ECOLE NATIONALE SUPÉRIEURE DE TÉLÉCOMMUNICATIONS
DE PARIS - Télécom Paris

École Doctorale STIC

THÈSE

Présentée pour obtenir le titre de :
Docteur en SCIENCES
de Télécom Paris

Spécialité : Informatique et Réseaux

par

Lamia ROMDHANI-FILALI

Conception inter-couche de réseaux ad-hoc IEEE 802.11 pour le provisonnement de la qualité de service

Soutenue publiquement le 22 Septembre 2006 devant le jury composé de :

Président :	Jean-Claude	BELFIORE	ENST Paris
Rapporteurs :	Eric	FLEURY	INSA de Lyon
	Pascal	LORENZ	IUT de Colmar
Examinateurs :	Houda	LABIOD	ENST Paris
	Thierry	TURLETTI	INRIA Sophia-Antipolis
Directeur de thèse	Christian	BONNET	Institut Eurécom

(14:00 - Institut Eurécom, Sophia-Antipolis)

Conception inter-couche de réseaux adhoc IEEE 802.11 pour le provisonnement de la qualité de service

Lamia ROMDHANI-FILALI

Cross-Layer Design of IEEE 802.11-based Mobile Ad Hoc Networks for QoS Provisioning

DÉDICACES

À ma douce et adorable mère,
À mon père qui a tout sacrifié pour que j'arrive là où je suis,
À mes trés chèrs Lina et Fethi pour leurs grand amours et soutiens,
À mes grands parents et à toute ma famille qui ont su être avec moi chaque jour,
malgré la distance qui nous sépare et qui n'hésitent pas à m'adresser toute sorte d'aide.

LAMIA

ACKNOWLEGEMENTS

I would like to thanks individuals who have helped me make my thesis undertaking much easier and more importantly, successfully completed.

First and Foremost, I would like to acknowledge my supervisor Prof. Christian Bonnet for his encouragement and support over the past years and for making this dissertation possible. This dissertation is greatly indebted to his remarkable insights given during our discussions.

In addition, many thanks go to my commitee menbers: Mme. H.Labiod, Mr. JC.Belfiore, Mr T.Turletti, Mr E.Fleury, and Mr P.Lorenz for their time and insights into this study.

I would like to thank the members of the mobile communication group for creating an intellectually stimulating and fun research environment. I would like to say a big “thank you” to all the people at Eurecom who agreed to be interviewed by me for this thesis in some form or another.

I extend my unending thanks and utmost appreciation to my husband Fethi, for his constant support and useful comments. I could never have reached this milestone without his believing in me and giving me the love I could never need.

Special thanks go to my daugther Lina. Her first few words and nice smiles have lightened the burden of my workload, making it all seen worthwhile.

My thanks are due to Imed who provided many valuable suggestions. Additionally, I offer thanks to my sister Besma and my friend Manel that have given me unconditional support.

Many thanks to all friends, Balkine, Hanene, Zohra, Loubna, Aicha, Halima, Emmanuel, Said, Jihene, Nouha, Manel..., and all my neighbors, that made Nice a very enjoyable place to live.

I owe a very special debt of gratitude to my grand parents and all my family, who have raised me with the **greatest love** in the world. Throughout all my life, your love, support, guidance and endless patience have been truly inspirational. “Thanks” will never suffice. I love you very very much!

Lamia ROMDHANI
Sophia-Antipolis
September 22, 2006

Résumé

Les réseaux ad-hoc sont une particularité de réseaux informatiques, constitués de noeuds mobiles qui utilisent un mode de communication sans infrastructure et des liaisons radios. Chaque noeud mobile communique dans son rayon de portée d'émission/réception, et est totalement autonome quant à son déplacement, son fonctionnement et sa participation à l'acheminement des informations du réseau. L'utilisation des réseaux ad-hoc présente de nouveaux défis de part les problèmes cruciaux qu'ils posent, notamment les problèmes liés au support de communication qui est hertzien et donc de qualité variable dans l'espace et dans le temps.

La qualité du lien sans fil reste inconnue, et sujet à des variations selon la configuration du réseau et l'état du canal radio. Les enjeux s'étendent également à la couche d'accès au canal (par exemple Wi-Fi), à la couche réseau (en particulier aux algorithmes de routage) et à la couche transport (le comportement de TCP est sensible aux variations de délai). L'utilisation des liaisons radios introduit des différences notoires et de nouvelles problématiques par rapport aux communications filaires telles que la limitation physique ou règlementaire de la capacité disponible pour l'accès radio, la qualité fluctuante des liens radios (influence des obstacles, du mouvement, des interférences, ..), la position des points d'accès inconnue à l'avance et variable dans le temps. Au regard de ces limitations, les protocoles du modèle en couches du réseau filaire ne peuvent être transférés dans l'environnement sans fil sans adaptation. Un des enjeux en terme de recherche qui est apparu, est l?optimisation du fonctionnement des réseaux ad-hoc à travers l'utilisation de techniques innovantes qui permettent d'améliorer leurs performances. Les techniques multi-niveaux appelées "cross-layer" sont ainsi apparues pour faciliter le partage d'information entre les couches du modèle OSI et s'appliquent à tous les protocoles de divers niveaux, tant qu'il existe des interactions pour les quelles les performances globales du système peuvent être améliorées.

Motivations

La qualité du lien sans fil reste inconnue, et sujet à des variations suivant la configuration du réseau et l'état du canal radio. Les conditions de liens variables constituent des caractéristiques inhérentes de la plupart des MANETs. De plus, la topologie d'un réseau ad-hoc change en raison de la mobilité des noeuds. Ceci nécessite un routage dynamique car les routes peuvent disparaître en cours d'utilisation.

Fournir de la qualité de service dans les réseaux ad-hoc est un sujet complexe à cause des contraintes liées au mode ad-hoc et au médium sans fil qui est peu fiable et la bande passante limitée. L'utilisation traditionnelle du plus court chemin en termes de nombre de sauts n'est pas représentative de la qualité des liens sans fil, ce qui entraîne une réduction du débit. Il

manque une solution définitive capable de gérer de façon efficace la mobilité des utilisateurs et garantir de la qualité de service.

Le paradigme actuel de conception en couches est inflexible pour les réseaux sans fil. En effet, le passage de l'environnement filaire à l'environnement sans fil, très différents l'un de l'autre dans leurs caractéristiques propres, a conduit à considérer l'expansion des réseaux ad-hoc sous l'angle de techniques innovantes destinées à améliorer leurs performances. Une bonne planification de réseau est exigée afin de répondre aux besoins d'exécution particulièrement quand le protocole IEEE 802.11 est employé pour le support des applications en temps réel qui ont besoin de grandes ressources en termes de bande passante et de délai. En effet, les protocoles de routage, MAC et physique doivent être adaptatifs et coopératifs afin de faire face à la variation de ressources et de la topologie. De cette orientation ont émergés les systèmes « cross-layer » avec pour objectif de fournir une adaptation efficace des protocoles du modèle en couches du réseau filaire à l'environnement sans fil.

De nos jours, cette approche de conception inter-couche, qui constitue le sujet général de notre travail, est le concept le plus approprié dans les réseaux ad-hoc mobiles. Elle est adoptée pour résoudre plusieurs problèmes de communication ouverts. Elle vise à surmonter les problèmes de QoS dans les MANETs en permettant aux protocoles appartenant à des différentes couches à coopérer et partager l'information concernant le statut de réseau tout en maintenant une architecture en couches.

Cette thèse se concentre sur la conception, le développement, et l'évaluation des mécanismes cross-layer basés sur le protocole 802.11 et qui visent à améliorer la qualité de service dans ce type de réseaux.

Avant d'étudier les approches cross-layer, nous avons exploré l'idée d'améliorer, séparément, le fonctionnement d'une couche de communication; à savoir le protocole MAC 802.11e qui est conçu pour garantir la différentiation de service. Bien qu'on ait abouti à des améliorations en utilisant un nouveau mécanisme adaptatif de la couche MAC, les résultats des études de performances montrent les limitations de l'architecture en couches qui a, par ailleurs, démontré ses bons résultats dans les réseaux filaires.

Il y a de nombreuses possibilités d'interaction inter-couche. Dans nos travaux, nous intéressons au partage de paramètres entre la couche MAC et la couche routage. Nous pensons que la coopération entre ces deux couches fournit une meilleure amélioration de performance que la coopération entre d'autres couches. En particulier, nous nous concentrons sur le problème de l'acheminement des paquets de données d'une manière qui prend en compte le niveau de congestion du canal, les caractéristiques de réseau, et les besoins de protocoles de couches hautes. Nous proposons un routage optimal en ce qui concerne la stabilité des liens, le délai de bout-en-bout, et la conservation d'énergie avec et sans assistance/initiation du réseau. Nous proposons plusieurs mécanismes cross-layer qui visent à résoudre le problème de routage dans MANETs tout en augmentant la performance des métriques de QoS importantes tel que la stabilité du chemin, la consommation d'énergie, le délai de bout-en-bout. À cet effet, nous identifions les paramètres adéquats à partir de la couche MAC et la couche réseau et nous les adaptons pour fournir de la qualité de service basée sur de nouveaux algorithmes de coopération inter-couches.

En outre, nous identifions les avantages et les inconvénients des architectures inter-couche par rapport à l'architecture en couche traditionnelle pour améliorer la performance de protocoles de communication dans un environnement sans fil.

Analyse des performances et amélioration du mécanisme EDCA du 802.11e

EDCA (Enhanced Distributed Coordination Access) est un mécanisme de qualité et de différentiation de service pour l'accès au canal radio. Cette approche est présentée dans l'algorithme de base de HCF décrit dans la norme d'IEEE 802.11e. Ce mécanisme définit quatre priorités de trafic (la voix, vidéo, best-effort, background). Il se base sur un algorithme probabiliste, pondéré par la priorité pour l'accès au médium alors que le contrôle de l'accès reste distribué. Il assigne des petites valeurs de CWmin, CWmax, et AIFS aux classes de haute priorité ce qui leur favorise d'occuper fréquemment le canal. Autre nouveauté d'EDCA est le TxOP (Transmission Opportunities). Ce mécanisme de gestion de transmission, définit le droit d'accès d'une station et son temps alloué en fonction de son niveau de priorité. Si plusieurs stations, de catégories de trafics différents, accèdent au support en même temps, le TxOP, qui est un temps prédéterminé (trafic de plus haute catégorie, temps le plus long), donnera l'accès à la catégorie la plus prioritaire. Ce temporisateur s'ajoute à la fin du temporisateur de backoff. La valeur de ce paramètre est fixe et ne peut être adaptée selon la charge du réseau.

Nous nous sommes intéressé à faire des études analytiques de ce paramètre afin d'analyser l'effet de sa valeur sur différentes métriques telle que le débit, le délai, et les taux de pertes au niveau des files d'attentes. Ensuite, nous avons proposé un mécanisme d'ajustement de TXOP selon la priorité du trafic et de la congestion dans le réseau. Les résultats obtenus ont montré l'efficacité de l'adaptation introduite, sur les performances de l'application.

Mécanisme de routage à la demande fondé sur la prise en compte de taux de consommation de l'énergie

Une caractéristique essentielle des réseaux ad-hoc est que, tout comme les autres réseaux sans fil, ils sont soumis aux aléas d'un comportement dynamique. Les conditions du canal de transmission sont imprévisibles et variables dans le temps. L'échec des routes ainsi que la dégradation de la qualité des liens à cause de la mobilité, les multi-trajets, les interférences, et le manque de ressources, sont fréquents.

Au regard de l'importance de la conservation d'énergie dans les réseaux mobiles sans fil, nous nous somme intéressés à décrire un algorithme de routage réactif qui se base sur l'interaction entre les protocoles MAC et routage. Nous proposons une nouvelle approche adaptative qui vise à incorporer la métrique de l'énergie dans le processus de la sélection de la route. En effet, on définit le taux de consommation de l'énergie pour chaque noeud qui permet d'estimer sa durée de vie. Ensuite, on définit un coût qui correspond à cette durée ainsi qu'au niveau d'énergie. Cette information est alors utilisée pour le calcul des routes. Pour cela, on suit les étapes suivantes:

- **Calcul de la durée de vie éventuelle d'un noeud**

Dans notre algorithme, nous ne considérons pas uniquement le niveau d'énergie dans chaque noeud comme le cas de plusieurs mécanismes dans ce domaine. Cependant, nous tenons en compte aussi de taux de consommation d'énergie à chaque période constante de temps (T_{Update}) :

$$E_{rate}(j) = \frac{E_{remain}(j) - E_{remain}(j-1)}{T_{Update}} \tag{1}$$

Où $E_{remain}(j)$ est le niveau d'énergie restante calculé dans la période j comme suit:

$$E_{remain}(j) = \max\left\{ E_{current}(j) - \sum_{i=1}^{i=N_{pkts}} E_{Tx}(i), 0 \right\} \tag{2}$$

Où $E_{current}(j)$ est le niveau courant de l'énergie et $E_{Tx}(i)$ est la quantité d'énergie nécéssaire pour transmettre les N_{pkts} paquets quit sont déja dans la file d'attente. En considérant ces paramètres, nous obtenons une valeur plus précise sur le niveau de l'énergie.

Pour minimiser les transisions de taux de consommations, nous utilisons la méthode EWMA pour donner plus de précisions sur les valeurs de taux de consommation d'énergie. Alors $\overline{E_{rate}(j)}$ est exprimé comme suit:

$$\overline{E_{rate}(j)} = (1-\alpha) * E_{rate}(j) + \alpha * \overline{E_{rate}(j-1)} \tag{3}$$

Avec α est dans l'intervalle $[0,1]$. [1]

Certes, la considération du taux de consommation d'énergie inclut, implicitement, le niveau du congestion du noeud tout en évitant des complexes mesures sur le montant de transmissions et de réceptions des données.

Le taux de consommation d'énergie $E_{rate}(j)$ ainsi que le niveau de l'énergie restante $E_{remain}(j)$ permet d'éstimer, dans chaque intervale de temps j, la durée de vie $T_{lifetime}(j)$ d'un noeud dans le réseau comme suit:

$$T_{lifetime}(j) = \frac{E_{remain}(j)}{\overline{E_{rate}(j)}} \tag{4}$$

- **Calcul du coût d'une route**

En utilisant la valeur $T_{lifetime}(j)$, chaque noeud j peut calculer le coût de la route à chaque réception d'un message de recherche de route (RREQ). Ce coût est calculé comme suit:

$$cost_{node}^{(j)} = 1 - \frac{T_{lifetime}(j) * W_k}{T_{MaxLifetime}} \tag{5}$$

On note que $T_{MaxLifetime}$ est la valeur maximale $T_{lifetime}(j)$ d'un noeud dans le réseau et W_k est un facteur multiplicatif dans l'intervalle [0, 1] définit pour chaque niveau d'énergie. Par conséquent, nous définissons quatre valeurs de W_k se référant à quatre intervalles d'énergie comme montre le Tableau 4.1. Le premier est de 50% à 100% de la valeur initiale d'énergie. Les noeuds qui ont un niveau d'énergie dans cet intervalle sont les plus favorisés pour participer à l'établissement d'une route. Ainsi, ils sont assignés avec le plus grand poids qui est égal à 1. Le deuxieme intervalle est de 30% à 50%. Les noeuds dans cet intervalle sont moins favorisés à participer a l'établissement d'une route que les noeuds dans le premier intervalle. Ils sont

---

[1] La valeur de α, que nous utilisons est égale a 0.25. Cette valeur est obtenue suite a des simulations extentives.

assignés avec un poids égal à 0.75. Le troisieme intervalle est de 10% à 30%. Les noeuds dans cet intervalle ont un faible niveau d'énegie. Le protocole devrait éviter l'utilisation de ces noeuds s'il y a d'autres possibilités de routage. Le poids correspondant à cet intervalle est égal à 0.5. Le dernier intervalle va de 0% à 10%. Dans notre approche, ces noeuds sont fortement évités. En effet, ils sont assignés avec le plus petit poids qui est égal à 0.25. Ainsi, nous visons à prolonger la vie de ces noeuds s'il y a d'autres alternatives de routage.

