B.4.1 The Node is Hidden from the Sender

In figure B.5, the handshake between node 0 and the AP is shown. Node 1 is hidden from node 0. It receives the CTS at $t_0 + RTS + CTS + P_0 + P_1 + sifs$. It sets its NAV to $t_1 = t_0 + RTS + CTS + P_0 + P_1 + sifs + (DATA + ACK + 2sifs)$, i.e., $t_1 = t_0 + RTS + CTS + P_0 + P_1 + 3sifs + DATA + ACK$. On the other hand, node 1 receives the first bit of the ACK at $t_2 = t_0 + RTS + CTS + DATA + 3P_0 + P_1 + 3sifs$. At this time, the carrier sensing prevents node 1 from sending any data. Hence, there is a problem if $t_1 < t_2$, i.e., if:

$$ACK < 2P_0 \ . \tag{B.3}$$

Remember that at 2 Mbps with a short preamble, $ACK = 152$ μs. As a conclusion, this case is not a problem from the NAV perspective for the rural application. It is also clear that if node 1 hears the data packet, the situation is even better. Thus, there is still a single case to study.

B.4.2 The Node is Hidden from the Receiver

In figure B.6, a typical message exchange is shown with node 1 hidden from the receiver, node 0. Node 1 receives the data packet at $RTS + CTS + DATA + 2sifs + 2P_0 + P_1$. It sets its NAV to $t_1 = t_{AP} + RTS + CTS + DATA + 2sifs + 2P_0 + P_1 + (ACK + sifs)$, i.e., $t_1 = t_{AP} + RTS + CTS + DATA + ACK + 3sifs + 2P_0 + P_1$. On the other hand, the AP receives the last bit of the ACK at $t_2 = t_{AP} + RTS + CTS + DATA + ACK + 4P_0 + 3sifs$. A problem may occur if node 1 does a wrong evaluation of the end time of the handshake between the AP and node 0, i.e., if $t_1 < t_2$, which is equivalent to:

$$P_1 < 3P_0 . \tag{B.4}$$

Thus, a problem may occur if node 1 is close to the AP and node 0 is far. However, this situation should not lead to collisions because any transmission is preceded by at least a DIFS period. Thus, in the worst case, the first bit of a packet sent by node 1 arrives to the AP at $t_1 + difs + P_1$. So, there is a collision if:

$$difs + 2P_1 < 3P_0 , \tag{B.5}$$

which is not supposed to occur in our model ($P_0 < 13$ μs).

As a conclusion, the utilization of IEEE 802.11b DCF in an outdoor environment has a small impact on the number of collisions.
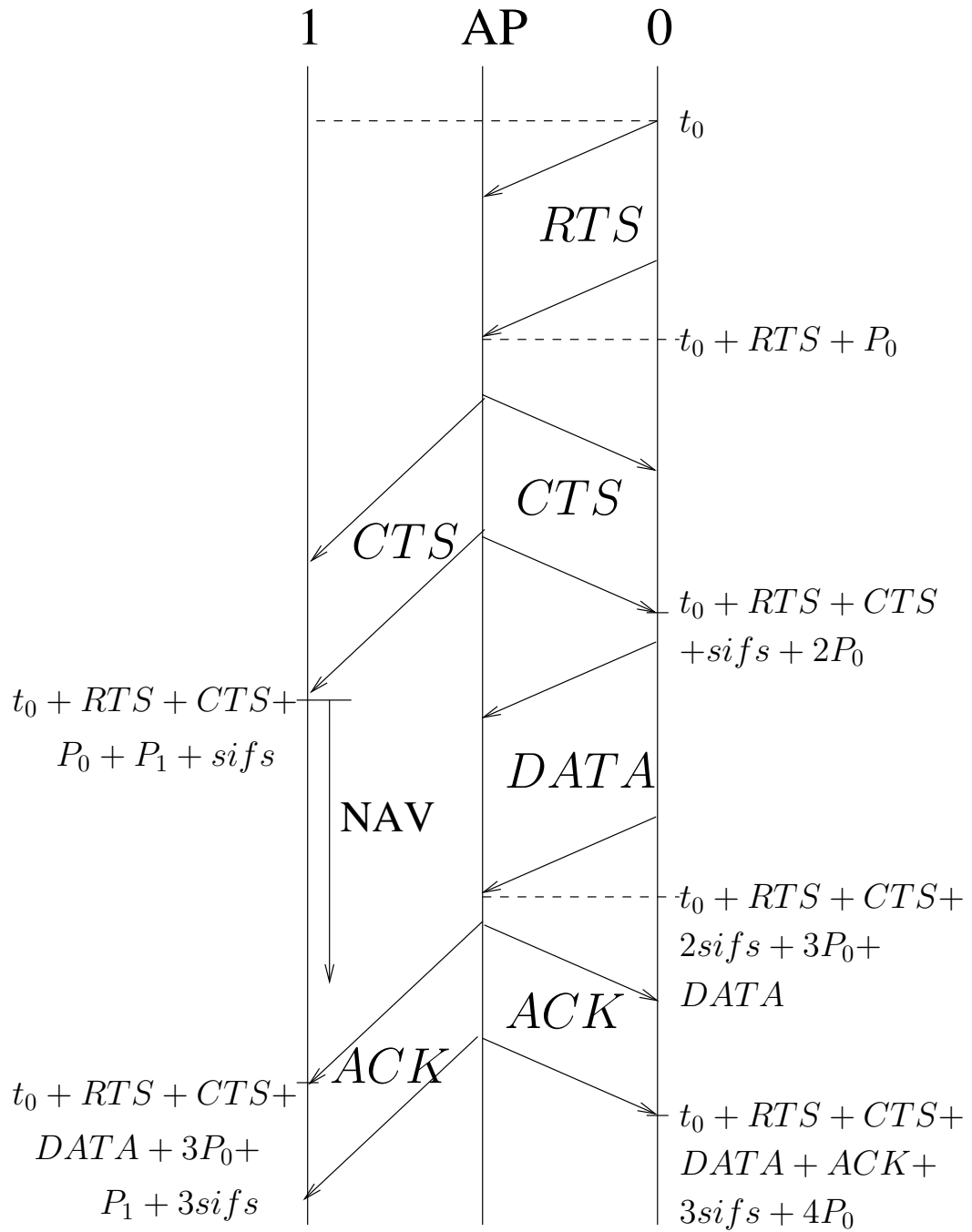
Figure B.5: NAV computation: node 1 is hidden from the sender.
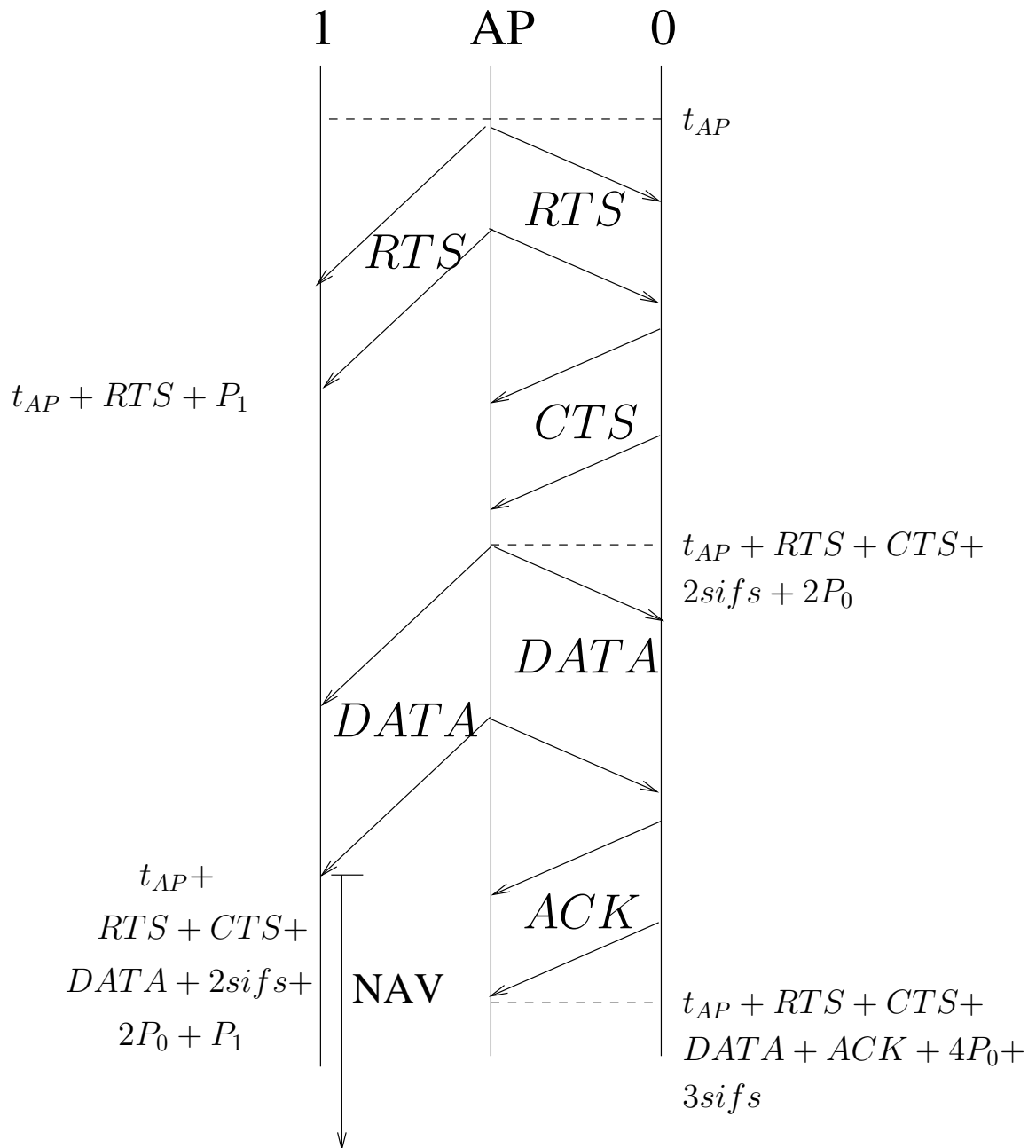
Figure B.6: NAV computation: Node 1 is hidden from the receiver.

Appendix C

Maximum Achievable Throughput of IEEE 802.11 PCF

In this appendix, we provide the simple analytical model that allows the computation of the maximum achievable throughput of the IEEE 802.11 PCF. Two transport protocols have been considered: TCP and UDP.

C.1   PCF Mode

Both simulations with the Network Simulator v.2 (ns2) and the analytical computation are based on a simplified version of the contention-free period (CFP) of PCF. This version is described in this section.