Normalized remain energy= $\frac{E_{remain}(j)}{E_{max}}$ with E_{max} is the maximum energy level	W_k
[0, 0.1[	0.25
[0.1, 0.3[	0.5
[0.3, 0.5[	0.75
[0.5, 1]	1

Table 1: Correspondance entre $|E_{remain}|$ et W_k.

Maintenant nous expliquerons les avantages et les raisons derrière l'utilisation des poids dans le calcul de coût. La durée de vie d'un noeud calculée est une valeur estimée basée sur le taux moyen de consommation d'énergie dans les périodes précédentes. Cette durée de vie change à travers le temps selon le taux de consommation d'énergie. Donc il y a quelques cas où les noeuds ont moins d'énergie mais ils produisent une longue durée de vie parce qu'ils n'ont pas participé au processus d'expédition de données dans les périodes précédentes. Par conséquent, nous devrions prendre en compte leur niveau d'énergie en tant qu'un deuxième critère pour produire leur cout. Nous donnons l'exemple suivant pour expliquer ces cas.

Dans la figure 4.1, le noeud S veut communiquer avec le noeud D. Il envoie un RREQ a A et B. Les Noeuds A et B ont des valeurs $T_{lifetime}(j)$ egales à 65 sec et 62 sec, respectivement. Leurs énergies sont égales a 20W et a 50W, respectivement (On suppose, pour cet exemple, que la valeur initiale d'énergie est égale à 100W). Chaque noeud recevant le paquet de RREQ calcule son propre coût et l'inclut dans le paquet de RREQ et le broadcast de nouveau. Si nous n'employons pas le poids dans le calcul de coût et nous considérons seulement la valeur de $T_{lifetime}(j)$, la communication entre les noeuds S et D sera par A. Cependant, même si $T_{lifetime}(A) > T_{lifetime}(B)$ (qui peut être expliqué par le fait que A n'a pas participé beaucoup aux communications dans les dernières périodes), $E_{remain}(A) < E_{remain}(B)$. Alors, il est plus juste de faire B participer à l'établissement de la route entre S et D que A.

- **Intégration de notre approche dans le protocole de routage AODV**

Dans cette thèse nous intégrons nos mécanismes cross-layer dans le protocole AODV. Ce protocole est utilisé pour le routage d'informations à la demande. L'émission d'un paquet par un noeud du réseau donne lieu à une diffusion du message RREQ de requête de route lorsque le noeud ne dispose pas de route valide menant à la destination recherchée. Il peut s'agir d'une destination pas encore connue du noeud source ou d'une route ayant expirée de la table de routage. La diffusion du message RREQ est précédée de la mise en cache de son numéro de séquence et de l'adresse du noeud émetteur pendant un temps déterminé, pour éviter de retraiter le même message acheminé par les autres noeuds. Après la diffusion du message RREQ, le noeud se met en attente des messages RREP de réponse de route pendant un intervalle de temps fixé. Les routes sont établies en se basant sur le principe du plus court chemin en termes de nombre de sauts, qui existe entre l'expéditeur et le récepteur. Cette

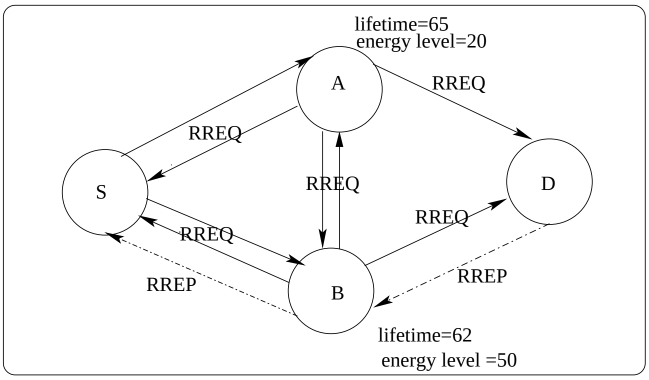

Figure 1: Example pour montrer les bienfaits d'introduire les poids utilisés dans le calcul du coût considéré dans l'établissement d'une route.

politique de sélection de route a été montrée moins pertinente. Ceci s'explique par le faite que la qualité des noeuds intermédiaires dans le chemin, en termes de niveau d'énergie, stabilité, congestion.., n'est pas prise en compte.

Nous intégrons notre mécanisme dans le protocole AODV. Cependant, le principe du plus court chemin n'est plus le critère du choix des routes. En effet, l'établissement des routes se fait en se basant sur le plus faible coût $\overline{cost_{path}}$ calculé de la source à la destination:

$$\overline{cost_{path}} = \frac{\sum_{i=1}^{i=\#int\_nodes} cost_{node}^{(i)}}{\#int\_nodes} \tag{6}$$

Où $\#int\_nodes$ est le nombre des noeuds intermédiaires.

• Evaluation des performances

L' environnement ns-2 est utilisé pour évaluer les performances de notre mécanisme. Nous considérons des sources de type CBR et un modèle de mobilité "random waypoint". Les scénarios utilisent une surface de simulations égale à : 1500 x 300 avec 50 noeuds. Le nombre de sources varient de 10 à 30, tous les liens sont considérés comme bi directionnels. Les critères de performance évalués sont:

- le gain en termes de conservation d'énergie
- le nombre de noeuds qui meurent à cause de leur faible énergie
- le ratio de paquets de données delivrés par rapport au pertes
- l'overhead de routage.

Nous présentons les résultats de performance obtenus avec plusieurs scénarios de mobilité. L'estimateur de la mobilité est le temps de pause.

La Figure 2 montre que la conservation de l'énergie augmente avec la charge de réseau et la mobilité. En effet, notre mécanisme conserve plus d'énergie que le protocole de base. Le gain atteind 26%, 17%, et 11% avec, respectivement 30, 20, et 10 sources. Nous présentons

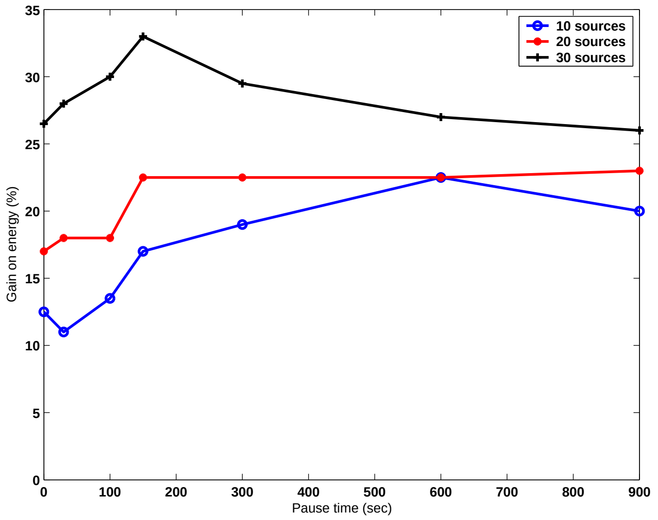

Figure 2: Gain en terme d'énergie

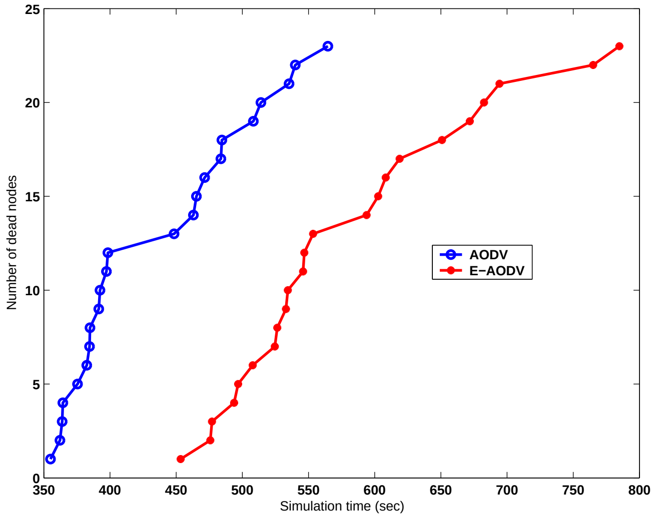

Figure 3: Nombre de noeud qui meurent durant la simulation

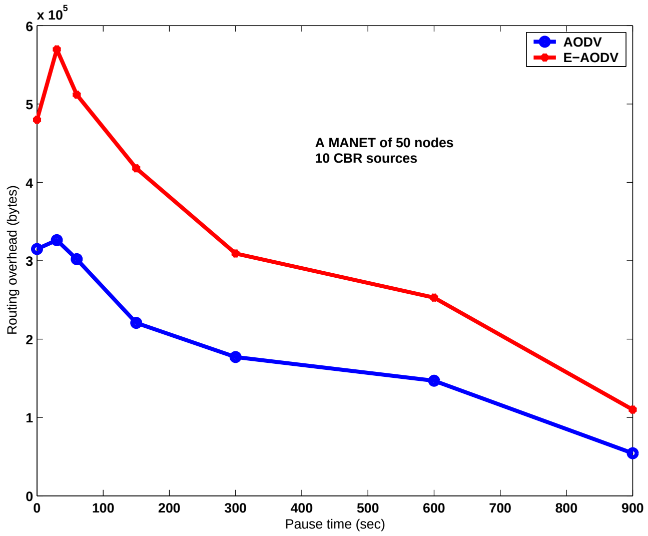

Figure 4: Overhead du routage obtenu avec 10 sources.

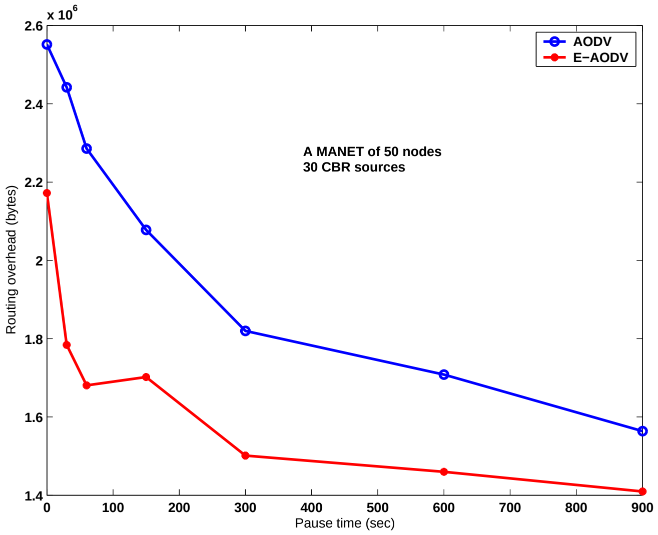

Figure 5: Overhead du routage obtenu avec 30 sources.

dans la figure 3 le nombre de noeuds qui meurent durant la simulation. cette métrique informe sur la connectivité du réseau. Avec notre mecanisme le premier noeud meurt à $350sec$ alors que avec AODV le premier noeud meurt à $450sec$.

Les résultats obtenus à partir du test des deux protocoles (Figures 4 et 5), E-AODV et AODV dans la configuration de 10 et 30 sources montrent que l'overhead du routage généré par E-AODV est moin faible que AODV uniquement lorsque la charge du réseau est forte (30 sources).

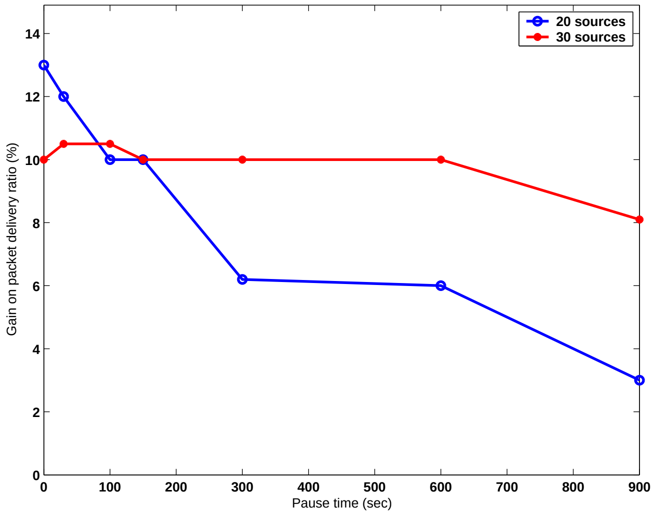

Figure 6: Gain en termes de ratio des paquets délivrés.

Le ratio des paquets délivrés à leurs destinations augmente avec notre approche par rapport au protocole de base, comme le montre la Figure 6. En effet, E-AODV prend avantage avec plus de sources (20 et 30). Il délivre entre 10 et 13% de paquets en plus qu'AODV, à forte mobilité (temps de pause trop court).

Les résultats que nous avons obtenus montrent l'efficacité de la métrique de l'énergie qu'on a introduit pour établir les routes entre les sources et les destinations. En effet, notre approche permet de conserver mieux de l'énergie et d'utiliser mieux les ressources de réseau qui sont affectées essentiellement par la forte mobilité et la congestion des noeuds.

Une approche cross-layer pour le cheminement de paquets dans les réseaux sans fil ad-hoc multi-sauts.

Les réseaux ad-hoc sont constitués de noeuds mobiles. Chaque noeud mobile communique dans son rayon de portée d'émission/réception. Deux noeuds qui ne sont pas dans le même rayon de portée l'un de l'autre, peuvent communiquer entre eux à travers d'autres noeuds. Le choix des noeuds intermédiaires est arbitraire. En effet, les routes sont établies en se basant sur le principe du plus court chemin. Ainsi, la qualité des noeuds qui participent à l'acheminement des informations du réseau, n'est pas prise en compte. Cela peut aboutir à un partage de charge inéquitable entre les noeuds du réseau. En effet, certains noeuds seront plus épuisés que d'autre. Pour éviter ce problème, nous proposons une nouvelle stratégie de cheminement des données dans les réseaux ad-hoc basé sur l'interaction entre le protocole MAC et le protocole de routage. Cette approche qu'on appelle F-AODV vise à réduire le nombre des noeuds expéditeurs dans le réseau afin d'alléger la contention pour l'accès au médium pour mieux utiliser le médium et optimiser la consommation d'énergie. En plus, on alloue plus de ressources pour les noeuds expéditeurs en leur adaptant les paramètres de la couche MAC suivant leur niveau de congestion. Cette adaptation leur permet d'accéder et utiliser plus le médium.

Les principes du mécanisme F-AODV

Nous présentons les principes du protocole cross-layer du F-AODV:

- Les noeuds intermédiaires (Forwarding Nodes :FN) sont sélectionnés suivant deux critères: leur niveau d'énergie $E[E]$ ainsi que leur niveau de charge ($E[B]$).
- Le niveau de charge de chaque noeud voisin est mesuré en observant tous les paquets reçus au niveau MAC. Cette information est ensuite passée au niveau réseau pour la considérer dans le processus de choix des routes.
- L'algorithme qu'on a adopté vise à réduire le nombre de FN afin de réduire les collisions et améliorer l'utilisation du médium.
- Le fonctionnement de cheminement est partagé périodiquement entre les noeuds qui figurent dans les routes alternatives. Les noeuds qui présentent un meilleur niveau d'énergie et faible charge sont choisis. Ceci est réalisé par le faite que la couche MAC partage avec la couche réseau des informations sur la qualité du noeud.
- Pour distinguer entre les noeuds qui participent au cheminement des données et qui ne le font pas, nous adaptons les paramètres de la couche MAC de telle sorte que les noeuds chargés sont plus favorisés d'occuper le médium plus que les autres. En effet, des petites valeurs de $CWmin$ et $CWmax$ leur sont assignées ainsi de plus longue durée de $TXOP$.

Evaluation des performances

Nous avons intégré notre mécanisme dans le simulateur ns2. Nous avons fait des extensions sur le protocole de routage AODV et le protocole EDCA de la couche MAC pour supporter la technique de l'interaction qu'on a proposé. L'évaluation des performances à laquelle nous procédons est destinée à montrer l'amélioration atteinte avec la nouvelle approche. En outre, elle est consacrée à comparer par simulation, la politique du cross-layer qu'on a présenté dans le mécanisme de F-AODV ainsi que celle introduite par E-AODV décrit dans la section précédente, avec celle du modèle en couche de base. Nous désirons montrer par cette comparaison les avantages ainsi que les limites d'un modèle cross-layer par rapport à un modèle de base.