C.1.1   Polling Process

The CFP begins with a Beacon and ends with the control packet CF-End. Three cases have been implemented for the polling (see figure C.1). In the first one, the access point (AP) sends a specific CF-Poll packet. In the second case, the poll packet is concatenated with a downlink packet and the mobile station acknowledges the data packet with an ACK packet. In the third case, the mobile station acknowledges the downlink packet while sending at the same time an uplink packet to the AP.
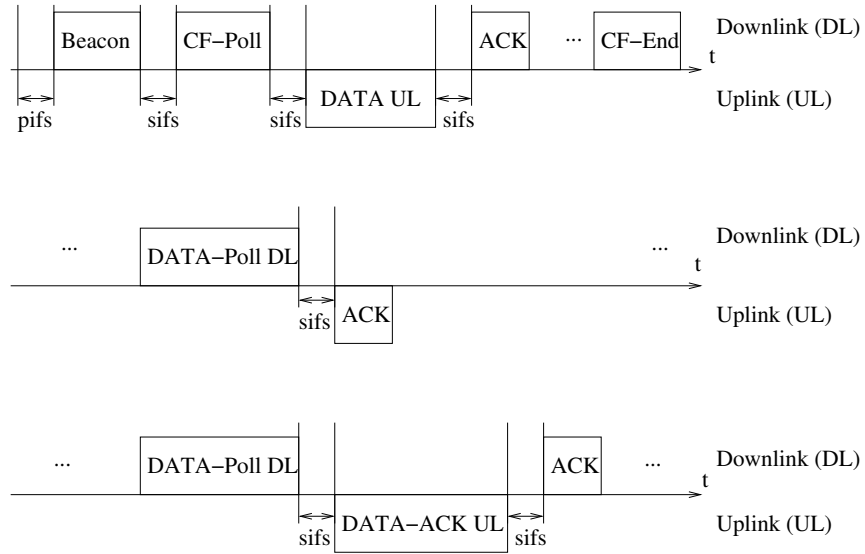
Figure C.1: Simplified version of the PCF implemented in ns2.

C.1.2 Multi-rate Support

Beacon and CF-End packet are sent at the basic rate. Data packets, CF-Poll, DATA+Poll, and DATA+ACK are sent according to the link adaptation mechanism (see appendix A). ACK packets are sent using the physical mode used for the transmission of the preceding data packet.

C.1.3 Beacon Interval and Duration of the CFP

According to the standard, the parameter $CFPMaxDuration$ gives the maximum duration of the CFP and is transported by the beacons. This value is used by stations to set their Network Allocation Vector (NAV) at the nominal start of the CFP ($TBTT$). So, a CFP ends no later than $TBTT + CFPMaxDuration$. An AP can poll a station if and only if the expected response arrives no later than $TBTT + CFPMaxDuration$.

The minimum value for $CFPMaxDuration$ is $2MaxMPDUTime + Beacon + CF - End$, with $MaxMPDU = 2346$ bytes. The maximum value is $CFPPeriod - [MaxMPDUTime + 2sifs + 2SlotTime + 8ACK]$. In our implementation, we chose $CFPPeriod = 100$ ms $= 98$ TU (a Time Unit is 1024 μs). For IEEE 802.11b with

246

a basic rate of 1 Mbps, we have:

$$ACK = 304\,\mu s\,,\qquad(C.1)$$

$$MaxMPDUTime = 18768\,\mu s\,,\qquad(C.2)$$

$$CFPMaxDuration = 77\ \text{TU}\,.\qquad(C.3)$$

If we assume that the CFP period lasts $CFPMaxDuration$, the contention period (CP) lasts: $CPPeriod = CFPPeriod - CFPMaxDuration = 21$ TU. The maximum time during which the CFP is dedicated to data transfer is:

$$\begin{aligned} MaxDataCFPTimeTransfer &= CFPPeriod - Beacon - pifs - \qquad(C.4)\\ &\quad CFEnd - sifs - CPPeriod\,. \end{aligned}$$

C.2 UDP Traffic

The main parameters used in this section are shown on figure C.2.
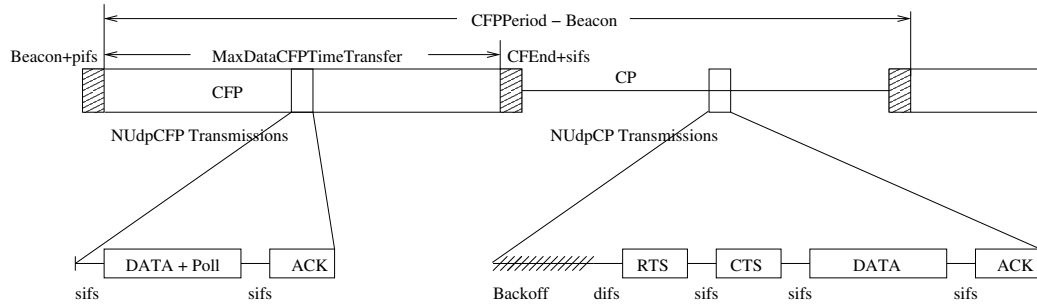
Figure C.2: Main parameters for the computation of the maximum achievable throughput in PCF.

C.2.1 Contention-free Period

During the CFP, UDP packets are sent in downlink according to the scheme: sifs / DATA + CF-Poll / sifs / ACK. Thus, the transmission time during the CFP is:

$$UdpCFPTransDur = sifs + MACPDU + sifs + ACK\,,\qquad(C.5)$$

where $MACPDU$ is the time needed to transfer a MAC Packet Data Unit. And the number of UDP packet sent during the CFP is:

$$NUdpCFP = \lfloor MaxDataCFPTimeTransfer/UdpCFPTransDur \rfloor . \quad \text{(C.6)}$$

The additional overhead due to the beacon and CF-End transfers is the following:

$$OvCFP = Beacon + pifs + CFEnd + sifs . \quad \text{(C.7)}$$

So, the equivalent transmission time of an UDP packet in downlink during the CFP is:

$$EqUdpCFPTransDur = UdpCFPTransDur + OvCFP/NUdpCFP . \quad \text{(C.8)}$$

C.2.2 Contention Period

During the CP, UDP packet are sent according to the following scheme: Backoff / difs / RTS / sifs / CTS / sifs / DATA / sifs / ACK. Thus, the transmission time is:

$$
\begin{aligned}
UdpCPTransDur \quad = \quad & CWmin/2 + difs + RTS + CTS + \quad \text{(C.9)} \\
& 3sifs + MACPDU + ACK .
\end{aligned}
$$

Now, the length of the CP is:

$$
\begin{aligned}
CPLength \quad = \quad & CFPPeriod - Beacon - pifs - CFEnd - sifs - \quad \text{(C.10)} \\
& NUdpCFP \times UdpCFPTransDur .
\end{aligned}
$$

And the number of packets sent during the CP:

$$NUdpCP = \lceil CPLength/UdpCPTransDur \rceil . \quad \text{(C.11)}$$

We now compute the equivalent transmission time of an UDP packet, taken into account the CFP:

$$
\begin{aligned}
EqUdpTransDur \quad = \quad & (EqUdpCFPTransDur \times NUdpCFP + \quad \text{(C.12)} \\
& UdpCPTransDur \times NUdpCP)/(NUdpCFP + NUdpCP) .
\end{aligned}
$$

As a consequence, the UDP throughput is given by:

$$ThUDP = UDPPacketSize/EqUdpTransDur \ . \qquad (C.13)$$

Analytical and simulation results are compared in table 2.5 of chapter 2.

C.3   TCP Traffic

C.3.1   Contention-free Period

During the CFP, TCP packets are sent in downlink according to the following scheme: sifs / TCP / sifs / ACKTCP + ACK / sifs / ACK. So, the transmission time of the CFP is:

$$
\begin{aligned}
TcpCFPTransDur \ &= \ sifs + MacTCPPdu + sifs + \qquad (C.14) \\
&\quad MacACKTCPPdu + sifs + ACK \ .
\end{aligned}
$$

Hence, the approximate number of TCP packets sent during the CFP is:

$$
\begin{aligned}
NTcpCFP \ &= \ \lfloor (CFPPeriod - CPPeriod - Beacon - \qquad (C.15) \\
&\quad pifs - CFEnd - sifs)/TcpCFPTransDur \rfloor \ .
\end{aligned}
$$

The overhead per data packet can also be computed:

$$OvCFP = \lceil (Beacon + pifs + CFEnd + sifs)/NTcpCFP \rceil \ . \qquad (C.16)$$

C.3.2   Contention Period

During this period, the TCP packets are sent according to this scheme: difs / Back-off / RTS / sifs / CTS / sifs / MacTCPPdu / sifs / ACK, while the ACKTCP packet is sent according to: difs / Back-off / ACKTCP / sifs / ACK. From these schemes, the value of $TcpCPTransDur$ can be derived (the computation is similar to the UDP case). Then, the approximate number of TCP packets sent during the

CP is:

$$NTcpCP \;=\; \lceil(CFPPeriod - Beacon - pifs - CFEnd - sifs - \quad \text{(C.17)}$$
$$NTcp \times TcpCFPTransDur)/TcpCPTransDur\rceil \,.$$

The transmission time of a TCP packet is then averaged over the CFP and the CP. Analytical and simulation results are compared in table 2.2 of chapter 2.

Appendix D

CROMA Correctness
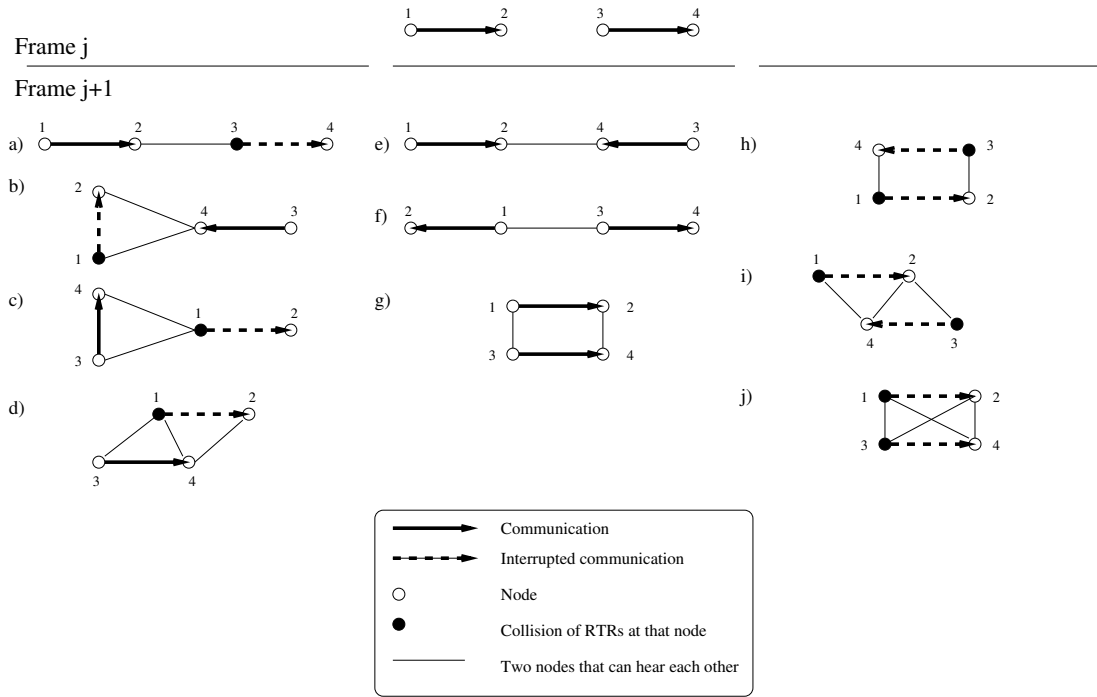
Figure D.1: Interference between two communications sharing the same slot.

In this section, we will show that CROMA is correct, i.e., that it is collision-free in both fixed and mobile environments. The capture effect is not considered here, so this section shows that CROMA is collision-free in the commonly encountered scenarios.

Let us first consider a fixed and multi-hop topology. We now prove that two data packets cannot collide.
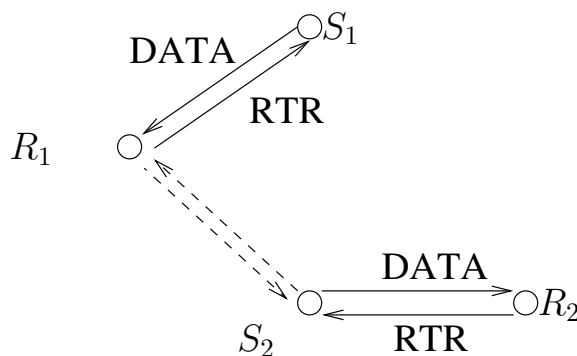
Figure D.2: Scenario for the demonstration of the correctness of CROMA.

We suppose that a collision of two data packets occurs at a receiver R_1 (see figure D.2). These packets have been sent by two different senders, namely S_1 and S_2. During the RTR-mini-slot, R_1 specified the MAC address of the sender, say S_1, that was allowed to send its data in the current slot. As the MAC address is unique, a single colliding data packet is destined to R_1. Therefore, we know that the data packet of S_2 was destined to another receiver, R_2.

Now, as R_1 has received a data packet from S_2 and links are bi-directional, S_2 has received the RTR of R_1. Moreover, S_2 has also received a RTR from R_2, since it sent a data packet destined to R_2. Thus, S_2 has detected a collision of RTRs in the current slot and nevertheless it has sent a packet to R_2. This is impossible. As a conclusion, no data collision can occur in a fixed topology.

Let us now consider the case of a dynamic topology. Two concurrent communications on a slot are shown on the top of figure D.1, from node 1 to node 2 and from node 3 to node 4. These communications are sharing the same slot in frame j and are far enough to avoid any mutual interference. For mobile nodes and at the next frame $j + 1$, node 3 can either stay out of range of nodes 1 and 2, enter the communication range of 1, 2, or both. Same alternatives can occur for node 4. Thus, after mobility, a total of 16 relative new positions are possible. Because of the symmetry of the problem, only 10 cases are shown in figure D.1.

The left hand side of figure D.1 shows situations, where a single communication is interrupted because the sender detected a collision of RTRs on the considered slot. For example, in case b, node 4 moved in the transmission range of nodes 1 and 2. In frame $j + 1$, nodes 2 and 4 send simultaneously an RTR. Node 3 receives correctly the polling of 4, whereas node 1 senses a collision during the RTR-mini-slot. Node 1 decides to interrupt the communication with node 2 and does not send any data

packet on this slot. If node 1 has still packets in its buffer, it has to enter a new reservation phase. The communication between nodes 3 and 4 continues normally.

The central part of figure D.1 shows exposed-terminal topologies, where both communications can still share the same slot. In case e, node 4 moved in the transmission range of node 2. In frame $j + 1$, node 1 (respectively 3) decodes the RTR of node 2 (respectively 4) because it is out of the transmission range of node 4 (respectively 2). Both nodes 1 and 3 can send data packet during the DATA-mini-slot.

The right hand side of figure D.1 shows topologies, where communications are released because both senders detected a collision of RTRs. Case j shows a configuration where the network of nodes is fully connected after mobility. Here, RTRs of nodes 2 and 4 collide at nodes 1 and 3. On detecting the collision, they decide to interrupt their communication.

So, in the commonly encountered cases, in both fixed and mobile environment, CROMA is collision-free. As in all protocols that rely on the exchange of short control packets, the capture effect may however affect this conclusion.

Appendix E

CROMA State Machine

In this appendix, we provide the finite state machine of CROMA as it is implemented in the Network Simulator v2 (figure E.1). The generic format for the messages attached to the arrows between states is [*event*|*action*], i.e., if an *event* occurs, the protocol behaves according to the *action*.
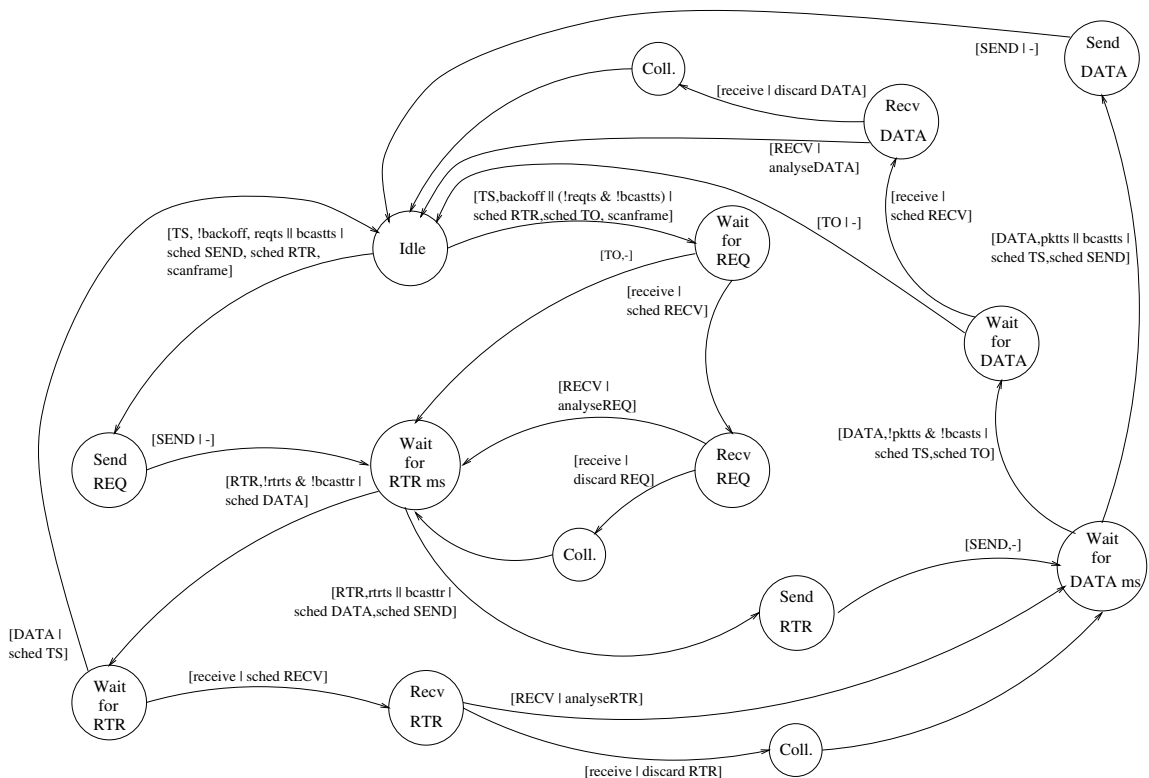
Figure E.1: CROMA finite state machine.

States

Idle: original state of CROMA.
Send REQ: CROMA is sending a REQ in the REQ-minislot.
Send RTR: CROMA is sending a RTR in the RTR-minislot.
Send DATA: CROMA is sending a DATA packet in the DATA-mini-slot.
Recv REQ: CROMA is receiving a REQ in the REQ-minislot.
Recv RTR: CROMA is receiving a RTR in the RTR-minislot.
Recv DATA: CROMA is receiving a DATA packet in the DATA-mini-slot.
Wait for REQ: CROMA is waiting for a REQ in the REQ-minislot.
Wait for RTR: CROMA is waiting for a RTR in the RTR-minislot.
Wait for DATA: CROMA is waiting for a DATA packet in the DATA-mini-slot.
Wait for RTR ms: CROMA has not received any REQ in the REQ-mini-slot and is waiting for the start of the RTR-mini-slot.
Wait for DATA ms: CROMA has not received any RTR in the RTR-mini-slot and is waiting for the start of the DATA-mini-slot.
Coll.: CROMA has detected a collision of packets.

Timers

TS: start of the time-slot.
RTR: start of RTR-mini-slot.
DATA: start of the DATA-mini-slot.
TO: TO is scheduled at the beginning of a mini-slot in states 'Wait for REQ', 'Wait for RTR', and 'Wait for DATA'. If CROMA has not received any packet before TO expires, it leaves the state and waits for the next mini-slot. The reception of a packet in this mini-slot is not possible any more.
back-off: end of the back-off.
RECV: end of the reception of a packet.
SEND: end of the transmission of a packet.
RTR: start of RTR-mini-slot.

Functions

scanframe: choose a slot for sending a REQ according to the scheduling policy.
sched: schedules a timer.
discard: drops a packet.

analyseREQ: analyses a received REQ.

analyseRTR: analyses a received RTR.

analyseDATA: analyses a received DATA packet.

Variables

reqts: A REQ has to be sent.

rtrts: A RTR has to be sent.

pktts: A DATA packet has to be sent.

bcastts: A broadcast packet has to be sent.

Other Event

receive: A packet is being received.