Scénarios de simulations

Pour la mise en oeuvre, Nous considérons des sources de type CBR et un modèle de mobilité "random waypoint". Les scénarios utilisent une surface de simulations égale à : 1000 x 1000. Les paramètres de simulation utilisés sont illustrés dans le tableau 2. Chaque point des courbes obtenues, représente la moyenne de 10 itérations de simulations alors que la barre d'erreur représente 95% de l'intervalle de confiance.

Simulation time	900s
traffic	CBR, 4pkt/s
Packet size	512 bytes
Mac rate	2 Mbps
Initial speed	$Sp_{min} = 5m/s$, $Sp_{max} = 25m/s$
Speed	Uniform
Density	#nodes $* \frac{\Pi * range^2}{X_{dim} * Y_{dim}}$
Range	250m
Simulation area	1000*1000m
#nodes	40, 50, 60, 70, 80
Confidence Interval	95%

Table 2: Paramètres de simulations

Nous étudions l'effet de la densité des noeuds ainsi que l'impact de la variation de taux de trafic de données sur les performances des protocoles E-AODV, F-AODV et le protocole de base AODV.

Les critères de performance évalués sont:

- le rapport de laivraison des paquets
- taux d'erreur de routes.
- le délai de bout en bout

Impact de la densité des noeuds

Nous illustrons dans ce premier ensemble de simulations, l'influence de la densité des noeuds (en termes de nombre de voisin par noeud) calculé comme montré dans le tableau 2, sur les performances de E-AODV, F-AODV et AODV.

La courbe 7 fait ressortir le fait que, la densité du réseau augmente, les trois protocoles délivrent de moin en moin de paquets. Cependant, la politique cross-layer augmente le rapport de la livraison des paquets par raptport à la politique traditionnelle. Ces résultats nous ont permis de confirmer l'efficacité du partage d'information et de la cooperation qu'on a adopté entre les couches MAC et réseau. En outre, F-AODV donne un meilleur rapport de livraison que E-AODV (9% en plus) et AODV (14% en plus) quand on considere un réseau à forte charge. Ceci s'explique par le fait que l'adaptation des paramètres MAC utilisée par F-AODV augmente l'utilization du canal et que la réduction de nombre de FNs dans le réseau déminue les contentions et donc minimise les collisions.

Une observation similaire peut être faite sur la Figure 8, où on montre le taux d'erreur de route obtenus avec chaque protocole. Nous concluons des résultats que F-AODV a un minimum taux d'erreur et donc les échecs de routes sont moin fréquents par rapport à E-AODV et AODV.

Les courbes de la figure 9 permettent de comparer les délais moyens des trois protocoles. Pour chaque protocole pris isolement, la tendance de la courbe augmente quand la densité du réseau augmente. Dans le cas d'un réseau à faible charge, AODV affiche une latence moyenne

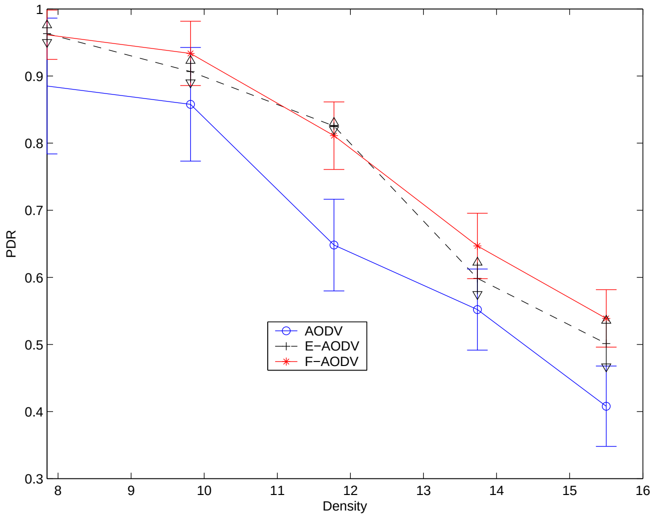

Figure 7: Résultats de ratio des paquets delivrés.

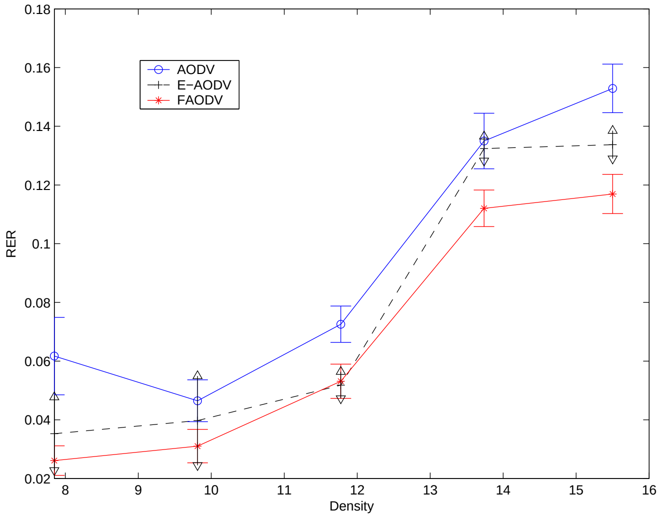

Figure 8: Taux d’erreur de route.

inférieure à celle de F-AODV et E-AODV. Cependant, quand la charge de réseau augmente, E-AODV démontre le plus court délai par rapport au deux autres protocoles.

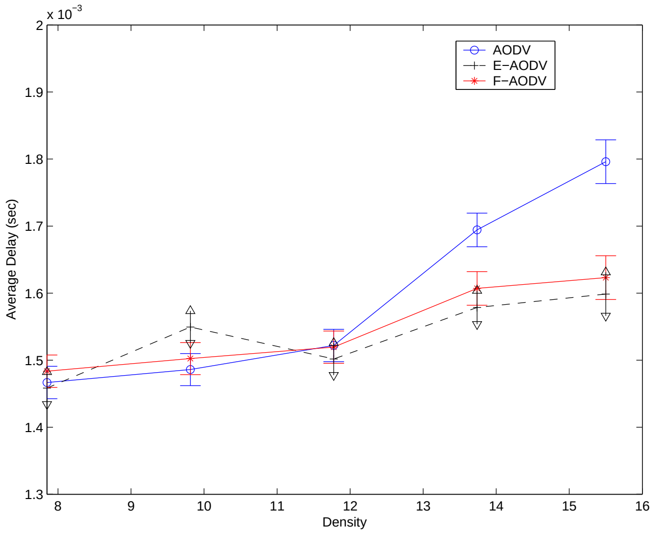

Figure 9: Résultat de délai.

Impact de la charge du trafic

Dans cet ensemble de simulations, on considère un scénario de 40 noeuds. On fait varier l'inter arrivée de paquets du trafic CBR pour chaque noeud. Les courbes de la figure 10 permettent d'illustrer l'influence de la variation de la charge sur la délivrance de paquets. On observe qu'à forte charge de réseau, qui correspond à un court inter arrivé de paquets, F-AODV et E-AODV donnent de meilleures performances par rapport à AODV. En effet, l'amélioration atteint 40% avec E-AODV et 30% avec F-AODV. Dans le cas d'un réseau à faible charge les trois protocoles affichent une performance similaire.

Impact de la mobilité

Dans cet ensemble de simulations, nous étudions l'influence de la mobilité des noeuds sur la performance de AODV, E-AODV et F-AODV. Pour cela, nous varions la vitesse initiale des noeuds mobiles ce qui fait varier la vitesse moyenne et donc varier la mobilité.

On note que tant que la mobilité des noeuds augmente, tant que la probabilité d'échec de route augmente. Par conséquent, le taux d'erreur indiquant la disparition des routes augmente aussi. De plus le manque d'énergie contribue à la rupture des liens entre les noeuds sources et destinations. Cependant, grâce aux techniques d'interaction qu'on propose pour E-AODV et pour F-AODV et qui considèrent les métriques d'énergie, leur taux d'erreur de routes affichés dans la Figure 11, sont inférieurs à ceux obtenus pour AODV.

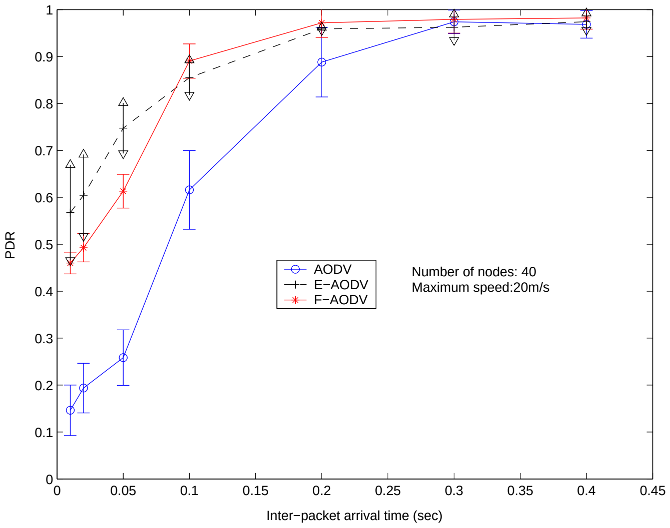

Figure 10: Rapport de la délivrance des paquets.

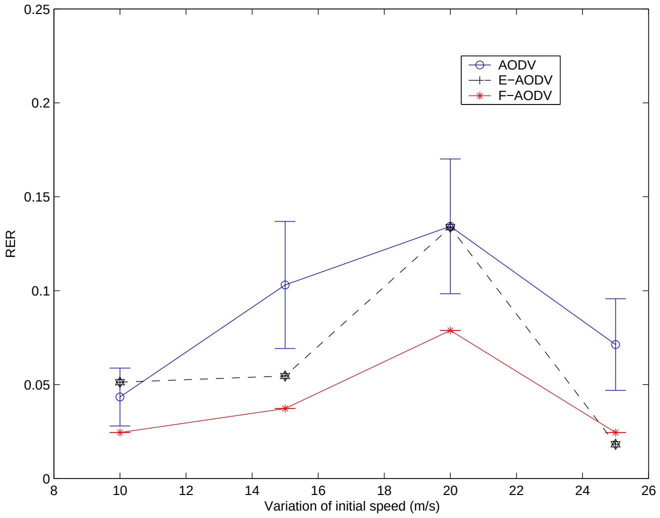

Figure 11: Résultat de taux d'erreur de route en fonction de la mobilité.

Un mécanisme de routage à la demande cross-layer basé sur la prise en compte de la stabilité des liens

La mobilité constitue un point important de la problématique du routage à QoS. Le choix du chemin stable est un défi important dans MANETs. Cependant, les protocoles traditionnels de cheminement ne tiennent pas compte des modèles de mouvement des noeuds et ils ne peuvent pas prévoir comment le voisinage va changer dans un proche avenir. Par conséquent, nous abordons ce problème en développant une nouvelle métrique, calculée grâce au partage d'information entre la couche MAC et routage. Ce nouveau paramètre permet de mesurer la stabilité des liens dans MANETs. La mise à jour de sa valeur est basée sur les mesures faites en captant les caractéristiques du réseau périodiquement. Nous développons un algorithme distribué (qu'on appelle S-AODV) permettant de calculer ce métrique de stabilité et le maintenir a jour dans chaque noeud sans fil dans le réseau. Ensuite, l'information de la stabilité est incorporée dans le processus de cheminement des données laissant optimiser le choix des noeuds composant un chemin entre chaque paire de noeuds.

Nous montrons l'éfficacité de cet algorithme par les simulations obtenus avec des différents scénarios testés dans le simulateur ns-2.

Principes de S-AODV

L'intérêt et l'importance de notre politique de routage réside dans la méthode de calcul de paramètre de stabilité. En effet, nous n'avons pas besoin de système de localisation supplémentaire(GPS, Galileo, etc.) comme le cas de plusieurs travaux précédents. Nous procédons à une intéraction entre la couche MAC et le protocole de routage pour construire les routes les plus stables.

Au niveau MAC, chaque noeud dans le réseau crée un vecteur de stabilité et un vecteur de charge. Dans ces deux vecteurs, on mesure la stabilité et la charge pour chaque voisin. La charge est calculée à base de l'observation des paquets reçus des voisins. Le calcul du paramètre du stabilité est fait suivant ces étapes suivantes:

- Pour chauque voisin i, un noeud j enregistre le nombre de messages de "HELLO" noté $n_{ij}(\triangle_{up})$ de i reçus pendant une période du temp $\triangle_{up}$. On a :

$$n_{ij}(\triangle_{up}) \leq n_{HELLO}^{max}, \tag{7}$$

Où n_{HELLO}^{max} est le nombre maximal de messages de "HELLO" qu'un noeud peut envoyer durant une période $\triangle_{up}$ car ces messages sont périodiquement émis. Ces messages peuvent être perdus à cause des collisions ou des corruptions.

- A chauque fin de la période $\triangle_{up}$ un noeud i calcule le taux de messages de "HELLO" reçu de chaque voisin j:

$$V_i(j) = s_{ij} = \frac{n_{ij}(\triangle_{up})}{n_{HELLO}^{max}} \tag{8}$$

  Absolument on a $0 \leq s_{ij} \leq 1$. On note qu'a cause du risque de collisions, le noeud i peut ne pas recevoir des messages de "HELLO" de j. Pour régler ce problème et augmenter la fiabilité de calcul, nous tenons en compte aussi des trames reçues des voisins si un "HELLO" n'est pas arrivé au temp attendu. Alors, si à un moment donné, un noeud

i qui espère recevoir un message d'un voisin j et ce message n'arrive pas, alors que i reçoit un autre paquet de contrôle ou un paquet de donnée de j. Dans ce cas la valeur de $n_{ij}(\triangle_{up})$ est incrémentée.

- Le paramètre de la stabilité du noeud j à i (ou la stabilité du lien $i \longrightarrow j$) est:

$$H_i(\triangle_{up}) = \frac{-\sum p_{ij} \log p_{ij}}{\log n_i}, \tag{9}$$

Où $p_{ij} = \frac{sij}{\sum s_{ik}}$, et k appartient à l'ensemble de voisins n_i du noeud i.

La couche réseau partage avec la couche MAC les données des deux vecteurs et elle les utilise dans le processus de routage.

L'inforamation de la plus petite valeur de stabilité est incorporée lors d'émission des messages RREQ et RREP. L'établissement de route est basé sur la principe de "max-min" de la valeur de stabilité observée au long du chemin. Deux noeuds qui ont la même valeur de stabilité, le noeud le moin chargé est choisi pour participer au cheminement des informations. Ceci est fait grâce au vecteur de charge calculé au niveau MAC.

Ceci permet de définir l'entropie d'un nombre d'évènements dans une base et d'en étudier la stabilité. Toutefois, le calcul d'entropie montre qu'étant donné deux noeud i et j ayant la meme valeur S_{ik} et S_{jk} vis à vis à un voisin k, le noeud qui a plus de voisin génère une plus grande entropie. Par la suite, ce noeud est plus favorisé d'être sélectioné pour l'établissement des routes. En effet, les courbes de la figure 12 démontrent cet aspect. L'entropie augmente lorsque le nombre de voisin augmente. Les noeuds qui ont une grande densité sont considérés les plus stables. En fait, en cas de répture de route il y aura plus de probabilté de rétablir le lien à cause de la forte densité.