Appendix F

Optimal Transmission Range

In this section, we provide the solutions of the following equation with unknown variable r:

$$\theta(1-\theta)(\lambda\pi r^2)^2 e^{-(1-\theta)\lambda\pi r^2} = 1 \ . \tag{F.1}$$

Let us recall first of all some preliminary results on the Lambert W-function [75, 199]. This function is the inverse of:

$$f(W) = We^W \ . \tag{F.2}$$

$W(x)$ is real for $x \geq -1/e$. W is two-valued for $-1/e \leq x < 0$. For $W(x) \geq -1$, the function is denoted $W_0(x)$ and is called the principal branch, for $W(x) \leq -1$, the function is denoted $W_{-1}(x)$. The Lambert W-function is plotted in figure F.1. Equation F.1 is equivalent to

$$t^2 e^{2t} = \frac{1-\theta}{4\theta} \ , \tag{F.3}$$

where $t = -(1-\theta)\lambda\pi r^2/2$. The left hand side of equation F.3 is exactly $[f(t)]^2$. As a consequence:

$$f(t) = \epsilon\sqrt{\frac{1-\theta}{4\theta}} \ , \tag{F.4}$$

where $\epsilon = \pm 1$. If $\epsilon = 1$, this equation has a unique solution given by the Lambert W-function. Otherwise, it has at least a real solution if

$$-\sqrt{\frac{1-\theta}{4\theta}} \ \geq \ -\frac{1}{e} \ , \text{i.e.,} \tag{F.5}$$

$$\theta \ \geq \ \frac{1}{4e^{-2}+1} \ . \tag{F.6}$$
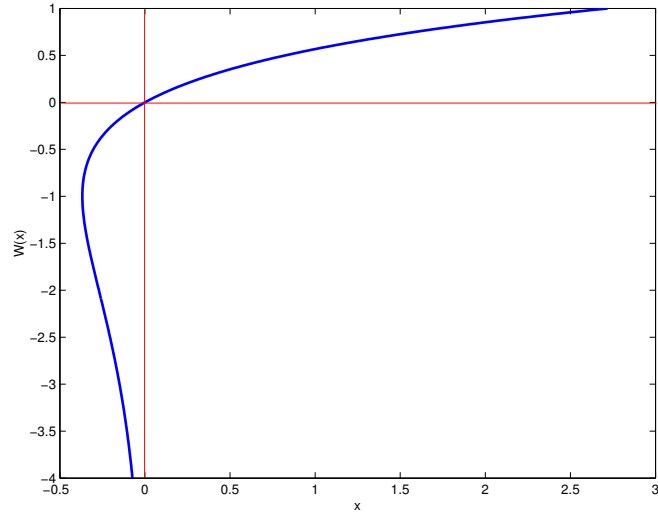
Figure F.1: Lambert W-function.

Hence, if a solution exists,

$$t \;=\; W\left(\epsilon\sqrt{\frac{1-\theta}{4\theta}}\right) \tag{F.7}$$

$$r^2 \;=\; \frac{2W\left(\epsilon\sqrt{\frac{1-\theta}{4\theta}}\right)}{\lambda\pi(\theta-1)} \tag{F.8}$$

As $r \geq 0$ and $\theta - 1 \leq 0$, $\epsilon = -1$. As explained before, in this case, W is two-valued, so that

$$r_{opt1} \;=\; \sqrt{\frac{2W_0\left(-\sqrt{\frac{1-\theta}{4\theta}}\right)}{\lambda\pi(\theta-1)}} \tag{F.9}$$

$$r_{opt2} \;=\; \sqrt{\frac{2W_{-1}\left(-\sqrt{\frac{1-\theta}{4\theta}}\right)}{\lambda\pi(\theta-1)}} \;, \tag{F.10}$$

Numerical procedures for evaluating the function can be found e.g. in [75] or [66].

Appendix G

Capacity with Optimal Power and Rate Adaptation

In this section, we look for the power allocation $S(\gamma)$ that maximizes the following function:

$$f(q) = \int_{\gamma_0}^{\infty} \log_2(1 + q\gamma)p_\gamma(\gamma)d\gamma \ , \tag{G.1}$$

where $q = S(\gamma)/S$. The optimization problem is subject to the average power constraint:

$$g(q) = \int_{\gamma_0}^{\infty} qp_\gamma(\gamma)d\gamma = 1 \ . \tag{G.2}$$

We assume that $\nabla g \neq 0$, so that this problem is equivalent to maximize

$$f(q) - \lambda g(q) = \int_{\gamma_0}^{\infty} (\log_2(1 + q\gamma) - \lambda q)p_\gamma(\gamma)d\gamma \ , \tag{G.3}$$

without restriction, where λ is the Lagrange multiplier.

Recall that q is a function of λ. Hence, the Euler-Lagrange differential equation applies:

$$\frac{\partial}{\partial q}L - \frac{d}{d\gamma}\left(\frac{\partial}{\partial \dot{q}}L\right) = 0 \ , \tag{G.4}$$

with $L(q, \dot{q}, \gamma) = (\log_2(1 + q\gamma) - \lambda q)p_\gamma(\gamma)$. Taken into account that $q \geq 0$, this equation reduces to

$$\forall \gamma \in [\max(\lambda, \gamma_0); \infty[, \ \frac{\gamma}{1 + q\gamma} - \lambda \ = \ 0 \tag{G.5}$$

$$\forall \gamma \in [\max(\lambda, \gamma_0); \infty[, \ q \ = \ \frac{1}{\lambda} - \frac{1}{\gamma} \ . \tag{G.6}$$

Replacing $q(\gamma) = S(\gamma)/S$ in equations G.1 and G.2, we get equations 4.26 and 4.27.

Appendix H

Additional Results for Section 4.5

H.1 Proof of Equation 4.72

In this section, we recall the proof of equation 4.72 given in [153]. A given node a is considered. We are interested in the probability that a sender t sends a packet to a (see figure 4.25). Let us define two events and a probability: (T) the event that t sends a packet to a; (D) the event that t sends in the half circle that contains a; $N(i)$, the probability that there are i other terminals besides a in the half circle of radius R_0 from t. We assume that all directions are equiprobable, so that $P[D] = 1/2$. Moreover, since we do not allow packet to go away from their destination:

$$
\begin{aligned}
P[T] &= P[T|D]P[D] & \text{(H.1)} \\
&= \frac{1}{2} \sum_{i=0}^{\infty} P[T|D, N(i)] P[N(i)|D] & \text{(H.2)} \\
&= \frac{1}{2} \sum_{i=0}^{\infty} \frac{1}{i+1} \frac{(\lambda \pi R_0^2/2)^i}{i!} e^{-\lambda \pi R_0^2/2} & \text{(H.3)} \\
&= \frac{1 - e^{-\lambda \pi R_0^2/2}}{\lambda \pi R_0^2} \ . & \text{(H.4)}
\end{aligned}
$$

This terminates the proof.

H.2 Proof of Equation 4.85

In this section, we recall the proof of equation 4.85 given in [153]. We consider a sender-receiver pair, t-a, as shown in figure 4.26. The forward progress of a packet destined to a final destination F is the projection of $\overrightarrow{ta}$ on $[t; F)$. Let us define Z to be the random variable denoting the forward progress. The angle θ is randomly distributed over $]-\pi/2; \pi/2[$. Now, for a given distance r between t and a:

$$Pr[Z \leq z] = Pr[|\theta| \geq \cos^{-1}(z/r)] . \tag{H.5}$$

Thus, letting $F(z)$ the cdf of Z, we have:

$$F(z|r) = \begin{cases} 1 & \text{for } r < z \\ 1 - \frac{2\cos^{-1}(z/r)}{\pi} & \text{for } 0 \leq z \leq r \leq R_0. \end{cases} \tag{H.6}$$

After differentiation, we obtain the conditional pdf of Z:

$$f(z|r) = \begin{cases} 0 & \text{for } r < z \\ \frac{2}{\pi r \sqrt{1-z^2/r^2}} & \text{for } 0 \leq z \leq r \leq R_0. \end{cases} \tag{H.7}$$

z takes values in $[0; r]$, so that the conditional expected forward progress is:

$$E[Z|r] = \frac{2}{\pi} \int_0^r \frac{z}{\sqrt{r^2 - z^2}} dz = \frac{2r}{\pi} . \tag{H.8}$$

If now $d(r)$ is the distribution of r:

$$E[Z] = \int_0^{R_0} \frac{2r d(r)}{\pi} dr . \tag{H.9}$$

This terminates the proof.

H.3 Integration of Equation 4.88

In this appendix, we prove the following expression: $I = k/n$ with

$$I = \int_0^{R_0} \frac{2r}{R_0^2} \sum_{i=0}^{k-1} \binom{n-1}{i} \left(\frac{r^2}{R_0^2} \right)^i \left(1 - \frac{r^2}{R_0^2} \right)^{n-1-i} \tag{H.10}$$

$$= \sum_{i=0}^{k-1} \sum_{j=0}^{n-1-i} \binom{n-1}{i} \binom{n-i-1}{j} \frac{2(-1)^j}{R_0^{2(i+j+1)}} \int_0^{R_0} r^{2(i+j)+1} dr \tag{H.11}$$

$$= \sum_{i=0}^{k-1} \sum_{j=0}^{n-1-i} \binom{n-1}{i} \binom{n-i-1}{j} \frac{(-1)^j}{i+j+1} \tag{H.12}$$

$$= \sum_{i=0}^{k-1} (n-i-1)! \binom{n-1}{i} \sum_{j=0}^{n-1-i} \frac{(-1)^j}{(i+j+1)(n-i-j-1)!j!} \tag{H.13}$$

$$= \sum_{i=0}^{k-1} \frac{(n-1)!}{i!} \sum_{j=0}^{n-1-i} \frac{(-1)^j}{(i+j+1)(n-i-j-1)!j!} \tag{H.14}$$

We will now prove that the last sum is exactly $i!/n!$, and this will terminate the proof of $I = k/n$. First of all, note that:

$$\frac{i!}{n!} = \prod_{k=1}^{n-i} \frac{1}{i+k} = \sum_{k=1}^{n-i} \frac{c_k}{i+k} \tag{H.15}$$

Let us determine the c_k. For that, we consider the following function:

$$f(x) = \prod_{k=1}^{n-i} \frac{1}{x+k} = \sum_{k=1}^{n-i} \frac{c_k}{x+k} \tag{H.16}$$

Now, it is clear that c_{k_0} verifies the following expression:

$$c_{k_0} = [f(x)(x+k_0)]_{x=-k_0} \tag{H.17}$$

$$= \left[(x+k_0) \prod_{k=1}^{n-i} \frac{1}{x+k} \right]_{x=-k_0} \tag{H.18}$$

$$= \prod_{k=1, k \neq k_0}^{n-i} \frac{1}{k-k_0} \tag{H.19}$$

$$= \frac{(-1)^{k_0-1}}{(k_0-1)!(n-i-k_0)!} \tag{H.20}$$

This expression terminates the proof because:

$$\frac{i!}{n!} = f(i) \tag{H.21}$$

$$= \sum_{k=1}^{n-i} \frac{(-1)^{k-1}}{(i+k)(k-1)!(n-i-k)!} \tag{H.22}$$

$$= \sum_{j=0}^{n-i-1} \frac{(-1)^j}{(i+j+1)(n-i-j-1)!j!} \tag{H.23}$$

H.4 Interference Characteristic Function in the MMSE Case

We are in the MMSE case. Let us look at the characteristic function of the interference for a given sender at a distance r from the receiver and for a SINR target γ_0. The interference created by the senders in a disk of radius R is:

$$Y_R = \sum_{r_i \leq R} \frac{PP_0}{P_0\gamma_0 + Pr_i^\alpha} \,. \tag{H.24}$$

Let $g(r) = PP_0/(P_0\gamma_0 + Pr^\alpha)$. The characteristic function of $Y = \sum_i I(P_i, P, \gamma_0) = \sum_i g(r_i)$ is:

$$\phi_Y(\omega) = \lim_{R\to\infty} E[e^{i\omega Y_R}] \,. \tag{H.25}$$

Let (D_R^k) be the event that there are k interferers in the disk of radius R. According to the Poisson distribution:

$$E[e^{i\omega Y_R}] = \sum_{k=0}^{\infty} E[e^{i\omega Y_R}|D_R^k]P[D_R^k] \tag{H.26}$$

$$E[e^{i\omega Y_R}] = \sum_{k=0}^{\infty} \frac{(\lambda p \pi R^2)^k}{k!} e^{-\lambda p \pi R^2} E[e^{i\omega Y_R}|D_R^k] \tag{H.27}$$

The characteristic function of a sum of independent random variables is the product of the characteristic functions:

$$E[e^{i\omega Y_R}|D_R^k] = \prod_{i=1}^{k} E[e^{i\omega g(r_i)}|D_R^k] \tag{H.28}$$

$$= \left(\int_0^R \frac{2r}{R^2} e^{i\omega g(r)} dr \right)^k \tag{H.29}$$

Replacing the above equation in equation H.27, we get:

$$E[e^{i\omega Y_R}] \;=\; exp\left(\lambda p\pi R^2\left(\int_0^R \frac{2r}{R^2}e^{i\omega g(r)}dr - 1\right)\right). \qquad \text{(H.30)}$$

Let us evaluate the following expression by successively integrating by parts and changing the integration variable:

$$\int_0^R 2re^{i\omega g(r)}dr - R^2 \;=\; R^2 e^{i\omega g(R)} - i\omega\int_0^R r^2 g'(r)e^{i\omega g(r)}dr - R^2 \qquad \text{(H.31)}$$

$$=\; R^2 e^{i\omega g(R)} + i\omega\int_{g(R)}^{g(0)}\left[g^{-1}(t)\right]^2 e^{i\omega t}dt - R^2 \qquad \text{(H.32)}$$

Note that $g(r) = PP_0/(P_0\gamma_0 + Pr^\alpha)$ with $\alpha > 2$. So, letting R grow indefinitely:

$$\lim_{R\to\infty}\int_0^R 2re^{i\omega g(r)}dr - R^2 = i\omega\int_0^{P/\gamma_0}\left[g^{-1}(t)\right]^2 e^{i\omega t}dt \qquad \text{(H.33)}$$

We can now conclude for the characteristic function of the interference:

$$\phi_Y(\omega) = exp\left(i\lambda p\pi\omega\int_0^{P/\gamma_0}\left(\frac{P_0}{t} - \frac{P_0\gamma_0}{P}\right)^{2/\alpha} e^{i\omega t}dt\right). \qquad \text{(H.34)}$$

Bibliography

Publications

[1] V. Capdevielle, T. Lestable, M. Coupechoux, M.-L. Alberi-Morel, and L. Brignol. Multi-Hop Coverage Extension of an IEEE 802.11b WLAN in a Corporate Environment. In *Proc. of WNCG Symposium, Austin, United States*, Oct. 2003.

[2] M. Cohen, L. Brignol, D. Rouffet, and M. Coupechoux. Internet rapide en zones peu denses. *Alcatel Telecommunications Review*, 3rd Quarter 2004.

[3] M. Coupechoux, B. Baynat, C. Bonnet, and V. Kumar. CROMA - a Slotted MAC Protocol for MANETs with Multi-slot Communications. In *Proc. of WNCG Symposium, Austin, United States*, Oct. 2003.

[4] M. Coupechoux, B. Baynat, C. Bonnet, and V. Kumar. Modeling of a Slotted MAC Protocol for MANETs. In *Proc. of MADNET'03, Sophia-Antipolis, France*, Mar. 2003.

[5] M. Coupechoux, B. Baynat, C. Bonnet, and V. Kumar. CROMA - an Enhanced Slotted MAC Protocol for MANETs. *Kluwer/ACM Mobile Networks and Applications (MONET)*, 2004. (to appear).

[6] M. Coupechoux, B. Baynat, T. Lestable, V. Kumar, and C. Bonnet. Improving the MAC Layer of Multi-Hop Networks. *Kluwer Wireless Personal Communications, Selected papers from the Strategic Workshop Sept. 2003, Helsingor, Denmark*, 29(1-2), April 2004.

[7] M. Coupechoux, C. Bonnet, and V. Kumar. A Scheduling Policy for Dense and Highly Mobile Networks. In *Proc. of IEEE Wireless'01, Calgary, Canada*, Jul. 2001.

[8] M. Coupechoux, C. Bonnet, and V. Kumar. CROMA - a New Collision-free Receiver-Oriented MAC Protocol for MANETs. In *Proc. of WTC'02, Paris, France*, Sept. 2002.

[9] M. Coupechoux, C. Bonnet, and V. Kumar. A Scheduling Policy for Dense and Highly Mobile Networks. *Gesellschaft für Informatik Edition Lecture Notes in Informatics, proc. of WMAN'02, Ulm, Germany*, 2002.

[10] M. Coupechoux, J. Brouet, L. Brignol, and V. Kumar. Suggested Solutions for the Near-far Effect in Multimode WLANs. In *Proc. of WWRF 10, New York, United States*, Oct. 2003.

[11] M. Coupechoux, V. Kumar, and L. Brignol. Voice over IEEE 802.11b Capacity. In *Proc. of ITCSS'04, Antwerp, Belgium*, Sept. 2004.

[12] M. Coupechoux and T. Lestable. On the Capacity of the Channel Aware Slotted Aloha over Rayleigh and Nakagami-m Channels. In *Proc. of WiOpt'04, Cambridge, United Kingdom*, Mar. 2004. (poster session).

[13] M. Coupechoux, T. Lestable, C. Bonnet, and V. Kumar. Throughput of the Multi-hop Slotted Aloha with Multi-packet Reception. *Springer Lecture Notes in Computer Science, proc. of IFIP WONS'04, Madonna di Campiglio, Italy*, Feb. 2004.

Public Presentations

[14] M. Coupechoux. A Collision-free Receiver-Oriented MAC (CROMA) Protocol for MANETs. In *IEEE MMT'02, Rennes, France*, June 2002. (invited presentation).

[15] M. Coupechoux. Un protocole MAC pour réseaux mobiles ad hoc. In *Laboratoire d'Informatique de Paris VI, Paris*, May 2002. (invited presentation).

[16] M. Coupechoux. Réseaux ad hoc et réseaux maillés. In *Les Rendez-Vous de l'Entreprise Innovante, Palaiseau, France*, Feb. 2004. (invited presentation).

[17] M. Coupechoux, T. Lestable, and V. Kumar. Improving the MAC Throughput of Multi-Hop Ad Hoc Networks. In *IEEE ComSoc/WWRF Workshop, New York, United States*, Oct. 2003. (invited presentation).

Alcatel Reports

[18] M. Coupechoux. Influence of ARP on Packet Delay in Ad Hoc Networks. *Alcatel R&I Report UTR/L/02/0043*, Nov. 2001.

[19] M. Coupechoux. Traffic Generators for ns2. *Alcatel R&I Report UTR/L/02/0044*, Mar. 2001.

[20] M. Coupechoux. Interference Computation in an Ad Hoc Cloud for Coverage Extension. *Alcatel R&I Report UTR/L/02/0047*, Sept. 2002.

[21] M. Coupechoux. Link Adaptation in ns2. *Alcatel R&I Report UTR/L/02/0045*, Sept. 2002.

[22] M. Coupechoux. Capacity Study of IEEE 802.11b. *Alcatel R&I Report UTR/L/03/0011*, Mar. 2003.

[23] M. Coupechoux. Near-far Effect in IEEE 802.11b. *Alcatel R&I Report UMS/L/03/0008*, Aug. 2003.

[24] M. Coupechoux. WLAN Coverage Extension with Relaying. *Alcatel R&I Report UMS/L/03/0010*, Aug. 2003.

[25] M. Coupechoux. High Speed Internet: Capacity Study. *Alcatel R&I Report UMS/L/04/0017*, Feb. 2004.

[26] M. Coupechoux. Mesh Networks Evaluation I. *Alcatel R&I Report UMS/L/04/0004*, Jan. 2004.

[27] M. Coupechoux. Mesh Networks Evaluation II. *Alcatel R&I Report UMS/L/04/0005*, Jan. 2004.

Filed Patents

[28] M.-L. Alberi-Morel, M. Edimo, H. Maillard, T. Lestable, and M. Coupechoux. Radio Frequency Device Allowing Rx/Tx Communications on the Same Antenna, Aug. 2003.

[29] M. Coupechoux and J. Brouet. Wireless Mobile Terminal and Telecommunication System, Apr. 2004.

[30] M. Coupechoux and V. Kumar. Method for Providing Access to a Data Network for a Vehicle Travelling on a Road, Feb. 2001.

[31] M. Coupechoux and V. Kumar. A Method for Selecting a Path to Establish a Telecommunication Link, Feb. 2003.

[32] V. Kumar and M. Coupechoux. Mobile Stations with Two Communications Interfaces, Sept. 2001.

[33] V. Kumar and M. Coupechoux. Fast Delivery of Multimedia Messages in Cellular Networks, Aug. 2003.

[34] V. Kumar, M. Coupechoux, and H. Maillard. Method for Establishing a Connection between Terminals Having a Short Range Interface, Jan. 2002.

References

[35] ANSI/IEEE Std 802.11. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 1999.

[36] ANSI/IEEE Std 802.11b. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Higher Speed Physical Layer Extension in the 2.4 GHz band, 1999.

[37] N. Abramson. The Aloha System - Another Alternative for Computer Communication. In *Proc. of Fall Joint Computer Conf., AFIPS Conf.*, 1970.

[38] M.-S. Alamouti and A. J. Goldsmith. Capacity of Rayleigh Fading Channels Under Different Adaptive Transmission and Diversity-combining techniques. *IEEE Trans. on Vehicular Technology*, 48(4), July 1999.

[39] G. Anastasi, L. Lenzini, and E. Mingozzi. Stability and Performance Analysis of HIPERLAN. In *Proc. of IEEE INFOCOM'98*, Mar. 1998.