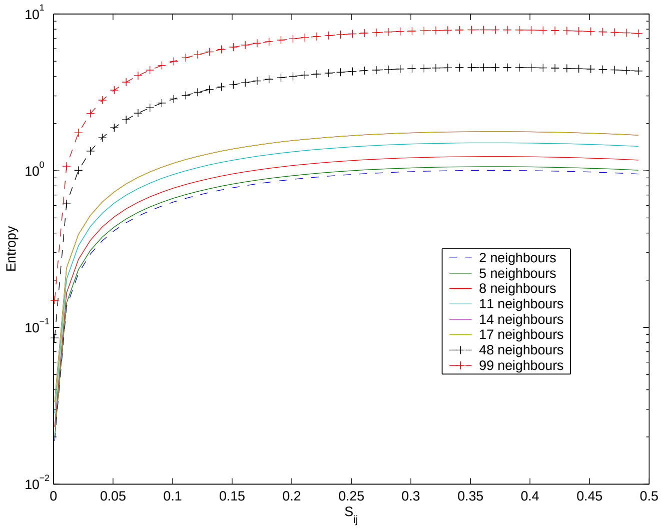

Figure 12: Entropie en fonction de nombre de voisins.

Evaluation des performances

Pour évaluer notre mécanisme, nous avons utilisé des paramètres et de scénarios de simulations similaires dans l'évaluation de performance de E-AODV. On considère 50 noeuds dont 30 émettent 4 paquets CBR chaque seconde. On fait varier le temps de pause des noeuds afin d'augmenter la mobilité. Nous présentons ici les résultats d'overhead de routage, le délai de bout en bout ainsi que le gain en termes d'énergie. Les trois métriques d'évaluations montrent un avantage de performance avec S-AODV. En effet, les figures 13, 14, et 15 affichent une amélioration de résultats avec S-AODV surtout dans le cas de scénario plus mobile. Ceci démontre l'efficacité du schéma d'interaction appliquée par S-AODV dans le processus de sélection des routes les plus stables. Cette politique favorise des liens avec plus de longévité et donc une meilleure utilisation de ressources.

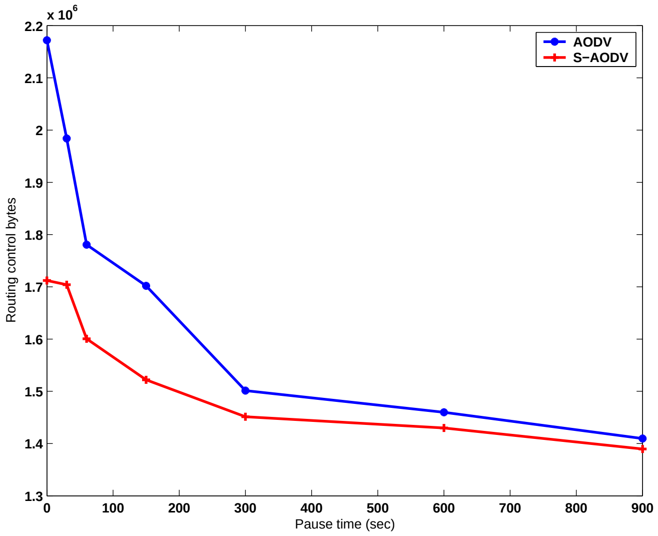

Figure 13: Overhead du routage.

Un mécanisme de routage cross-layer pour minimiser le délai de bout en bout pour les applications temp réel

Les applications en temps réel tôlèrent généralement la perte de paquets mais elles sont sensibles pour le délai de bout en bout et sa variation. Pour encourager le développement et le déploiement de telles applications dans MANETs, nous considérons également le délai, comme une métrique additive à optimiser en établissant un chemin entre deux noeuds dans le réseau. En effet, de même que les deux contributions précédentes, nous proposons également une interaction du MAC et du routage (qu'on appelle D-AODV) pour estimer dynamiquement et efficacement le délai de bout en bout pour chaque route candidate à l'établissement de communication entre les noeuds. Pour cela, chaque noeud intermédiaire dans la route doit mesurer et enregistrer le délai moyen d'accès au médium qu'elle observe au cours du temps.

Principes de D-AODV

- Objecti: Fournir un minimum délai de bout en bout pour des applications temps-réel.
- Au niveau MAC:

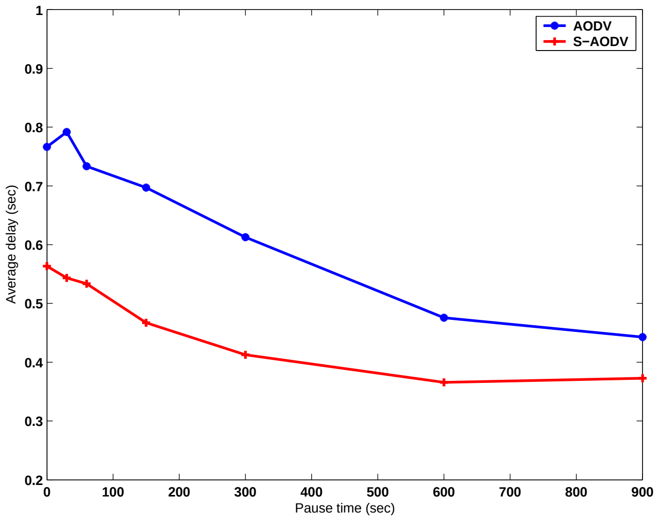

Figure 14: Délai de bout en bout.

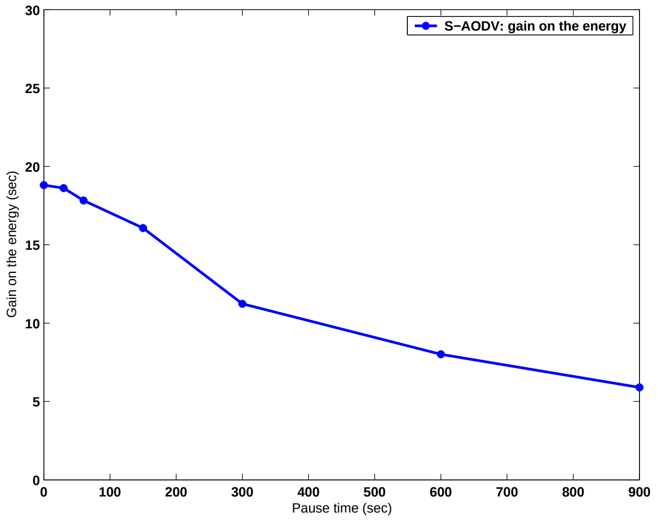

Figure 15: Résultats de gain d'énergie

  - un noeud mesure le délai moyen d'accès au médium en observant les trames de données transmises et les acquittements reçus.
  - appliquer une stratégie de scheduling favorisant les paquets de l'application qui a une contrainte de temps.

- Au niveau routage:
  - Un noeud dispose de la valeur de délai d'accès calculé au niveau MAC.
  - Ce délai est utilisé pour choisir et établir les routes entre les noeuds.
  - Les routes sont sélectionnées suivant le critère du minimum délai de bout en bout.

- Une autre politique appliquée dans notre approche cross-layer:
  - Chaque paquet circulant dans le réseau est assigné par sa durée de vie qui commence depuis son émission de la source
  - Un paquet est jeté si le délai nécessaire pour arriver à la destination est supérieur à celui toléré par l'application.

Evaluation des performances

Nous comparons dans cette partie, les résultats de la distribution de latence de D-AODV vis à vis à AODV de base. Nous considérons deux types de scénarios, un scénario stable (temp de pause=900) et un autre mobile (temp de pause = 0) avec trois types de trafics. Il ya 50 noeuds dans le réseau qui emèttent: 30 flux audios, 25 flux védios, et 29 flux de trafics background.

Les figures 17 et 16 montrent que D-AODV délivrent les paquets dans un court délai par rapport à AODV. Les résultats sont de plus, meilleurs lorsqu'on considère un scénario stable.

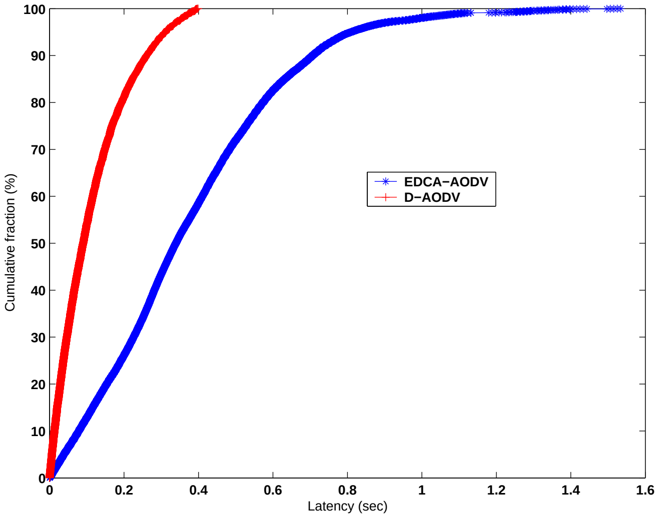

Figure 16: Distribution de la latence (temp de pause=0).

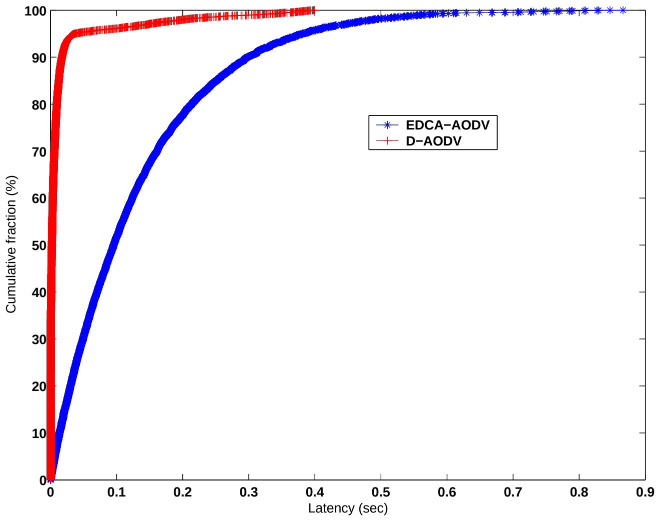

Figure 17: Distribution de la latence (temp de pause=900).

Comparaison des performances de S-AODV et D-AODV

Pour quantifier l'importance de tenir en compte les caractéristiques du réseau ainsi que la bonne métrique qui doit être optimisée dans le processus de routage, nous fournissons dans cette section des simulations de comparaison analysant le comportement de S-AODV et D-AODV dans différents scénarios

0.0.0.1 Environement de simulation

Nous considérons des scénarios formés de 50 noeuds localisés dans une surface rectangulaire de 1500x300m. Les sources sont CBR et génèrent un paquet UDP de taille 512 bytes chaque les 0.25 sec. On note que le nombre de source égal à 30. Les noeuds circulent dans le réseau avec une vitesse moyenne de $15m/s$. Nous faisons varier le temp de pause et nous analysons les résultats obtenus.

La figure 18, montre les résultats de délai moyen de bout en bout de deux mécanismes cross-layer que nous avons proposés. Les courbes affichent que le délai est amélioré avec D-AODV dans le cas d'un réseau static (temp de pause = 900). En effet cet algorithme permet d'acheminer les paquets à travers les noeuds les moin chargés. Cependant, cette bonne performance se dégrade quand la mobilté des noeuds augmente. En outre, le mécanisme S-AODV donne de meilleur résultats dans ce type de scénario dynamique.

De même, les courbes de la figure 19, représentants les résultats de livraison de paquets, font ressortir les mêmes remarques. En effet, avec D-AODV, le rapport de livraison de paquets par rapport aux pertes augmente lorsque les noeuds sont plus stables. Alors que S-AODV délivre plus de paquets que D-AODV quand les noeuds deviennent de plus en plus mobiles. Ces résultats démontrent l'importance de bien sélectionner le mécanisme cross-layer le plus adéquat et qui répond aux caractéristiques de réseaux ainsi qu'aux besoins de l'application. D'une part, l'efficacité de S-AODV est démontrée dans le cas des liens dynamiques car ce mécanisme est capable de sélectionner les routes les plus stables. D'autre part, avec D-AODV, une amélioration de performance est achevée uniquement en considérant un réseau statique.

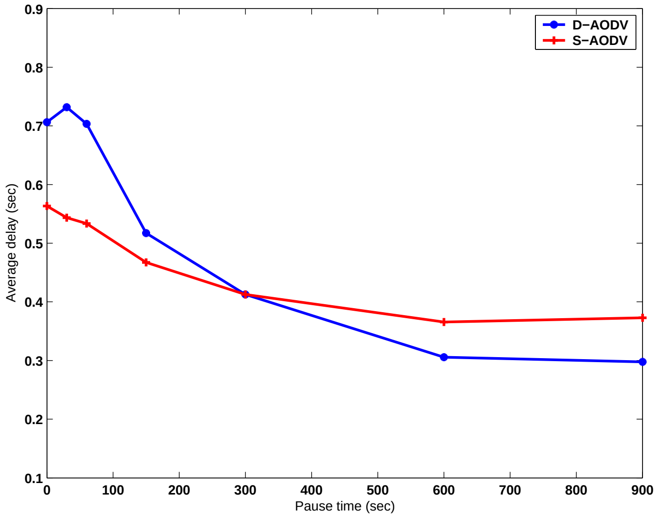

Figure 18: Délai de bout en bout

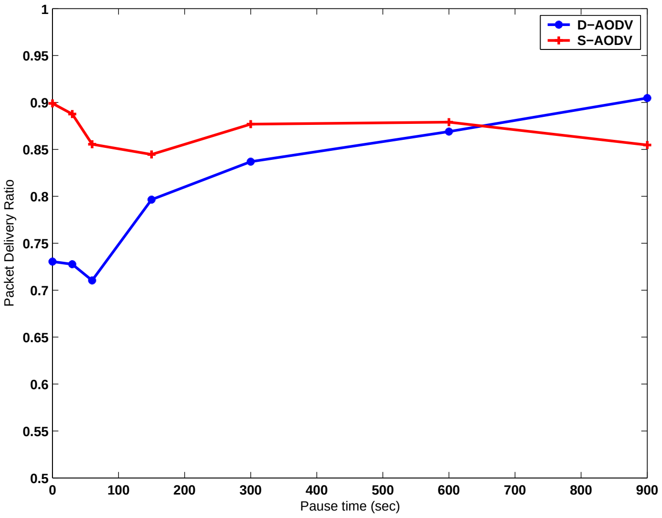

Figure 19: Fraction de délivrance de paquets

Par conséquent, nous concluons à partir des analyses de résultats, que la performance dépend fortement de caractéristiques de réseau.

Par cette étude comparative, nous démontrons que la conception cross-layer peut ne pas être efficace lorsque le comportement de réseau change fréquemment. Dans ce cas les différentes métriques utilisées pour permettre l'interaction entre les couches peuvent ne pas être fiables.

CrossAid (XAID): Une architecture cross-layer pour les réseaux mobiles ad-hoc

Ayant proposé plusieurs approches qui se basent sur l'interaction entre la couche MAC et le protocole de routage, nous évoquerons les avantages ainsi que les limites de performance obtenus avec ces mécanismes par rapport aux protocoles de base.

Tableaux récapitilatifs des performanes des interactions de nos cross-layer mecanismes

A cette dernière étape de la thèse, il nous est possible de déduire des tableaux de comparaison des performances de différentes interactions considérées dans notre travail par rapport à l'architecture traditionnelle en couche. Chaque tableau permet d'expliciter les bénéfices ainsi que les inconvénients obtenus par chaque protocole proposé. Nous présentons des évaluations dans les cas de fortes et faibles, mobilité et charge de réseau. Nous cherchons à identifier avec quels scenarios de réseau les modèles de cross-layer performent mieux que le modèle en couche.

Analyse des principaux résultats obtenus

- **Récaputilatif des résultats obtenus en terme de rapport de paquets delivrés**

Nous illustrons les résultats de comparaison en terme de Fraction de Paquets Délivrés (FPD) dans les cas d'un réseau à faible et à forte charge respectivement dans les tableaux 8.2 et 8.3.

Protocoles cross-layer	**mobilité faible**	**mobilite moyenne**	**mobilté forte**
E-AODV	-	$\simeq$	+
F-AODV	+ +	+ +	+
S-AODV	$\simeq$	+	+ +
D-AODV	+	-	-

Table 3: Résultats de FPD dans le cas d'un réseau à faible charge.