[40] F. Anjum, M. Elaoud, D. Famolari, A. Ghosh, R. Vaidyanathan, A. Dutta, P. Agrawal, T. Kodama, and Y. Katsube. Voice Performance in WLAN Networks - An Experimental Study. In *Proc. of IEEE GLOBECOM'03*, Dec. 2003.

[41] H. Anouar, Y. Souilmi, and C. Bonnet. Self-Balanced Receiver-Oriented MAC for Ultra-Wide Band Mobile Ad Hoc Networks. In *Proc. of ACM MOBIHOC'03*, June 2003.

[42] E. Arikan. Some Complexity Results about Packet Radio Networks. *IEEE Trans. on Information Theory*, 30(4), July 2004.

[43] M. G. Arranz, R. Agüero, L. Munoz, and P. Mähönen. Behavior of UDP-based Applications over IEEE 802.11 Wireless Networks. In *Proc. of IEEE PIMRC'01*, Sept. 2001.

[44] B. Awerbuch, D. Holmer, and H. Rubens. High Throughput Route Selection in Multi-rate Ad Hoc Wireless Networks. *Springer Lecture Notes in Computer Science, proc. of IEEE WONS'04, Madonna di Campiglio, Italy*, 2004.

[45] F. Baccelli, B. Blaszczyszyn, and P. Mühlethaler. A Spatial Reuse ALOHA MAC Protocol for Multihop Wireless Mobile Networks. Technical Report 4955, INRIA, Oct. 2003.

[46] F. Baccelli, K. Tchoumatchenko, and S. Zuyev. Markov Path on the Poisson-Delaunay Graph with Applications to Routing in Mobile Networks. Technical Report 3420, INRIA, May 1998.

[47] J. Q. Bao and L. Tong. A Performance Comparison Between Ad Hoc and Centrally Controlled CDMA Wireless LANs. *IEEE Trans. on Wireless Communications*, 1(4), Oct. 2002.

[48] L. Bao and J. J. Garcia-Luna-Aceves. A New Approach to Channel Access Scheduling for Ad Hoc Networks. In *Proc. of ACM/IEEE MOBICOM'01*, July 2001.

[49] C. M. Barnhart, J. E. Wieselthier, and A. Ephremides. A Neural Network Approach to Solving the Link Activation Problem in Multihop Radio Networks. *IEEE Trans. on Communications*, 43(2/3/4), Feb./Mar./Apr. 1995.

[50] J. P. Barthélemy, G. Cohen, and A. Lobstein. *Complexité algorithmique et problèmes de communications*. Masson, Paris, France, 1992.

[51] B. Baynat. *Théorie des files d'attente, des chaînes de Markov aux réseaux à forme produit*. Hermes Science Publications, Paris, 2000.

[52] P. Bergamo, A. Giovanardi, A. Travasoni, D. Maniezzo, G. Mazzini, and M. Zorzi. Distributed Power Control for Energy Efficient Routing in Ad Hoc Networks. *ACM/Kluwer Wireless Networks*, 10(1), Jan. 2004.

[53] C. Bettstetter and C. Wagner. The Spatial Node Distribution of the Random Waypoint Mobility Model. *Gesellschaft für Informatik Edition Lecture Notes in Informatics, proc. of WMAN'02, Ulm, Germany*, 2002.

[54] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang. MACAW: a Media Access Protocol for Wireless LAN's. In *Proc. of ACM SIGCOMM*, Aug. 1994.

[55] G. Bianchi. Performance Analysis of the IEEE 802.11 Distributed Coordination Function. *IEEE Journal on Selected Areas in Communications*, 18(3), Mar. 2000.

[56] L. Blazevic, L. Buttyan, S. Capkun, S. Giordano, J.-P. Hubaux, and J.-Y. Le Boudec. Self Organization in Mobile Ad Hoc Networks: the Approach of Terminodes. *IEEE Communications Magazine*, 39(6), June 2001.

[57] J. Broch, D.A. Maltz, D.B. Johnson, Y.-C. Hu, and J. Jetcheva. A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols. In *Proc. of IEEE MOBICOM'98*, Oct. 1998.

[58] S. Buchegger and J.-Y. Le Boudec. Performance Analysis of the CONFIDANT Protocol: Cooperation of Nodes - Fairness in Dynamic Ad Hoc Networks. In *Proc. of ACM MOBIHOC'02*, June 2002.

[59] Z. Cai and M. Lu. SNDR: a New Medium Access Control for Multi-Channel Ad Hoc Networks. In *Proc. of IEEE VTC'00*, May 2000.

[60] F. Cali, M. Conti, and E. Gregori. Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit. *IEEE/ACM Trans. on Networking*, 8(6), Dec. 2000.

[61] V. Capdevielle. Point-to-Multipoint Propagation in Rural Environment. Technical Report UTR/L/03/007, Alcatel R&I Report, Aug. 2003.

[62] V. Capdevielle and M.-L. Alberi-Morel. Coverage Simulator. Technical Report UTR/L/02/0086, Alcatel R&I Report, Dec. 2002.

[63] V. Capdevielle, T. Lestable, and L. Brignol. Cellular Capacity Enhancement thanks to Link Adaptation in HIPERLAN/2. In *Proc. of IST Mobile Communication Summit*, Sept. 2001.

[64] G. Chakraborty and T. Hirano. Genetic Algorithm for Broadcast Scheduling in Packet Radio Networks. In *Proc. of IEEE ICEC'98, part of WCCI'98*, May 1998.

[65] A. Chandra, V. Gummala, and J. O. Limb. Wireless Medium Access Control Protocols. *IEEE Communication Surveys*, Second Quarter 2000.

[66] F. Chapeau-Blondeau and A. Monir. Numerical Evaluation of the Lambert W Function and Application to Generation of Generalized Gaussian Noise with Exponent 1/2. *IEEE Trans. on Signal Processing*, 50(9), Sept. 2002.

[67] G. Chelius and E. Fleury. IETF MANET Working Group, Internet-Draft, IPv6 Addressing Architecture Support for Mobile Ad Hoc Networks, Sept. 2002.

[68] H. S. Chhaya. Performance Evaluation of the IEEE 802.11 MAC Protocol for Wireless LANs. Master's thesis, Illinois Institute of Technology, Chicago, USA, 1996.

[69] I. Chlamtac and A. Faragó. Making Transmission Schedules Immune to Topology Changes in Multi-Hop Packet Radio Networks. *IEEE/ACM Trans. on Networking*, 2(1), Feb. 1994.

[70] I. Chlamtac, A. Faragó, and H. Y. Ahn. A Topology Transparent Link Activation Protocol for Mobile CDMA Radio Networks. *IEEE Journal on Selected Areas in Communications*, 12(8), Oct. 1994.

[71] I. Chlamtac, A. Faragó, and H. Zhang. Time-Spread Multiple-Access (TSMA) Protocols for Multihop Mobile Radio Networks. *IEEE/ACM Trans. on Networking*, 5(6), Dec. 1997.

[72] A.-M. Chou and V. O. K. Li. Slot Allocation Strategies for TDMA Protocols in Multihop Packet Radio Networks. In *Proc. of IEEE INFOCOM'92*, May 1992.

[73] I. Cidon and M. Sidi. Distributed Assignment Algorithms for Multihop Packet Radio Networks. *IEEE Trans. on Computers*, 38(10), Oct. 1989.

[74] T. Clausen, P. Jacquet, and Projet INRIA Hipercom. IETF Network Working Group, Experimental RFC 3626, Optimized Link State Routing Protocol (OLSR), Oct. 2003.

[75] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert W function. *Advances in Computational Mathematics*, 5, 1996.

[76] S. Corson and J. Macker. IETF Network Working Group, Informational RFC 2501, Mobile Ad Hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations, Jan. 1999.

[77] S.R. Das, C.E. Perkins, and E.M. Royer. Performance Comparison of two On-Demand Routing Protocols for Ad Hoc Networks. In *Proc. of IEEE INFOCOM'00*, Mar. 2000.

[78] S. Diggavi, M. Grossglauser, and D. N. C. Tse. Even One-Dimensional Mobility increases Ad Hoc Wireless Capacity. In *Proc. of IEEE ISIT'02*, June 2002.

[79] J. Elson and D. Estrin. Time Synchronization for Wireless Sensor Networks. In *Proc. of the 15th Inter. Parallel and Distributed Processing Symposium*, Apr. 2001.

[80] J. Elson, L. Girod, and D. Estrin. Fine-Grained Time Synchronization using Reference Broadcasts. In *Proc. of the 5th Symposium on Operating System Design and Implementation*, Dec. 2002.

[81] V. Emamian, P. Anghel, and M. Kaveh. Multi-User Spatial Diversity in a Shadow-Fading Environment. In *Proc. of IEEE VTC'02*, Sept. 2002.

[82] A. Ephremides and B. Hajek. Information Theory and Communication Networks: an Unconsummated Union. *IEEE Trans. on Information Theory*, 44(6), Oct. 1998.

[83] A. Ephremides and T. V. Truong. Scheduling Broadcasts in Multihop Radio Networks. *IEEE Trans. on Communications*, 38(4), Apr. 1990.

[84] P. Bender et al. CDMA/HDR: a Bandwidth Efficient High Speed Wireless Data Service for Nomadic Users. *IEEE Communications Magazine*, 38(7), July 2000.

[85] esa European Commission. Galileo Mission High Level Definition v3.0, Sept. 1999. available on http://europa.eu.int/comm/dgs/energy_transport/galileo/index_en.html.

[86] M. Fainberg. A Performance Analysis of the IEEE 802.11b Wireless LAN-standard in the Presence of Bluetooth PAN. Master's thesis, Polytechnic University, New York, USA, 2001.

[87] R. H. Frenkiel, B. R. Badrinath, J. Borràs, and R. D. Yates. The Infostations Challenge: Balancing Cost and Ubiquity in Delivering Wireless Data. *IEEE Personal Communications*, 7(2), Apr. 2000.

[88] A. Frey and V. Schmidt. Marked Point Processes in the Plane I. *Notable Publications Inc. Advances in Performance Analysis*, 1, 1998.

[89] C. L. Fullmer and J. J. Garcia-Luna-Aceves. Solutions to Hidden Terminal Problems in Wireless Networks. In *Proc. of ACM SIGCOMM'97*, Sept. 1997.

[90] N. Funabiki and J. Kitamichi. A Gradual Neural-Network Algorithm for Jointly Time-Slot/Code Assignment Problems in Packet Radio Networks. *IEEE Trans. on Neural Networks*, 9(6), Nov. 1998.

[91] N. Funabiki, Y. Takenaka, and T. Higashino. A Gradual Neural-Network Approach for Broadcast Scheduling in Packet Radio Networks. In *Proc. of IEEE IJCNN'99*, July 1999.

[92] ITU-T Recommendation G.107. The E-Model, a Computational Model for Use in Transmission Planning, Mar. 2003.

[93] ITU-T Recommendation G.113. Transmission Impairments due to Speech Processing, Feb. 2001.

[94] ITU-T Recommendation G.114. One Way Transmission Time, May 2003.

[95] J. J. Garcia-Luna-Aceves and C. L. Fullmer. Floor Acquisition Multiple Access (FAMA) in Single-Channel Wireless Networks. *Kluwer/ACM Mobile Networks and Applications (MONET)*, 4(3), Jan. 1999.

[96] S. Garg and M. Kappes. Can I add a VoIP call? In *Proc. of IEEE ICC'03*, May 2003.

[97] S. Garg and M. Kappes. An Experimental Study of Throughput for UDP and VoIP Traffic in IEEE 802.11b Networks. In *Proc. of IEEE WCNC'03*, Mar. 2003.

[98] M. Gerla, R. Bagrodia, L. Zhang, and K. Tang. TCP over Wireless Multihop Protocols: Simulations and Experiments. In *Proc. of IEEE ICC'99*, June 1999.

[99] A. J. Goldsmith and P. P. Varaiya. Capacity of Fading Channels with Channel Side Information. *IEEE Trans. on Information Theory*, 43(6), Nov. 1997.

[100] B. Goode. Voice over Internet Protocol. *Proc. of the IEEE*, 90(9), Sept. 2002.

[101] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, 1980.

[102] M. Grossglauser and D. N. C. Tse. Mobility Increases the Capacity of Ad Hoc Wireless Networks. *IEEE/ACM Trans. on Networking*, 10(4), Aug. 2002.

[103] P. Gupta and P. R. Kumar. *Stochastic Analysis, Control, Optimization and Applications: a Volume in Honor of W. H. Fleming*, chapter Critical power for asymptotic connectivity in Wireless Networks. Springer Verlag, 1999.

[104] P. Gupta and P. R. Kumar. The Capacity of Wireless Networks. *IEEE Trans. on Information Theory*, 46(2), Mar. 2000.

[105] P. Gupta and P. R. Kumar. Towards an Information Theory of Large Networks: an Achievable Rate Region. *IEEE Trans. on Information Theory*, 49(8), Aug. 2003.

[106] Z. J. Haas and J; Deng. Dual Busy Tone Multiple Access (DBTMA): a Multiple Access Control Scheme for Ad Hoc Networks. *IEEE Trans. on Communications*, 50(6), June 2002.

[107] Z. J. Haas, M. Gerla, D. B. Johnson, C. E. Perkins, M. B. Pursley, M. Steenstrup, and C.-K. Toh. Guest Editorial, Wireless Ad Hoc Networks. *IEEE Journal on Selected Areas in Communications*, 17(8), Aug. 1999.

[108] B. Hajek and G. Sasaki. Link Scheduling in Polynomial Time. *IEEE Trans. on Information Theory*, 34(5), Sept. 1988.

[109] M. O. Hasna and M.-S. Alouini. End-to-end Performance of Transmission Systems with Relays over Rayleigh-Fading Channels. *IEEE Trans. on Wireless Communications*, 2(6), Nov. 2003.

[110] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda. Performance Anomaly of 802.11b. In *Proc. of IEEE INFOCOM'03*, Apr. 2003.

[111] G. Holland, N. Vaidya, and P. Bahl. A Rate Adaptive MAC Protocol for Multi-hop Wireless Networks. In *Proc. of ACM MOBICOM'01*, July 2001.

[112] G. Holland and N.H. Vaidya. Analysis of TCP Performance over Mobile Ad Hoc Networks. In *Proc. of IEEE MOBICOM'99*, Aug. 1999.

[113] Z. Huang and C.-C. Shen. A Comparison Study of Omnidirectional and Directional MAC Protocols for Ad Hoc Networks. In *Proc. of IEEE GLOBE-COM'02*, Nov. 2002.

[114] Z. Huang, C.-C. Shen, C. Srisathapornphat, and C. Jaikaeo. A Busy-Tone Based Directional MAC Protocol for Ad Hoc Networks. In *Proc. of IEEE MILCOM'02*, Oct. 2002.

[115] K.-W. Hung and T.-S. Yum. Fair and Efficient Transmission Scheduling in Multihop Packet Radio Networks. In *Proc. of IEEE GLOBECOM'92*, Dec. 1992.

[116] Texas Instruments. Voice over Packet. Technical Report SPEY005, White Paper, Jan. 1998.

[117] A. Iwata, C.-C. Chiang, G.Pei, M. Gerla, and T.-W. Chen. Scalable Routing Strategies for Ad Hoc Wireless Networks. *IEEE Journal on Selected Areas in Communications*, 17(8), Aug. 1999.

[118] R. Jain, D. Chiu, and W. Hawe. A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer Systems, DEC Research Report TR-301, Sept. 1984.

[119] J. Janssen and D. De Vleeschauwer. Designing Voice over IP Networks Based on Tolerable Mouth-to-Ear Delay. Technical Report TTD 454, Alcatel Corporate Research Center, Apr. 1999.

[120] J. Janssen, D. De Vleeschauwer, M. Büchli, and G. H. Petit. Assessing Voice Quality in Packet-based Telephony. *IEEE Internet Computing*, 6(3), May/June 2002.

[121] S. Jiang, J. Rao, D. He, X. Ling, and C. C. Ko. A Simple Distributed PRMA for MANETs. *IEEE Trans. on Vehicular Technology*, 51(2), Mar. 2002.

[122] D. B. Johnson, D. A. Maltz, and Y.-C. Hu. IETF MANET Working Group, Internet-Draft, The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR), Apr. 2003.

[123] J.-H. Ju and V. O. K. Li. An Optimal Topology-Transparent Scheduling Method in Multihop Packet Radio Networks. *IEEE/ACM Trans. on Networking*, 6(3), June 1998.

[124] J. Jubin and J. D. Tornow. The DARPA Packet Radio Network Protocols. *Proceedings of the IEEE*, 75(1), Jan. 1987.

[125] R. E. Kahn. The Organization of Computer Resources into a Packet Radio Network. *IEEE Trans. on Communications*, 25(1), Jan. 1977.

[126] A. Kamerman and L. Monteban. Wavelan II: a High-Performance Wireless LAN for the Unlicensed Band. *Bell Labs Technical Journal*, Summer 1997.

[127] P. Karn. MACA - a New Channel Access Method for Packet Radio. In *Proc. of ARRL/CRRL*, Apr. 1990.

[128] P. R. Karn, H. E. Price, and R. J. Diersing. Packet Radio in the Amateur Service. *IEEE Journal on Selected Areas in Communications*, 3(3), May 1985.

[129] L. Kleinrock and J. Silvester. Spatial Reuse in Multihop Channel Access Protocol. In *Proc. of IEEE INFOCOM'83*, Apr. 1983.

[130] L. Kleinrock and J. Silvester. Spatial Reuse in Multihop Packet Radio Networks. *Proceedings of the IEEE*, 75(1), Jan. 1987.

[131] L. Kleinrock and F. A. Tobagi. Packet Switching in Radio Channels: Part I - Carrier Sense Multiple-Access Modes and their Throughput-Delay Characteristics. *IEEE Trans. on Communications*, 23(12), Dec. 1975.

[132] L. Kleinrock and F. A. Tobagi. Packet Switching in Radio Channels: Part II - The Hidden Terminal Problem in Carrier Sense Multiple-Access and the Busy-Tone Solution. *IEEE Trans. on Communications*, 23(12), Dec. 1975.

[133] R. Knopp and P. A. Humblet. Information Capacity and Power Control in Single-Cell Multi-User Communications. In *Proc. of IEEE ICC'95*, June 1995.

[134] G. Kulkarni, V. Raghunathan, M. Srivastava, and M. Gerla. Channel Allocation in OFDMA based Wireless Ad Hoc Networks. In *Proc. of SPIE International Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations*, July 2002.

[135] Y. H. Kwon and D. C. Lee. An Uplink Packet Relay Protocol for CDMA Cellular-like Systems. In *Proc. of MILCOM'02*, Oct. 2002.

[136] B. M. Leiner, D. L. Nielson, and F. A. Tobagi. Scanning the Issue. *Proceedings of the IEEE*, 75(1), Jan. 1987.

[137] T. Lestable. Preliminary Cellular Performance with Link Adaptation. Technical Report UTR/L/01/0029, Alcatel R&I Report, Aug. 2002.

[138] C.-H. Lin and C.-Y. Liu. Media Access Control Schemes for Mobile Ad Hoc Networks. *Springer Lecture Notes in Computer Science, proc. of IEEE WONS'04, Madonna di Campiglio, Italy*, 2004.

[139]   C. R. Lin and M. Gerla. Asynchronous Multimedia Multihop Wireless Networks. In *Proc. of IEEE INFOCOM'97*, Apr. 1997.

[140]   C. R. Lin and M. Gerla. Real-Time Support in Multihop Wireless Networks. *ACM/Kluwer Wireless Networks*, 5(2), Mar. 1999.

[141]   F. Liu, K. Xing, X. Cheng, and S. Rotenstreich. *Resource Management in Wireless Networking*, chapter Energy Efficient MAC Layer Protocols in Ad Hoc Networks. Kluwer Academic Publisher, 2004. (to appear).

[142]   H.-H. Liu and J.-L. C. Wu. Packet Telephony Support for the IEEE 802.11 Wireless LAN. *IEEE Communications Letters*, 9(4), Sept. 2000.

[143]   J. Liu and S. Singh. ATCP: TCP for Mobile Ad Hoc Networks. *IEEE Journal on Selected Areas in Communications*, 19(7), July 2001.

[144]   Q. Liu, E.-H. Yang, and Z. Zhang. Throughput Analysis of CDMA Systems Using Multiuser Receivers. *IEEE Trans. on Communications*, 49(7), July 2001.

[145]   H. Luo, S. Lu, and V. Bharghavan. A New Model for Packet Scheduling in Multihop Wireless Networks. In *Proc. of ACM/IEEE MOBICOM'00*, Aug. 2000.

[146]   R. Lupas and S. Verdù. Near-Far Resistance of Multiuser Detectors in Asynchronous Channels. *IEEE Trans. on Communications*, 38(4), Apr. 1990.

[147]   H. Maillard. Mutual Time Synchronization for TDMA Support. Technical Report UMS/C/04/0029, Alcatel R&I Report, Jan. 2004.

[148]   D.A. Maltz, J. Broch, J. Jetcheva, and D.B. Johnson. The Effects of On-Demand Behavior in Routing Protocols for Multihop Wireless ad Hoc Networks. *IEEE Journal on Selected Areas in Communications*, 17(8), Aug. 1999.