Le tableau 8.2 récapitule les résultats de FPD obtenus avec nos propositions dans le cas d'un réseau à faible charge. Nous observons une bonne amélioration de performance achevée avec F-AODV quand la mobilté des noeuds est faible ou moyenne. D'autre part, S-AODV donne une bonne performance dans le cas de forte mobilté. Cependant, il ya une dégradation de performance avec E-AODV dans le cas d'un scénario à faible mobilité. De même, avec D-AODV pour une moyenne et forte mobilité.

Le tableau 8.3 indique que F-AODV performe mieux que les protocoles de base spécialement dans le cas d'un réseau à faible et à moyenne mobilité. Bien que D-AODV performe

Protocoles cross-layer	mobilité faible	mobilite moyenne	mobilté forte
E-AODV	+ +	+ +	+ +
F-AODV	+ + +	+ + +	+ +
S-AODV	$\simeq$	+ +	+ + +
D-AODV	+	-	- -

Table 4: Résultats de FPD dans le cas d'un réseau à forte charge

moins bien dans le cas d'un scénario à forte mobilité, le mécanisme S-AODV fournisse une bonne amélioration en termes de FPD alors qu'il présente une performance similaire aux protocoles de base dans le cas d'un scénario à faible mobilité. Nous remarquons qu'avec E-AODV la performance est améliorée pour les différents niveaux de mobilité.

- **Récaputilatifs des résultats obtenus en termes de délai de bout en bout**

Protocoles cross-layer	mobilité faible	mobilite moyenne	mobilté forte
E-AODV	$\simeq$	-	-
F-AODV	-	-	+
S-AODV	-	-	+
D-AODV	+ +	+ +	+

Table 5: Résultats de délai dans le cas d'un réseau à faible charge

Protocoles cross-layer	mobilité faible	mobilite moyenne	mobilté forte
E-AODV	+ +	+ +	+ +
F-AODV	+	+ +	+ +
S-AODV	+	+	+ +
D-AODV	+ + +	+ +	+

Table 6: Résultats de délai dans le cas d'un réseau à forte charge

Le délai de bout en bout donne un avantage de rapidité en faveur du protocole D-AODV comme le montre le tableau 8.4. Cependant, avec les autres mécanismes de cross-layer, la tendance est inversée, AODV est plus rapide et affiche des meilleures performances en particulier avec un temps de pause long.

Dans le cas d'un réseau très chargé, le délai de tous nos protocoles proposés est moins court que celui des protocoles de base même avec une faible mobilité comme indique le tableau 8.5.

Les exemples des tableaux récapitulatifs présentés au dessus montrent que ces nouvelles approches sont efficaces dans certain cas alors que dans d'autres cas il n'y aura pas de besoin de considérer ces interactions vu qu'il n'y a pas une amélioration de performances. Par conséquent, nous proposons une nouvelle architecture (CrossAid:XAID) qui intègre le modèle en couche ainsi que les nouveaux mécanismes cross-layer. Le choix entre les différents mécanismes est basé sur les informations qui peuvent être collectés par les noeuds et qui concernent les caractéristiques du réseau en termes de charge et mobilité. Dans la Figure 8.4, nous présentons la conception de notre nouvelle architecture XAID.

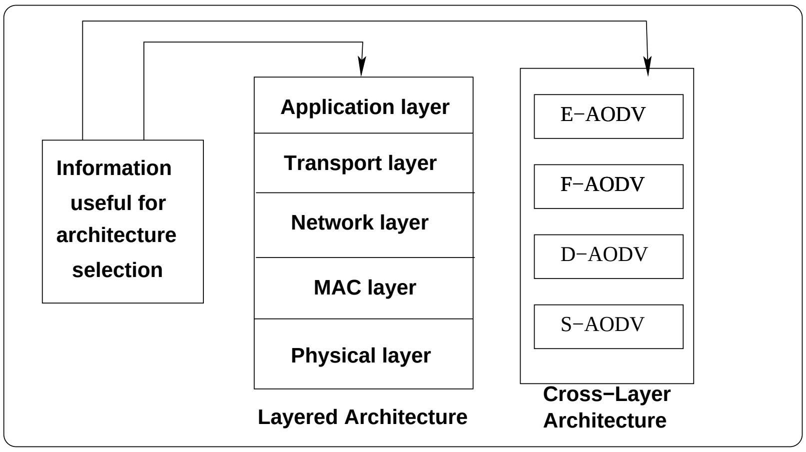

Figure 20: XAid: La nouvelle architecture proposée.

La conception cross-layer est une nécessité pour les réseaux ad-hoc parce qu'elle permet d'améliorer leurs performances comparativement aux réseaux câblés qui ne sont pas handicapés par les défaillances de même nature. Les évaluations des performances obtenues avec les mécanismes cross-layer proposés, montrent l'efficacité des adaptations appliquées par la politique de sélection des routes pour le partage de charge et pour fournir de la qualité de service. Cependant, face au manque des améliorations avec certains scénarios, il est indispensable d'appliquer juste le modèle traditionnel en couche afin d'éviter la complexité. Pour cette fin, nous avons décrit une nouvelle architecture qui mêle les modèles conceptuels favorisants des interactions entre les protocoles ainsi que la conception modulaire classique. Un module supplémentaire est introduit permettant de faire le choix entre un mécanisme ou un autre suivant les objectifs des applications, l'utilisation et les caractéristiques des réseaux, faisant intervenir pour cela des statistiques sur l'état de réseau.

L'importance de la méthode de conception se traduit également par la nécessité de conserver les acquis de l'architecture dont par exemple, la conception modulaire, la définition systématique des interactions entre les composants, la poursuite des objectifs à long terme quant à l'utilisation des réseaux, etc.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Internet today consists of thousands of access networks, which vary in scale from large wireline access networks, such as campus wide area networks or Internet Service Providers, supporting tens of thousands of users, to smaller wireless access networks supporting tens to hundreds of users. These access networks are interconnected by core networks which support hundreds of millions of users. All of these networks primarily offer two types of services: guaranteed service and best effort service.

In guaranteed service, the network provides some sort of service guarantee to individual users or groups of users. These guarantees are often in the form of ensuring that the throughput for a group of users is greater than some minimum value or that the delay experienced is smaller than some threshold. Often times, the guarantee is offered in a statistical sense, e.g. providing a particular data rate x% of the time.

In best effort service, the network makes no promises. This service is typically used by elastic traffic which consists of traffic where users do not necessarily have any minimum requirements, but would like to get as much data through to their respective destinations as quickly as possible. Individual user data flows react to congestion in the network and adapt their transmission rate with the aim of minimizing congestion. Email service is often provided on a best effort basis.

In parallel to the evolution of user services, there continues to be rapid adoption of wireless technology, and when coupled with the explosive growth of the Internet, it is clear that there will be increasing demand for wireless data services. Traffic on future wireless networks is expected to be a mix of real-time traffic such as voice, multimedia teleconferencing, and games, and data-traffic such as WWW (World Wide Web) browsing, messaging and file transfers. All of these applications will require widely varying and very diverse quality of service (QoS) guarantees for the different types of offered traffic, and we are now in the early days of this eventual amalgamation. Various mechanisms have been proposed and recently deployed to support data traffic over wireless media. This has ranged from Wireless Local Area Networks (WLANs), mainly based on the IEEE 802.11b or HiperLAN standards, to wireless wide area networks (WWANs) where data services are supported in the WiMAX (Worlwide Interoperability for Microwave Access), 2.5G, and 3G and beyond system versions.

In this thesis we explore the support of QoS in IEEE 802.11-based wireless networks through cross layer design perspectives. This standard defines two operation modes: infrastructure and ad hoc mode. The 802.11 infrastructure mode requires an access point implementing specific algorithms to control and coordinate the access to the WLAN and serves as a gateway to the Internet. On the other hand, an ad hoc wireless network consists of a group of mobile nodes and all communication is carried out through wireless links in a distributed fashion without a centralized controller. It has different properties when operating in different nodal movement patterns, performing different tasks and carrying varieties of patterns of traffic. The topology of an ad hoc network varies as a result of the mobility of its mobile hosts and the links break down and set up more frequently. A number of factors such as limited transmission range and power limitations, force long-distance communication in ad hoc networks to go through multi hops and each intermediate node is not responsible for the traffic it relays. Routing in ad hoc networks has to adapt to the unexpected link breakage and topology changes. To discover and maintain the routes in ad hoc networks requires more control traffic, which makes the task of performing ad hoc network routing more complex and less efficient. Indeed, due to the random movement of nodes, the bandwidth and power limitations, and the lack of fixed infrastructure, the development of efficient protocols to support the various networking operations in mobile ad hoc networks (e.g., routing, resource allocation, quality of service (QoS) support, etc.) presents many issues and challenges.

A fundamental issue in such multihop wireless environments is that network performance can degrade rapidly as the number of hops increases. Major problems to transmit data over available radio channels exist in every layer of the protocol stack. In one hand, issues like adaptive rate selection, adaptive antenna pattern, and adjust power control are addressed at the physical layer. In the other hand, link reliability, admission control, and access control to the shared channel are some issues of both routing and MAC layers. Moreover, there are several real-time application requirements that have to be guaranteed in order to provide QoS support and achieve service differentiation.

Obviously, all new protocols and mechanisms proposed for MANETs, including those presented in this thesis, have to take into account the basic inherent characteristics of these networks which are:

- **Dynamic topology**: Nodes are free to move about arbitrarily. In addition, radio propagation conditions change rapidly over time. Thus, the network topology may change randomly and rapidly over unpredictable times. The main design objective of our proposals is to efficiently operate for variant nodes movement patterns.

- **Bandwidth constraints and variable link capacity**: Wireless links have significantly lower capacity than wired links. Due to the effects of multiple access, multipath fading, noise, and signal interference, the capacity of a wireless link can be degraded over time and the effective throughput may be less than the radio's maximum transmission capacity. As we tackle the problem of QoS provisioning, we are aware that this characteristic shall to be taken into account in the design process of all our proposals.

- **Energy constrained nodes**: Mobile nodes rely on batteries for proper operation. Since an ad hoc network consists of several nodes, depletion of batteries in these nodes will have a great influence on overall network performance. Therefore, one of the most important protocol design factors is related to device energy conservation which representing an important design goal and evaluation metric of the different mechanisms described in this thesis.

- **Multi-hop communications**: Due to signal propagation characteristics of wireless transceivers, ad hoc networks require the support of multihop communications; that is, mobile nodes that cannot reach the destination node directly will need to relay their messages through other nodes. The routing schemes proposed here are used to establish multi-hop routes between each pair of nodes.

- **Limited security**: Mobile wireless networks are generally more vulnerable to security threats than wired networks. The increased possibility of eaves dropping, spoofing, and denial-of-service (DoS) attacks should be carefully considered when an ad hoc wireless network system is designed. This issue is out of the scope of our work discussed in this dissertation.

To support mobile computing in ad hoc wireless networks, a mobile host must be able to communicate with other mobile hosts, which may not lie within its radio transmission range. To support ad hoc mobile communications under the influence of the above-mentioned factors, an ad hoc routing protocol will need to perform four functions, namely: (1) Determining/detecting changing network topology, (2) Maintaining network topology/connectivity, (3) Scheduling of packet transmission and channel assignment, and (4) Routing.

Though we mainly focus, in this dissertation, on mobile ad hoc networks (MANETs), we also consider infrastructure mode of 802.11 especially when we design solutions for QoS provisioning through the optimisation of the MAC layer.

Current layered design paradigm is inflexible and sub-optimal for wireless networks. A good network planning is required in order to meet the performance expectations especially when IEEE 802.11 is used with real-time applications. Indeed, multimedia processing and transmission are delay sensitive that require considerable battery power as well as network bandwidth. Furthermore, the routing, mac, and physical protocols that support QoS must be adaptive and cooperative to cope with the time-varying topology and time-varying network resources.

In the past, a lot of research have been conducted to address the MANET-related issues separately. One new research direction to optimize data transfer in ad hoc networks is the cross-layer architecture. It does not respect the original layered approach in which each layer operates independently. The layered approach is simple, flexible, and scalable as the case in the Internet, but it led to poor performance in ad hoc network even with the optimization applied to the evolved protocols because they do not take into account network and application constraints.

Nowadays, the **cross-layer design** approach, which is the general topic of our work, is the most relevant concept in mobile ad-hoc networks which is adopted to solve several open issues. It aims to overcome MANET performance problems by allowing protocols belonging to different layers to cooperate and share network status information while still maintaining separated layers. The central key of related research studies is what information can be shared and how it used in cross-layer architecture to provide QoS enhancement and enable an efficient resource utilization.

This dissertation emphasis on the cross-layer design of 802.11-based wireless networks for QoS provisioning by first presenting a comprehensive design methodology and functionality of each proposed mechanism and protocol and then analysing its performance from a set of simulation tests using several communication scenarios as well as analytical analysis.

As **802.11e** is the emerging MAC layer for QoS support in 802.11-based wireless networks, most of our contributions assume the use of this protocol at the MAC layer. Additionally, due to the large number of cross-layer cooperation possibilities, we emphasize, in this work, on the **cooperation between the MAC and the network layers**. We believe that the cooperation between these two layers provides better performance improvement than the cooperation between other layers. Indeed, they run the two main communication protocols for routing and medium access control. Basically, we design new cross-layer schemes to enhance packets routing in MANETs with QoS considerations thanks to useful metrics measured at the MAC layer. Furthermore, we decided to incorporate our proposals into the **AODV routing protocol** as an example for on-demand routing protocols. We also show that our schemes are generic enough to be included in other routing protocols such as OLSR.

The following section summarizes the key contributions of this thesis.

1.1 Thesis contributions

This dissertation focuses on the design, implementation, and evaluation of 802.11-based cross-layer mechanisms for the enhancement of the support of the QoS feature. We identify the challenges that face the cross-layer architectures comparing to the traditional layered architecture for enhancing communication protocols. In particular, we focus on the problem of routing data packets in a way that takes into account channel contention level, network characteristics, and higher-layer protocol requirements. We address the optimal routing with regard to links stability, average end-to-end delay, and energy conservation with and without assistance/initiation from the network. We design several cross-layer mechanisms that aim to overcome the issue of routing in MANETs while enhancing important QoS metrics (path stability, energy consumption, end-to-end delay, etc.). To this end, we extract the adequate parameters from both MAC and network layers and adapt them to provide QoS enhancement based on new inter-layer cooperation algorithms. Before the cross-layer study, we have also explored the idea of enhancing separately a communications layer; namely the existing 802.11e MAC protocol [33] which is designed for the QoS support. Although the improvements achieved, this study shows the limitations of the layered architecture that demonstrated its good performance in wired networks.

Regarding the several issues addressed in this dissertation, we were able to make a number of contributions, which we summarize here and discussed in more detail in the following chapters:

- **Analysis and optimization of the Enhanced Distributed Contention Access (EDCA) scheme:** EDCA is a contention-based HCF channel access specified in IEEE 802.11e standard. The proposed scheme provides capability for up to four types of traffic classes. It assigns a short $CWmin$, $CWmax$, and $AIFS$ to classes that should have higher priority in order to ensure that in most cases, higher-priority classes will be able to transmit before the lower-priority ones. To decrease delay, jitter, and achieve higher medium utilization, packet bursting is proposed in IEEE 802.11e standard. So, once a station has gained access to the medium, it can be allowed to send more than one frame without contending for the medium again. After getting access to the channel, the station is allowed to send as many frames it wishes as long as the total access time does not exceed a certain limit ($TXOPLimit$) and no collision occurs. There is no need to use RTS/CTS frames for the transmitted packets. The $TXOPLimit$ parameter is fixed and could not react to traffic load variation and medium utilization. We propose to adapt the EDCA parameters in order to take into account the traffic load while decreasing the contention rate by reducing the number of forwarding nodes in MANETs. In particular, we interest in enhancing the dynamic tuning of the $TXOPLimit$ parameter after a mathematical analysis of the performance of the basic EDCA scheme. The proposed ATXOP (Adaptive Transmission Opportunity) scheme leads to a good tradeoff between complexity and performance in terms of several metrics including overall goodput and average end-to-end delay.