[149]   S. Marti, T. J. Giuli, K. Lai, and M. Baker. Mitigating Routing Misbehavior in Mobile Ad Hoc Networks. In *Proc. of ACM/IEEE MOBICOM'00*, Aug. 2000.

[150]   P. Michiardi and R. Molva. CORE: a Collaborative Reputation Mechanism to Enforce Node Cooperation in Mobile Ad Hoc Networks. In *Proc. of IFIP CMS'02*, Sept. 2002.

[151] V.N. Muthiah and W.C. Wong. A Speech-Optimized Multiple Access Scheme for a Mobile Ad Hoc Network. In *Proc. of ACM MOBIHOC'00*, Aug. 2000.

[152] T. Nandagopal, T.-E. Kim, X. Gao, and V. Bharghavan. Achieving MAC Layer Fairness in Wireless Packet Networks. In *Proc. of ACM/IEEE MOBI-COM'00*, Aug. 2000.

[153] R. Nelson and L. Kleinrock. The Spatial Capacity of a Slotted ALOHA Multihop Packet Radio Network with Capture. *IEEE Trans. on Communications*, 32(6), June 1984.

[154] R. Nelson and L. Kleinrock. Spatial TDMA: a Collision-Free Multihop Channel Access Protocol. *IEEE Trans. on Communications*, 33(9), Sept. 1985.

[155] ns2 web page. http://www.isi.edu/nsnam/ns.

[156] R. Ogier, F. Templin, and M. Lewis. IETF Network Working Group, Experimental RFC 3684, Topology Dissemination Based on Reverse-Path Forwarding (TBRPF), Feb. 2004.

[157] D. Olsen and N. Dave. Dynamically Reserved Slot Management (DRSM) for Packet Radio Networks. In *Proc. of IEEE MILCOM'91*, Nov. 1991.

[158] ITU-T Recommendation P.59. Artificial Conversational Speech, Mar. 1993.

[159] C. Perkins, E. Belding-Royer, and S. Das. IETF Network Working Group, Experimental RFC 3561, Ad hoc On-Demand Distance Vector (AODV) Routing, July 2003.

[160] A. Polydros and J. Silvester. Slotted Random Access Spread-Spectrum Networks: an Analytical Framework. *IEEE Journal on Selected Areas in Communications*, 5(6), July 1987.

[161] L. C. Pond and V. O. K. Li. A Distributed Time-Slot Assignment Protocol for Mobile Multi-Hop Broadcast Packet Radio Networks. In *Proc. of IEEE MILCOM'89*, Oct. 1989.

[162] A. R. Prasad. Performance Comparison of Voice over IEEE 802.11 Schemes. In *Proc. of IEEE VTC'99*, Sept. 1999.

[163] J. G. Proakis. *Digital Communications, 3rd Ed.* MacGraw-Hill, Inc, 1995.

[164] A. Qayyum. *Analysis and Evaluation of Channel Access Schemes and Routing Protocols for Wireless Networks.* PhD thesis, Université Paris Sud - Paris XI, 2000.

[165] X. Qin and R. Berry. Exploiting Multiuser Diversity for Medium Access Control in Wireless Networks. In *Proc. of IEEE INFOCOM'03*, Apr. 2003.

[166] S. Ramanathan. A Unified Framework and Algorithm for (T/F/C)DMA Channel Assignment in Wireless Networks. In *Proc. of IEEE INFOCOM'97*, Apr. 1997.

[167] R. Ramaswami and K. K. Parhi. Distributed Scheduling of Broadcasts in a Radio Network. In *Proc. of IEEE INFOCOM'89*, Apr. 1989.

[168] V. Rodoplu and T. Meng. Minimum Energy Mobile wirless Networks. *IEEE Journal on Selected Areas in Communications*, 17(8), Aug. 1999.

[169] R. Rom and M. Sidi. *Multiple Access Protocols Performance and Analysis.* Springer-Verlag New York, Inc., New York, USA, 1990.

[170] E.M. Royer and C.E. Perkins. Multicast Operation of the Ad Hoc On-Demand Distance Vector Routing Protocol. In *Proc. of IEEE MOBICOM'99*, Aug. 1999.

[171] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly. Opportunistic Media Access for Multirate Ad Hoc Networks. In *Proc. of ACM MOBICOM'02*, Sept. 2002.

[172] C. Sankaran and A. Ephremides. The Use of Multiuser Detectors for Multicasting in Wireless Ad Hoc CDMA Networks. *IEEE Trans. on Information Theory*, 48(11), Nov. 2002.

[173] F. Schiavini. Comparison of Routing Protocols for Mobile Ad Hoc Networks. Master's thesis, University of Ferrara, Italy and Alcatel R&I, France, 2002.

[174] C. E. Shannon. Communication in the Presence of Noise. *Proceedings of the IEEE*, 86(2), Feb. 1998. (paper reprinted from the Proceedings of the IRE 37(1), Jan. 1949.).

[175] S. Singh and C. S. Raghavendra. PAMAS - Power Aware Multi-Access Protocol with Signalling for Ad Hoc Networks. *ACM Computer Communications Review*, 28(3), July 1998.

[176] S. Singh, M. Woo, and C. S. Raghavendra. Power-Aware Routing in Mobile Ad Hoc Networks. In *Proc. of ACM/IEEE MOBICOM'98*, Oct. 1998.

[177] N. Smavatkul, Y. Chen, and S. Emeott. Voice Capacity Evaluation of IEEE 802.11a with Automatic Rate Selection. In *Proc. of IEEE GLOBECOM'03*, Dec. 2003.

[178] J. L. Sobrinho and A. S. Krishnakumar. Distributed Multiple Access Procedure to Provide Voice Communications over IEEE 802.11 Wireless Networks. In *Proc. of IEEE GLOBECOM'96*, Nov. 1996.

[179] E. S. Sousa. Interference Modeling in a Direct-Sequence Spread-Spectrum Packet Radio Network. *IEEE Trans. on Communications*, 38(9), Sept. 1990.

[180] E. S. Sousa. Performance of a Spread Spectrum Packet Radio Network Link in a Poisson Field of Interferers. *IEEE Trans. on Information Theory*, 38(6), Nov. 1992.

[181] E. S. Sousa and J. A. Silvester. Optimum Transmission Ranges in a Direct-Sequence Spread-Spectrum Multihop Packet Radio Network. *IEEE Journal on Selected Areas in Communications*, 8(5), June 1990.

[182] D. S. Stevens and M. H. Ammar. Evaluation of Slot Allocation Strategies for TDMA Protocols in Packet Radio Networks. In *Proc. of IEEE MILCOM'90*, Sept. 1990.

[183] W J. Stewart. *An Introduction to the Numerical Solution of Markov Chains*. Princeton University Press, New Jersey, 1994.

[184] K. Sundaresan, V. Anantharaman, H.-Y. Hsieh, and R. Sivakumar. ATP: a Reliable Transport Protocol for Ad Hoc Networks. In *Proc. of ACM MOBI-HOC'03*, June 2003.

[185] H. Takagi and L. Kleinrock. Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals. *IEEE Trans. on Communications*, 32(3), Mar. 1984.

[186] F. Talucci, M. Gerla, and L. Fratta. MACA-BI (MACA by Invitation) - A Receiver Oriented Access Protocol for Wireless Multihop Networks. In *Proc. of IEEE PIMRC'97*, Sept. 1997.

[187] Z. Tang and J. J. Garcia-Luna-Aceves. A Protocol for Topology-Dependent Transmission Scheduling in Wireless Networks. In *Proc. of IEEE WCNC'99*, Sept. 1999.

[188] S. Toumpis and A. Goldsmith. Capacity Regions for Wireless Ad Hoc Networks. *IEEE Trans. on Wireless Communications*, 2(4), July 2003.

[189] D. N. C. Tse and S. V. Hanly. Linear Multiuser Receivers: Effective Interference, Effective Bandwidth and User Capacity. *IEEE Trans. on Information Theory*, 45(2), Mar. 1999.

[190] ETSI EN 300 352 v1.2.1. Broadband Radio Access Networks (BRAN); High Performance Radio Local Area Networks (HIPERLAN) type 1; functional specification, July 1998.

[191] ETSI TR 101 112 (UMTS 30.03) v.3.2.0. Selection Procedure for the Choice of Radio Transmission Technologies of the Universal Mobile Telecommunication System, Apr. 1998.

[192] N. H. Vaidya, P. Bahl, and S. Gupta. Distributed Fair Scheduling in a Wireless LAN. In *Proc. of ACM/IEEE MOBICOM'00*, Aug. 2000.

[193] A. van Moffaert and D. De Vleeschauwer. Adaptive Dejittering Mechanisms. Technical Report TTD 461, Alcatel Network Strategy Group, Nov. 2000.

[194] M. Veeraraghavan, N. Cocker, and T. Moors. Support of Voice Services in IEEE 802.11 Wireless LANs. In *Proc. of IEEE INFOCOM'01*, Apr. 2001.

[195] S. Verdù. *Multiuser Detection*. Cambridge University Press, 1998.

[196] P. Viswanath, D. Tse, and R. Laroia. Opportunistic Beamforming using Dumb Antennas. *IEEE Trans. on Information Theory*, 48(6), June 2002.

[197] D. De Vleeschauwer and J. Janssen. Voice Performance over Packet-based Networks. Technical Report 02340, Alcatel Technology White Paper, Oct. 2002.

[198] G. Wang and N. Ansari. Optimal Broadcast Scheduling in Packet Radio Networks Using Mean Field Annealing. *IEEE Journal on Selected Areas in Communications*, 15(2), Feb. 1997.

[199] E. W. Weisstein. *CRC Concise Encyclopedia of Mathematics*. CRC Press, 2002. available on mathworld.wolfram.com.

[200] C.-S. Wu and V. O. K. Li. Receiver-Initiated Busy-Tone Multiple Access in Packet Radio Networks. *ACM Computer Communication Review*, 17(5), Oct./Nov. 1987.

[201] K. Xu, M. Gerla, and S. Bae. How Effective is the IEEE 802.11 RTS/CTS Handshake in Ad Hoc Networks? In *Proc. of IEEE GLOBECOM'02*, Nov. 2002.

[202] S. Xu and T. Saadawi. Does the IEEE 802.11 MAC Protocol Work Well in Multihop Wireless Ad Hoc Networks? *IEEE Comm. Magazine*, 39(6), June 2001.

[203] H. Yang, H. Luo, F. Ye, S. Lu, and L. Zhang. Security in Mobile Ad Hoc Networks: Challenges and Solutions. *IEEE Wireless Communications*, 11(1), Feb. 2004.

[204] C. Zhu and M.S. Corson. A Five-Phase Reservation Protocol (FPRP) for Mobile Ad Hoc Networks. In *Proc. of IEEE INFOCOM'98*, Mar. 1998.