- **An energy consumption rate-based cross-layer routing mechanism for MANETs:** we propose a new approach that aims to incorporate energy-related metrics in the decision of determining the optimal route between each pair of wireless devices. We propose a new framework to compute a novel metric called energy-consumption rate which reflects how fast a node is consuming its remaining energy. This metric takes into account by nature the traffic load in the node and its contribution on the data forwarding process in the network. We also propose the required modifications of the AODV (Adhoc On-demand Distance Vector) routing protocol in order to make it energy-aware by considering the metric we design. The extensions of OLSR (Optimized Link Source Routing) is also described. As the optimal path is decided at the source side and intermediate nodes help only on providing the updated measurement of the energy metric, this scheme can be classified as *source-initiated and network-assisted* technique.

- **A cross-layer data forwarding approach for MANETs:** in mobile environments a node is not aware of the mobility patterns of other nodes in the network. Hence, allowing intermediate nodes to contribute on dynamically optimizing the path for each pair of source and destination may be of an interest for several user applications as the decision may not only taken by the source but also by each intermediate node. We introduce a new proposal called CLFS (Cross-Layer Forwarding Scheme) to optimize the data forwarding in the network thanks to the assistance of intermediate nodes. The selection of the forwarding node is based on comparing the energy consumption (and so the expected residual lifetime) of all candidate nodes toward the destination. After selecting the "best" next-hop, a node has to effectively tune the parameters of corresponding EDCA Traffic Categories in a cross-layer manner. The routing and the MAC modules cooperate to optimize the data forwarding. Contrary to the above contribution, CLFS is a *source-assisted and network-initiated* approach.

- **A stability-based cross-layer routing mechanism for MANETs:** selecting the stable path is a major challenge in MANETs. However, traditional routing protocols do not take into account the movement patterns of nodes and they are not able to predict how the neighborhood is going to change in the near future. Hence, we tackle this problem by developing a new "cross-layer metric" for measuring the stability of links in MANETs. This metric is updated based on the measurements done in network and MAC layers. We develop a distributed algorithm allowing to compute this metric and maintaining it up-to-date in each wireless node in the network. Incorporating this metric on routing protocols such as AODV and OLSR allowing to optimize the selecting of nodes composing a path between each pair of nodes. We show through ns-2 simulations how these stability-aware routing protocols outperform the basic ones for several performance parameters and different scenarios.

- **A cross-layer routing mechanism for delay-sensitive applications:** real-time applications generally tolerate packet loss but they are sensitive for delay and delay variation. To encourage the development and deployment of such applications in MANETs, we consider also the delay as an example of additive metrics to optimize when establishing a route between two nodes in the network. Similarly to the two previous contributions, we also propose a scheme to dynamically and efficiently estimate the average end-to-end delay along each candidate path by tacking into account the observed MAC access delay in each intermediate node and the size of each MAC queue. Again, we simulate the proposal using ns-2 and we demonstrate that it largely outperforms the basic routing protocol where the route selection procedure consists only on selecting the shortest path in term of the number of hops.

- **XAid (CrossAid) - a cross-layer architecture for 802.11-based MANETs:** having proposing several cross-layer approaches for QoS-based routing in MANETs, we designed a new architecture called XAid (CrossAid) that incorporates all our proposals. We detail how each cross-layer approach is selected based on the requirements of the user application as well as the measurement gathered from the network. Hence, we provide a mapping between the set of QoS requirements and one of the four protocols we proposed.

Briefly, during this Ph.D., we focused on the cross-layer design of 802.11-based wireless networks in order to provide an efficient service differentiation. We believe that the set of protocols and mechanisms that we have designed and extensively evaluated will encourage the deployment of MANETs which efficiently supporting the QoS in their networks in order to excite the development of a wide range of new user applications.

1.2 Dissertation outline

The remainder of this dissertation is organized as follows.

In Chapter 2, we present an overview of the existing works that are related to our field of study. We provide the trends and challenges of designing new cross-layer protocols for wireless networks. We also enumerate the limitations of previous approaches and determine the position our research work regarding to the state of the art in this field.

In Chapter 3, we provide **a novel analytical analysis** of the EDCA 802.11e protocol and we derive some guidelines for the optimal tuning of 802.11e parameters mainly the transmission opportunity parameter for single hop and multihop networks. The proposed Adaptive Transmission Opportunity (**ATXOP**) scheme is described. We show through ns-2 simulations how the performance of the EDCA function can be improved using the ATXOP mechanism.

In Chapter 4, we describe an energy consumption rate-based routing approach, called **E-AODV** (Energy-AODV), and provide simulations results and analysis. The required extensions to the AODV and OLSR routing protocols is also subject of this chapter.

In Chapter 5, we tackle the problem of enhancing the data forwarding in MANETs. We describe the CLFS (Cross-Layer Forwarding Scheme) and we analysis the obtained results through ns-2 simulations of the new routing protocol **F-AODV** obtained by incorporating CLFS in the AODV routing protocol. We evaluate the performance of F-AODV under different network scenarios.

In Chapter 6, we focus on developing a new stability metric that evaluates the lifetime of links in MANETs. We describe a cross-layer stability-based routing protocol as an extension for

AODV, called **S-AODV**. We demonstrate using ns-2 simulations how S-AODV dramatically outperforms the basic AODV protocol for several performances metrics.

In Chapter 7, the delay is considered as the main metric for establishing routes between nodes in MANETs. We detail the approach for estimating the end-to-end delay and how it is incorporated in the AODV protocol (the resulting protocol is called **D-AODV**). Thanks to ns-2 simulations, we show that D-AODV is very useful for delay-sensitive applications.

In Chapter 8, we address the tradeoff between the improvements achieved thanks to our proposals and the complexity to implement them and to exchange useful information between wireless devices. We also expose **XAid (CrossAid)**, a novel framework for cross-layer QoS provisioning in multihop wireless networks. It includes all proposals described from Chapter 4 to Chapter 7. We provide a guideline for mapping user applications to the "optimal" cross-layer routing scheme according to their QoS requirements.

In Chapter 9, we present our conclusions and we outline potential future works.

Chapter 2

Cross-layer design of wireless networks: proposals, trends and challenges

A wireless ad hoc network is a self-organizing collection of user nodes that must cooperate in order to provide basic networking functionality. In general, the nodes of an adhoc network are mobile and rely entirely on wireless transmission without any fixed infrastructure or dedicated communications devices. Consequently, packet-switched routing is required to manage limited device power, unpredictable variation in channel quality, and to reduce media access contention. Hence, an ad hoc network effectively consists of a set of mobile wireless routers. As such, the user nodes must participate in an adaptive routing algorithm that is responsible enough to meet application requirements without over-utilizing limited resources.

Applications for ad hoc networks range from rapidly deployable networks for military and civil operations, to networks of intelligent sensor devices. Ad hoc networks may also be utilized commercially to increase the capacity range and quality-of-service of infrastructured wireless networks, and to support intelligent highway systems. To achieve their full potential, however, many challenges remain unaddressed or incompletely resolved, namely, scalability with respect to size, traffic load, and mobility, power efficiency and control, efficient multicast routing, improved transport-layer effectiveness, cross-layer interaction and optimization, security and service availability.

Many subsystems of appliance operating systems are implemented as stacked modules. For example, the TCP/IP subsystem consists of the link layer, the network layer (IP), the transport layer (TCP and UDP) and the application layer organized as a protocol stack. One new research direction to optimize data transfer in wireless ad hoc networks is the cross-layer design without respecting the original layered design approach in which each layer operates independently. The central theme of this dissertation focuses on the cross-layer design of mobile wireless networks for QoS provisioning. We aim to develop a set of mechanisms and proposals operating between communication layers to enhance the service differentiation in 802.11-based MANETs. Specifically, this thesis investigates the improvement of supporting the QoS at MAC and network layers.

Before embarking into the details of the different contributions, we devote this chapter for reviewing proposals, trends and challenges on designing cross-layer communication protocols for mobile wireless networks. Indeed, lot of research works have been presented in the open literature that introduce several coupling ways and solutions between different communication layers [40, 41, 69, 57].

The remainder of this chapter is organised as follows. In Section 2.1, we enumerate the limitations of the layered architecture on providing an acceptable efficiency in wireless networks. For each layer, going from the physical layer to the application layer, we provide the restrictions which make it not able to accommodate its services with similar accuracy as in wired networks. We also identify the most important parameters in each layer to be managed in a cross-layer architecture. We review the most works that have been conducted to study the cross-layer design in mobile ad hoc networks in Section 2.2. In Section 2.3, we outline our observations on the limitations of cross-layer mechanisms proposed to enhance the support of the QoS in MANETs and enumerate the motivations behind the contributions presented here as well as the design goals of the different designed mechanisms. We conclude the chapter in Section 2.4.

2.1 Why the layered architecture is inefficient in mobile wireless networks?

A Mobile Ad-hoc NETwork (MANET) is a set of wireless mobile nodes dynamically forming a temporary communication network. The goal of this architecture is to provide communication facilities between end-users without any centralized infrastructure. In a such network, each mobile node operates not only as a host but also as a router. It is also possible to have an access to some hosts in a fixed infrastructure depending on the kind of the available mobile ad hoc network.

As it is well known, networks are organized as a series of layers, each one built upon the one below it. The goal of this architecture is to split the network into smaller modules with different functionalities and deal with more manageable design and implementation. The purpose of each layer is to offer certain services to the higher layers, shielding these layers from the details of how the services are implemented. So the advantage behind the layered protocol architecture is to reduce complexity by dividing and conquering approach. This simplicity ensure an easy way to standardize, and to deploy new flexible protocols (easy upgradeable). However, ad hoc networks have several characteristics that reduce the simplicity and the flexibility of the layered architecture. Indeed, wireless networks do not come with links and the wireless channel quality changes dynamically. Moreover, in contrast to the Internet, routes are unstable and may change frequently as nodes move in an unpredictable manner. Additionally, the applications require a minimum of QoS that could not be achieved in such very dynamic capacity networks. A fundamental issue in such multihop wireless environments is that network performance can degrade rapidly as the number of communications increases. Major problems to transmit data over available radio channels exist in every layer of the protocol stack. In one hand, adaptive rate selection, adaptive antenna pattern, adjust power control are issues of the physical layer. In the other hand, the link reliability, the admission control, and the access control to the shared channel are some issues of both routing and MAC layers. Furthermore, there are several real-time application requirements that have to be respected in order to provide QoS support and achieve the expected service differentiation.

One of the most important design goals for wireless ad hoc networks is to reduce the energy consumption due to the operations of the distributed communication protocols, which must be met in order to allow networks to operate unattended for extended periods of time.

Indeed, there are some scenarios, as emergency disaster relief personnel coordinating efforts after a natural disaster such as a hurricane, earthquake or flooding, where recharging battery is very difficult (hard) to do in such conditions. Hence, the network connectivity is strictly related to the possibility of routing between each node in the network. The goal behind the routing process is to minimize the energy consumption while maintaining the existing of routes between nodes. Moreover, energy conservation is used in power control mechanism by reducing transmission range in order to decrease the contention probability while allowing the control of the network topology [23, 67, 74]. The energy exhaustion issue leads to network disconnection and resource unavailability problems.

Hereafter, we briefly present the limitations of the layered architecture in wireless networks and we emphasize on the impact of MANET environment on the QoS support.

2.1.1 Limitations related to the physical layer

The wireless channel varies over time and space and has short-term (or small-scale) memory due to multipath. The channel variation meets the amount of contention, time-varying fading, multi-path, variation of the SNR. Indeed, these variations are caused either due to motion of the wireless device, or due to changes in the surrounding physical environment, and lead to detector errors. This causes bursts of errors to occur during which packets cannot be successfully transmitted on the link. Fast channel variations due to fading are such that states of different channels can asynchronously switch from good to bad within a few milliseconds and vice-versa. Furthermore, very strong forward error correction codes (i.e. very low rates) cannot be used to eliminate errors because this technique leads to reduced spectral efficiency. In addition to small-scale channel variations, there is also spatio-temporal variations on a much greater time-scale. Large-scale channel variation means that the average channel state condition depends on user locations and interference levels. Thus, due to small-scale and large-scale changes in the channel, some users may inherently demand more channel access time than others based on their location or mobile velocity, even if their data rate requirement is the same as or even less than other users. The techniques that may be used to adapt to rapid SNR changes in wireless links and mobility include: power control, multiuser detection, directional antenna, adaptive modulation and software radio. However, sharing these information with high layers, has a big benefits on performance as shown in [23, 15, 41]. For example, characterizing the application requirements help to use the adaptive modulation, the knowledge of channel quality help to avoid useless MAC retransmission..etc.

In the layered architecture it is not possible for upper layers to ask the physical layer for the wireless channel characteristics described above and to adapt their operations accordingly.

2.1.2 Limitations related to the MAC layer

Due to the high difference between transmitted and received power levels, traditional random channel access mechanisms used in wired networks as CSMA/CD are not applicable in wireless networks. To deal with this problem, contention-based random-access multiple access protocols have been commonly used since they are simple to implement. To further increase the efficiency of the operation, carrier sense based MAC algorithms are used, requiring the mobile terminal to first sense the channel to determine that it is idle and only then attempt its packet transmission. The latter attempts can still results in a collision event (when the intended receiver detects multiple transmissions at such power levels that it may not be able to correctly receive any of them). CSMA-based MAC protocols can yield an efficient operation (under proper loading levels) when the carrier sensing operation is spatially effective. Unfortunately, stations may be geographically located in a manner that induced blocking, leading to masked terminal scenarios. In this case, two major problems have been identified: hidden terminal and exposed terminal conditions. Despite of introducing RTS/CTS handshaking scheme, leading to the Multiple Access Collision Avoidance (MACA) protocol [38], the MAC layer still suffers from the problems of interference resolution, exposed terminal, efficient medium utilization. Indeed, the optimal strategies of resource sharing issue among different classes of users, still the main challenge also for the FDMA, TDMA techniques.

The MAC layer can be optimized if it could gather information from the physical layer. Additionally, it may also be tuned optimally if it is aware about the operations of upper layers as well as the QoS requirements of user applications. These two features cannot be supported by the layered architecture.

2.1.3 Limitations related to the routing layer

The functions of the network layer are to provide (IP) addresses to end hosts, and set up routes between sources and destinations. Routing protocols for ad hoc networks require to consider the reasons for link failure to improve its performance. Link failure stems from node mobility and lack of network resources. Therefore, it is essential to capture the aforesaid characteristics to identify the quality of links. Furthermore, the routing protocols that support QoS must be adaptive to cope with the time-varying topology and time-varying network resources. For instance, it is possible that a route that was earlier found to meet certain QoS requirements no longer does so due to the dynamic nature of the topology. In such a case, it is important that the network intelligently adapts the session to its new and changed conditions. So, the goal of QoS routing is to optimize the network resource utilization while satisfying application requirements. Indeed, it is not enough to find a shortest path but also with available resources as battery, bandwidth, and buffer. Note that the factors that can change the topology of an ad-hoc network are: the mobility of nodes, change of power, the MAC layer mechanism because different schedule for the contending nodes, results in different topology, the flow dynamics that flows come and go; if a node has nothing to transmit, its links are gone from the topology, and finally the mode of nodes: sleeping or active mode. If a node goes to a sleeping mode, its links are gone from the topology and hence it can't participate to route establishment and communication.

The network layer requires to identify the different routes parameters and share it with the other neighbors. This helps to use efficiently the links and establish paths with an economic manner that takes into account the changes in the network topology and resources. However, efficiency and fast convergence rate are two conflicting objectives. It is hard to achieve the trade off between communication overhead and computation effort. The more routing information distributed, the less computation required at each node. So, there are three main functions of the network layer: provide (IP) addresses to end hosts, set up routes between sources and destinations, pro-actively (routes ready-to-use) or reactively (routes on-demand). To set up a route, we need route discovery; to make routes ready to use, we need route maintenance. Within these functions, there are several objectives that have to be achieved: the efficiency that consists in minimizing signal overhead in route discovery and route maintenance and minimizing convergence time. Furthermore, providing routes that support requested QoS is very important. Then, make sure if the protocol is scalable that's mean whether the network is able to provide an acceptable level of service to data packets even in the presence of a large number of nodes in the network. Finally, introduce energy efficiency in route establishment.

Obviously, the routing component is one of the major modules to handle data forwarding in MANETs. A layered architecture can only offer the basic routing functionality with the minimum guarantee for routes establishment. Including information from lower and upper layers can lead to a non-negligible optimization of the selection of the optimal paths each pair of nodes.

2.1.4 Limitations related to the transport layer

TCP combines error control (ARQ), flow control that are not over-running the receiver buffer, and congestion control that is not clogging the network, and not overloading the capacity in the routers. Moreover, TCP enjoys simplicity of control and gains widest acceptance. However, this simplicity of control is at the cost of efficiency loss. TCP is not able to distinguish the presence of congestion in wired networks, mobility, collision in wireless links, and bit errors due to poor quality of wireless links. Single bit error could trigger congestion control mode (TCP getting into slow start phase); even fast retransmit/fast recovery is not effective in coping with packet/bit errors. So, TCP needs to handle delay (RTT) and packet loss statistics that are very different from those in wired networks.

To overcome the issues related to the TCP operations in MANETs, it is mandatory to have an knowledge about the wireless channel conditions. Getting also details about the MAC layer such as the number of retransmissions and the observed frame loss rate may also be used by TCP to efficiently tune its retransmission timeout parameter.

2.1.5 Limitations related to the application layer

There are some application's requirements that should be considered in order to maintain as good as possible the performance and offer a minimum service delivery according to their constraints. Indeed, there are time-bounded applications that are sensitive to delay and others require high throughput and/or less packet loss rate. For example audio traffics should reach destinations at most up to 400 sec. The corresponding packets are almost short. They could have the highest priority: minimum waiting time in the queue, and so short medium access time (e.g. short contention window size). Moreover, they require short and less congested routes to reach destinations within a short delay. Throughput-constrained applications require less congested routes and available queue to enqueue packets. Hence, successful transmission should be assured and they are less sensitive to delay comparing to above described class. TCP traffics are very sensitive to both packet loss and delay. Background traffic should not be starved and so a minimum service has to be guaranteed.

For the past few years, significant work has been done to find efficient wireless communication techniques, but most of the work has concentrated on optimizing a single layer in the protocol stack. Recent results indicate that local optimization of all layers may not lead to global optimization. The key question is how to adapt physical layer parameters, distribute fairly the access to the medium, find the best routing path with the better QoS, and achieve an efficient bandwidth sharing while providing service differentiation and application requirements with the less possible complexity?

In the next section, we describe how these problems and required information, related to each layer, are exchanged over the different protocols in the layered stack in order to address cross-layer optimization not only for QoS provisioning but also for other objectives. Hence, we discuss some of the cross-layer approaches that have been proposed in the open literature.

2.2 Review and analysis of cross-layer proposals

Each layer of a stacked set of modules maintains an independent set of statistics for error conditions and performance metrics. When a problem occurs, it may manifest itself as aberrant statistic values in multiple layers in the system. In classical systems, there is no logic that correlates these aberrant statistic values across different system layers. This lets thinking about alternative solution as cross-layer design. The main feature of the proposed studies in the literature is the determination of what information could be shared and how is it used in a cross-layer architecture to provide QoS enhancement and enable an efficient resource utilization? Hereafter, we describe some examples of cross-layer integration for ad-hoc networks for several objectives and between different layers.

2.2.1 Cooperation between physical and MAC layers

2.2.1.1 Making the modulation adaptive

A cross layer networking system is described in [75]. The authors propose a coordination between routing, MAC, and physical layers. Indeed, the scheme consists of considering three signal strength attenuation factors, namely, path loss, shadowing and multi-path fading. The authors suggest that for channel-adaptive protocols, a good time-varying channel model is needed for simulation. So, a correlated shadowing channel model is proposed. At MAC layer, a rate adaptation scheme is described. The RTS, CTS and ACK packets are sent at nominal rate. When a node receives the RTS packet, it estimates the SNR. Then, the transmission rate is mapped from the estimated SNR, and appended to the CTS packet. So, the sender transmits data at the adapted rate. An M-QAM scheme is used in which the constellation size changed with SNR. At routing layer, the source node considers the MAC delay of every RREP packets and chooses the route with the minimum delay. The RREP packets are unicast packets to the source node using rate adaptation based on the SNR information along the route. However, no rate adaptation is us‘ed in RREQ packets. The routing decision is made based on three metrics. The first one is the bandwidth that represents the rate of link between node i and j. The second one is the interference duration that is the interval from the time when the RTS packet is sent to the time when the data packet is received. The third one, is the congestion that is the queuing delay in the buffer of the transmitting node.

2.2.1.2 Making the power control adaptive

There are several works that integrate adaptive modulation, MAC functions, and routing metrics to introduce cross-layer cooperation. Note that, adaptive modulation needs channel estimation: When channel is good (high SNR) that means higher order modulation and so, higher rate. The metric of channel estimation link gain is used in [23]. The proposed solution is based on MAC and physical layer cooperation. It estimates the channel using RTS packet and transmit the information using the CTS packet. Then, an adaptive power control mechanism is described according to the obtained information from MAC layer. The used system model considers n nodes in ad-hoc network. All nodes share the same frequency bands: TDMA or TDMA/CDMA, and one hop transmission (no routing) is considered. The interference is reduced by scheduling and power control. In the first step, the MAC layer scheme has to determine the optimum set of scheduled users. This decision changes according to the propriety of MAC layer: The simple TDMA eliminates self interference half duplex transmissions. This criteria is the same for TDMA/CDMA. Moreover, a node cannot receive from more than one neighbor simultaneously. Furthermore, a receiving node should be spatially separated from any other transmitter by at least a distance D that leads to spatial separation. Note that this later criteria doesn't mandatory for TDMA/CDMA. So, the D parameter greatly influences the amount of interference suppression: If we have low D, more users are selected in the valid set, but a lot of interference must be managed by the power control. The drawback is that power control may not be feasible. If we have high D, there are less users scheduled and so easier job for the power control. However, the scheme may be too conservative higher delays resulted from scheduling because of only these users can run the power control algorithm. In the second step, the power control mechanism optimizes the power allocation among different users. If we have few scheduled users, the MAC layer does not re-optimize its selection based on information from the physical layer, the loop is not closing. Concluding this work, we mention that the main challenge is to select the optimum valid subset of users that gives the maximum number of users in a valid configuration. Indeed, the optimal solution presents an exponential complexity in the number of users that leads to a combinatorial optimization problem. The suboptimal solution selects the users sequentially and decides if they can be added or not to the valid subset. This solution could lead to deferring more transmissions than needed. Moreover, we have the same problem for the optimum admissible set: If power control is not convergent how to determine the optimal subset of users that leads to convergence? This is an exponential complexity NP hard problem. The suboptimal solution is to defer transmission for the user having the minimum SIR (or SINR signal to interference + noise ratio).

2.2.2 Cooperation between physical, MAC, and network layers

2.2.2.1 Exploiting adaptive beamforming for routing enhancement

The above discussed research works were not specifically consider routing. In general, MAC protocols differ based on: How RTS/CTS, is transmitted (omni, directional), transmission range of directional antennas, channel access schemes, and omni or directional NAVs. Furthermore, the antenna gains are different for omnidirectional (Go) and directional transmission (Gd): Gd > Go. Moreover, if an idle node listens omnidirectionally, it does not know who is going to transmit to it. The directional antennas provide a good spatial reuse, higher gains and better links. However, higher gains mean also high interference at distanced nodes. There are three types of links. The first type called Omnidirectional Omnidirectional (OO) which is characterized by smallest range. The second is Directional Omnidirectional (DO) links usually used in the protocols discussed up to now because the node listens omnidirectionally. The third type is Directional Directional (DD) which has the largest range and so the least number of hops. However, the problem is that the nodes listen omnidirectionally. In [15], the proposed mechanism describes a cooperation between adaptive beamforming, MAC, and routing. The scheme uses the same MAC for directional antennas, but transmits RTS over multiple hops (MMAC protocol). If source 1 (see Figure 2.1) wants to communicate with node 6. It transmits a forwarding RTS with the profile of node 6, using DO links. Then, when node 6 gets the RTS, it beamforms in the direction of 1, forming a DD link. Moreover, the transmission from

1 to 9 on DD links requires only 2 hops. The presented network performance results depend on the simulated network topology. There are several cases that were studied: Manhattan networks with aligned routes, Manhattan networks with random routes, and Random configuration. For all three cases, the numerical parameters chosen are: Antenna beamwidth equal to 45 Omnidirectional, transmission range equal to 250 m, Directional transmission range (DD link) equal to 900 m. The performance measure only average throughput. It shows that in general MMAC, better than DMAC, better than 802.11. However, when the routes are aligned, using MAC and directional antennas degrades the performance, compared to the case with omnidirectional antennas (802.11). For Manhattan networks, more directional interference occurs, due to the aligned paths. The gain is better if we can actually exploit the spatial reuse property of the directional antennas. If not, the performance will be worse because of the increased directional interference (higher gain for the directional antennas).

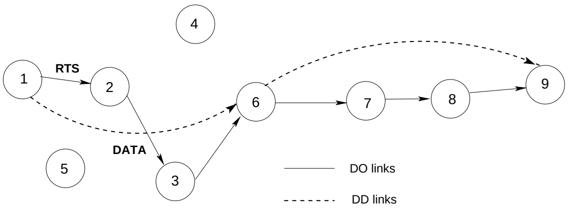

Figure 2.1: Example for cross-layer scheme work

2.2.2.2 Using physical power control and enhanced MAC scheduling for routing optimization

The cross-layer approach introduced in [69], is presented at four levels: First, the proposed adaptive MAC protocol is sensitive to contention. Second, influence of network layer FIFO queuing on better bandwidth utilization. Third, importance of transmission scheduling. Fourth, routing and power control interactions.

Firstly, the authors study the effects of queue management, routing protocol, power control, and medium access mechanism on the capacity of the networks. Based on a novel frame format based loosely on the CSMA/CA protocol and TDMA, a contention slot is splitted into m mini-slot pairs (slotting is done at a much finer level). These contention intervals finally result in a *(src,dst)* pair agreeing to exchange data during the data slot portion of the frame. The duration of a mini-slot pair depends on several parameters, but as described in the paper it is equal to 80 micro-seconds. In the proposed Progressive Back Algorithm (PBOA) Protocol, nodes contend during every contention period. Unsuccessful nodes progressively backoff during progression of contention period. Successful nodes use remaining contention interval to discover minimum power needed to transmit their data. There are two benefits of this approach: the first that energy conservation is enhanced because of tuning transmission range as possible. The second benefit is that both interference and collisions are reduced thanks to the proposed backoff procedure. The PBOA performance is much better than CSMA/CA in throughput, power. But, in PBOA nodes can still waste contention time by transmitting an RTS to a node who has already sent a CTS to another originator.

The PBOA's authors proposed a second algorithm called Progressive Ramp Algorithm (PRUA). In PRUA, nodes monitor the channel for favorable conditions and send an RTS with some probability p at the beginning of the next RTS minislot. If the transmission succeeds, the node continues to send RTS packets to notify others to backoff, otherwise it backs off and tries again at a later RTS minislot. No transmitter power control mechanism was used. The benefits of this approach is that it does not obey to FIFO queuing, so that the best reachable destination has a high chance to receive its packets. PRUA employees carrier sense but it is tuned to detecting CTS and hence nodes will avoid extra RTS transmissions and unnecessary interference and collisions experienced by PBOA. The simulation results show that PRUA has better uniform capacity than PBOA as well as delay and throughput, but lack of power control costs more energy.

The performance of these interactions depends on several constraints. Indeed, The cost of packet control overhead and packet lost could be more significant than the performance improvement in an arbitrary mobile node in a particular scenario. The key observation is that the protocol performance looks worse than some optimal choices because these two protocols are distributed and hence require global knowledge to schedule their transmissions which is hard to achieve in a very high mobile and distributed networks.

2.2.3 Cooperation between MAC and network layers

2.2.3.1 Using MAC utilization and interface queue information for reactive routing protocols optimization

In [29], the authors propose a mechanism for detecting network congestion, in order to improve the performance of all types of traffic. Indeed, there are two metrics which are used to measure the congestion level. The first one, is the average MAC layer utilization around each node. Instantaneously, this metric can be equal to 1 or 0. It is equal to 1 if the MAC layer is utilized (there is at least one packet in the transmission queue, during backoff decrease period, inter-frame space, detection of physical carrier). The second metric, is the instantaneous interface queue length. This metric is used to avoid nodes that are congested even there is no contention. The proposed mechanisms aim to influence routing decisions that will follow other route discovery scheme either than the short hops count used traditionally. Indeed, it is unsuitable to establish routes over nodes that are already busy. However, if we avoid busy nodes in route establishment, there are some routes that cannot be established even they exist. The congestion information is also used when the medium utilization is high, to influence the setting of the Explicit Congestion Notification (ECN) bits in the IP header of packets at each node. ECN is used to prevent the loss of packets along that flow. At transport layer, the MAC layer utilization metrics measured around the node allow TCP sender to tune its parameters according to these metrics since they represent a recent level to the wireless medium utilization. At higher layer, these metrics can be used to decide or not data compression. Indeed, when the medium is busy the sender can decide to compress the data. However, the compression should represent a trade off between bandwidth consumption and the CPU time used for compression and decompression.

2.2.3.2 Using topology information and enhanced back-off for proactive routing protocols optimization

In [17], a detailed discussion of cross-layering design has been presented in the context of research project called MobileMAN. This project investigates a local interaction among protocols in a manet node. For example, MAC layer exploits the topology information collected by network layer to achieve fair channel scheduling and fix the problem related to hidden and exposed terminals. An enhanced backoff scheme is introduced. At transport layer, the different events occurring at lower layers such as route failure, route changes and congestion, are analyzed in order to minimize the useless data retransmissions. Moreover, MobileMan considers routing according to the cross-layering principal. Indeed, a path per-formability index is computed using congestion, link quality, and other parameters that can influence system performance. Furthermore, the MobileMan transport protocol exploits information reported by the routing and Wi-Fi layers in the Network Status component to avoid useless data retransmission.The authors suppose that a node has a knowledge of the hole network topology and so a proactive routing protocol should be used. Hence, it seems that for some scenarios, it is very hard, costly, and not efficient to address this cross layer architecture regarding the dynamic traffic nature and the high mobile node speed. Any information has been provided to how compute the path per-formability index or other cross layer parameters considered in this project.

2.2.3.3 Using probability of successful transmission and energy conservation for route selection optimization

In [57], tow cross-layer designs based on energy consumption are presented for wireless ad hoc and sensor network. The proposed schemes, namely, Energy-Constrained Path Selection (ECPS) and Energy-Efficient Load Assignment (E2LA), employ probabilistic dynamic programming (PDP) techniques and utilize cross-layer interactions between the network and MAC layers. They aim to enhance the operation of existing power-based multi-path routing protocols via cross-layer designs and optimal load assignments. The Energy-Constrained Path Selection (ECPS) consists of maximizing the probability of successful transmission in at most n retry. That is mean that the total n transmissions do not exceed a total amount of energy equal to γ. Furthermore, the authors developed four distinct reward schemes for which E2LA assigns routing loads accordingly. This objective is achieved by applying PDP techniques and assigning a unit reward if the favorable event (in this case, reaching the destination in n, or less transmissions) occurs, and assigns no reward otherwise. Hence, it can be shown that maximizing the expected reward is equivalent to maximizing the probability that the packet reaches the destination in at most n transmissions. In ECPS mechanism, the MAC sublayer provides the network layer with information pertaining to successfully receiving a CTS or an ACK frame, or failure to receive one. ECPS, in turn, chooses the route that will minimize the probability of error or, equivalently, maximize the expected reward.

The proposed medium access control (MAC)-based performance studies, revealed that battery capacity may not be efficient for achieving energy-based fairness and system longevity for wireless mobile multi-hop ad hoc and sensor networks. However, energy conservation may be attained only if valuable MAC (and PHY) input is passed to the network layer. In addition, illustrative examples of E2LA were presented, and its diverse properties were introduced and validated.

2.2.3.4 Using the topology information for power control enhancement

In [40], a study of cross-layer design based on power conservation, and congestion information in ad hoc network have been presented. The authors describe a power control based cross layer architecture. Indeed, they detail the significant impact of power control on all protocol stack above the physical layer. Furthermore, they summarize several works that have been done to address power saving in the protocol stack and show how the power information could be considered at each layer. Moreover, the work claims that, exchanging the topology information between different layers through their interfaces, is very important to support QoS such as geometric location, channel, link conditions. A proposed mechanism, that uses the number of neighbors around the node to adjust transmission power, has been presented.

2.2.4 Cooperation between physical, MAC, and application layers

2.2.4.1 Using the SNR information and MAC retransmission to make the FEC scheme adaptive

Real-time applications, such as audio and video streaming over wireless links, suffer from bandwidth variation, packet losses, and heterogeneity of the receivers. To overcome, bit error problem, many works have addressed adaptive error-control strategies at the application layer. However, in existing WLAN environments, different protection strategies exist at the various layers of the protocol stack and, hence, a joint cross-layer consideration is desirable in order to provide an optimal overall performance for the transmission of video. In [58], the authors propose to exploit the mechanisms available at the lower layers of the protocol stack in order to address an adaptive cross-layer protection strategies for robust scalable video transmission. This mechanism uses a multipath channel model to simulate the wireless indoor channel. This channel model provides the bit error rate (BER) of the link for the eight different PHY modes of 802.11a under different channel signal-to-noise ratio (SNR) conditions. Then, the authors analytically derive the packet loss ratios and throughput efficiency at various channel conditions, considering a given packet size, a given number of retransmissions at MAC layer, and an application layer FEC. These parameters are dynamically adapted according to an end-to-end distortion model in order to achieve an efficient transmission of video streams. The presented algorithm presents a good performance for video streaming. However, it is only centralized.

In the next section, we discuss the constraints of introducing cross layer architecture and the recommendations to achieve a good and optimized solution.

2.3 Discussion and motivations of our work

The layered approach is simple, flexible, and scalable for several networks as the case in the Internet, but it leads to poor performance in ad hoc wireless network even with the optimization separately applied to the evolved protocols because they are not taking into account network and application constraints. For instance, each layer has to react in route failures and collisions in its own way and there are no coupling of different layer information to meet some parameters in order to address a good coordination of the efforts satisfying as well as possible the application requirements. Another example showing the importance of cross-layer design is when a MAC layer tries several times to transmit a packet to a destination which is out of its transmission range. In this case, if the network layer informs the MAC layer that the destination is unreachable, the useless frames retransmissions could be avoided. As conclusion, the cooperation between layers to enable performance enhancement is very important and useful in wireless ad hoc networks. The global objective of such cooperation is to achieve a reliable communication-on-the-move in highly dynamic environments as well as QoS provisioning.

Cross-layer protocol interactions (also known as interlayer protocol interactions), when used appropriately, can lead to increased network efficiency and better QoS support. In MANETs, cross-layer protocol interactions can be used to improve the routing mechanism or to provide service discovery support. Cross-layer design is particularly important for any network using wireless technologies, since the state of the physical medium can significantly vary over time. Layers can exchange information to make more optimal usage of the network. We are looking at optimizing across layers, not merely within single layers. The advantages of cross layer design is also to exploit the interactions between layers in order to improve QoS support and optimize resource utilization. Moreover, this new architecture promotes adaptability at all layers based on the exchanged information and tight their interdependence. However, understanding and exploiting the interactions between different layers is the core of the cross-layer design concept. For example, the cross layer models introduced in [29, 17] require respectively the congestion information and global topology information to build routes using inter-layer cooperation mechanisms. Hence, if we consider high variable scenario in term of mobility and traffic load, the collected metrics will be inaccurate and so become inefficient and very costly. Indeed, it is hard to characterize and identify the required interactions between protocols at different layers. Moreover, joint optimization across layers may lead to complex algorithms. Note that the complexity criteria consists on consuming more resources for computing required information and introducing scalability problem in terms of the number of nodes and the traffic load. So, we have to answer the following question: is cross-layer design suitable for all types of wireless networks and all types of applications? If yes, that means that we have to throw away the OSI reference model and we do not need to consider a network architecture anymore? This is clearly impractical and disaster in terms of implementation, debugging, upgrading and standardization.

As we have mentioned above, a vast amount of cross-layer adaptations and optimizations have been proposed in the recent past years spanning all layers of popular protocol stacks. Some of these mechanisms consider a centralized architecture or assume that every node has a global knowledge of the network. Most of them do not consider energy consumption, links stability, and end-to-end delay parameters. The actual cross-layer interactions are not the primary focus of most publications and are generally assumed to be somehow available. The sheer amount of cross-layer adaptations suggests though that a sound architectural design would not only greatly reduce the design and implementation complexity of new and existing approaches but it would also clearly be a necessity to have a common architectural framework to support these cross-layer optimizations efficiently. The architectures presented above try to support cross-layer protocols in different ways and they also differ in their application scenarios. Furthermore, they have distinct ability to detect and prevent potential risks that exist when weakening the strict layer separation principle. Therefore, a direct comparison is not always a straightforward process but can be done using criteria that abstract from the architectural details.

In this work, we are interesting in developing new solutions for data communications in a fully distributed wireless environment. We are looking for new algorithms, mechanisms, and protocols by cooperating **MAC and network layers** in order to overcome three main issues in MANET: energy exhaustion, frequent link breaks, and large delay for multimedia applications. Moreover, we emphasize our study on the **emerging 802.1e standard**, which was designed explicitly for QoS support in 802.11-based wireless networks. At the network layer, we decide to select **the AODV routing protocol** in order to demonstrate how our proposals are able to improve the determination of routes between each pair of nodes in the network.

One of the main features of this thesis, is that it considers several optimization metrics which have not been addressed previously and it provides a comparison of our proposals with the basic ones. Furthermore, we analyze their performance under several network and application characteristics. Our study of the obtained results show that although the objective is to provide a gain in terms of packet delivery ratio, average delay, routing overhead, it is possible to not provide a performance enhancement when applying these layer interactions. This is because of some network parameters as node mobility and variation of traffic load that affect the applications end-to-end performance metrics.

In order to understand the importance of interlayer interaction and to figure out what are the limitations of the layered architecture, we start by studying the performance of the new EDCA scheme that is introduced in IEEE 802.11 MAC layer [33] to support service differentiation at the MAC layer. We focus on the EDCA performance in both single hop and multihop network. We discuss how its parameters could be adapted to application and the wireless environment characteristics. Our observations show that this mechanism provides a good QoS support for low loaded networks. However, when the traffic load increases, we have already shown in a previous work that it performs poorly [53]. These results, which are also confirmed here using a new analytic framework, are due to the dynamic characteristics of mobile ad hoc networks and the lack of adaptation of tuning parameters through the inter-layer cooperation in order to adapt to these characteristics. Hence, we are looking at optimizing across MAC and routing layers the performance of data communications in MANETs.

Then, we study how the cooperation between communication layers can enhance the service differentiation. We develop several routing approaches that require information from network and MAC layers. Our proposals introduce several new metrics which are important for mobile wireless networks such as the stability of wireless links between mobile devices, the rate of consuming the energy, and the end-to-end delay based on the MAC layer measurements. We also explain new sophisticated algorithms in order to estimate these metrics in an efficient way.

In order to achieve interlayer protocol interaction, we modified the 802.11 MAC layer and routing protocols so that it provides various information about the status of the wireless channel to other layers. The modified MAC and routing protocols provide several useful measurements which will be detailed in the next chapters. The changes are implemented in the ns-2 simulator. These changes could be eventually ported to any open-source 802.11 Linux driver such as MADWIFI Linux driver of 802.11g cards based on Atheros 5212 chipset.

2.4 Chapter summary

Wireless ad hoc networks are collections of nodes communicating over a wireless channel. Since wireless signal power decays with distance and in the presence of obstructions, each node can communicate directly with only some of the other nodes that typically lie in its vicinity. On the other hand, the traffic requirements of the nodes are taken to be arbitrary, therefore it is necessary that nodes cooperate to forward packets to their final destinations. The lack of any wired infrastructure, the nature of the wireless channel, and the need for robustness and scalability create many challenging design problems in the link, network, and higher layers of the OSI model's hierarchy.

Both wireless and wireline networks need to support best effort and guaranteed network data service types. The focus of this thesis is on supporting such services over wireless networks, and to quantify the gains that can be achieved by cross-layer techniques which do away with the firm boundary that currently exists between the PHY and MAC layers, and the higher layers of the network protocol stack.

Before presenting the different proposals and mechanisms by which data services can be supported over wireless networks using the cross-layer paradigm, we described, in this chapter, the limitations of each communication layer in mobile wireless systems. We discussed the most important features based on cross-layer exchanged information, introduced for mobile ad hoc networks. Overall, we can conclude from the analysis presented in this chapter that despite the performance improvement that this new paradigm can achieve, there are some risks of changing the legacy layered architecture. Indeed, several issues need to be talcked before these interactions can be successfully exploited such as implementation, debugging, upgrading and standardization. We have to specify and explain whether cross-layer paradigm is suitable for all types of wireless networks and applications or not. Even if the answer is yes, it is necessary to maintain the layered approach, while enabling interactions between various protocols at different layers.

The aim of the research activities presented in this dissertation is to come up with new proposals or enhancing the existing ones in order to design a new cross-layer framework for enhancing the QoS support in MANETs. In the next chapter of the thesis, we propose a new analytical model for the EDCA 802.11e mechanism and we propose new schemes to enhance this standard in order to improve the service differentiation without cooperating with upper and lower layers. From Chapter 4 to Chapter 7, we describe our contributions which are based on the interactions between MAC and network layers.

Chapter 3

Performance analysis and enhancement of the 802.11e EDCA mechanism

The IEEE 802.11 Task Group E is working on a QoS-aware MAC protocol involving EDCA [33]. This scheme was developed to offer QoS capabilities to WLAN, providing significant improvements to multimedia traffic. MANETs will also benefit from this new technology since the most widely deployed and used wireless cards are those based on IEEE 802.11. The use of transmission opportunity (TXOP) EDCA differentiation parameter, received a little attention in the previous analytical studies in contrast to the contention window size-based differentiation mechanism.

In the fist part of this chapter, an analytical model is proposed to study and analyze the TXOP bursting transmission scheme. We expose formulates for saturation throughput, delay, and frame-dropping of different priority classes. The analysis allows us to derive the optimal durations of the packet bursting regarding application and network load parameters. We validate the accuracy of our model through ns-2 simulations.

The second part of this chapter focuses on performance enhancement of the contention based access mechanism based on the EDCA scheme. Our approach, called Adaptive Transmission Opportunity (ATXOP), aims to share the transmission channel efficiently and reduce the overhead cost. Relative priorities are provisioned by adjusting the transmission opportunity duration of each traffic class taking into account both application requirements and medium utilization. We evaluate through simulations the performance of ATXOP scheme and compare it with the basic EDCA. The results show that our new approach outperforms the basic EDCA, especially at high traffic load conditions. Indeed, ATXOP increases efficiently the medium utilization ratio and it can provide an overall goodput up to 25% higher than EDCA while achieving the desired delay differentiation.

The remainder of this chapter is organized as follows. In Section 3.1 provides a detailed description of the EDCA scheme. The description of the proposed EDCA transmission opportunity analytical models in both WLAN and multihop network are given in Section 3.2 and Section 3.3, respectively. It derives the formulates of saturation throughput, delay, and dropping-packets rate. The analytic results are studied in Section 3.4 to give a guide to configure $TXOP_{Limit}$. In Section 3.5, we give the most important works that have been described to address QoS support in WLAN. The description of the proposed ATXOP scheme is given in

Section 3.6. Simulation methodology and performance evaluation of our proposal are detailed in Section 3.7. Section 3.8 concludes this chapter by summarizing results and outlining the main remarks.

3.1 The IEEE 802.11e EDCA scheme

IEEE 802.11 wireless LAN specification defines two different ways to configure a wireless network: ad-hoc and infrastructure mode. In infrastructure mode an Access Point (AP) is needed to connect wireless stations to a Distribution System (DS), whereas in ad-hoc mode all wireless stations are distributed without access coordinator. In this chapter, we focus on ad-hoc networks since distributed random access control are often preferred to centrally coordinated access control. Distributed Coordination Function (DCF) is the basic medium access mechanism of 802.11 for both ad-hoc and infrastructure modes [42]. It uses CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) protocol. In this mode, if the medium is found idle for longer than a DIFS (Distributed InterFrame Space) then the station can transmit a packet. Otherwise, a backoff process is started. More specifically, the station computes a random value called backoff time, in the range of 0 and CW (Contention Window) size. The backoff timer is periodically decremented by one for every time slot the medium remains idle after the channel has been detected idle for a period greater than DIFS. As soon as the backoff timer expires, the station can access the medium. If no acknowledgment is received, the station assumes that collision has occurred, and schedules a retransmission by re-entering the backoff process [42].

The performance evaluation results in [11] show that DCF suffers from significant throughput degradation and high delay at high load conditions, which are caused by the increasing time used for channel access negotiation. Many medium access schemes have been proposed for IEEE 802.11 WLAN to provide some QoS enhancements for real-time audio and video traffics. Previous research works mainly focus on the station-based DCF enhancement scheme [1, 19, 64, 70]. When two or more TCP senders share the same receiver, they all receive TCP-ACKs with the same priority (limited to the same receiver priority). This tends to reduce the service differentiation. Furthermore, if the shared receiver is slow, the observed relative priority will also be reduced [2, 48]. This motivates the use of queue-based differentiation where a shared node handles simultaneously several flows with different priorities. There are many recent works that focus on the queue-based enhancement schemes [14, 54, 10, 53] since they perform more efficiently.

A new access method called Hybrid Coordination Function (HCF) is introduced by the IEEE working group in [33]. It combines functions from DCF and PCF mechanisms [42]. HCF provides QoS STAtions ($QSTA$) with prioritized and parameterized QoS access to the wireless medium. Indeed, the main purpose of IEEE 802.11e is to give multimedia streams higher priority when accessing the medium, decreasing end-to-end delay and allocating more bandwidth to such traffic if necessary.

EDCA is a contention-based HCF channel access. It has received most attention recently. This mechanism provides a priority scheme by differentiating between different Access Categories (AC) by applying different access parameters. Indeed, there is a set of EDCA parameters associated to each AC. Each $QSTA$ has four ACs to support eight User Priorities (UPs) as shown in Figure 3.1. Therefore, one or more UPs are mapped to the same AC queue. This comes from the observation that usually eight kinds of applications do not transmit frames simultaneously, and using less ACs than UPs reduces the MAC layer overheads. Each AC queue works as an independent DCF STA and uses its own backoff parameters. In EDCA, different $AIFS$, CW_{min}, CW_{max}, and $TXOP_{Limit}$ values are introduced to support service differentiation. Thus, the AC with the smaller AIFS has the higher priority. Moreover, assigning a short CW_{min} size to a high priority AC ensures that in most cases, high-priority AC is able to transmit packets ahead of low-priority one. If the backoff counters of two or more parallel ACs in one QSTA reach zero at the same time, a scheduler inside the QSTA will avoid the virtual collision by granting the EDCA-TXOP to the highest priority AC. At the same time, the other colliding ACs will enter a backoff process and double the CW sizes as if there is an external collision (see Figure 3.2). By this way, EDCA is supposed to improve the performance of DCF under congested conditions. However, simulation results show that although internal collision rates are low for EDCA, external collisions between the same priorities in different QSTAs are still high [53].

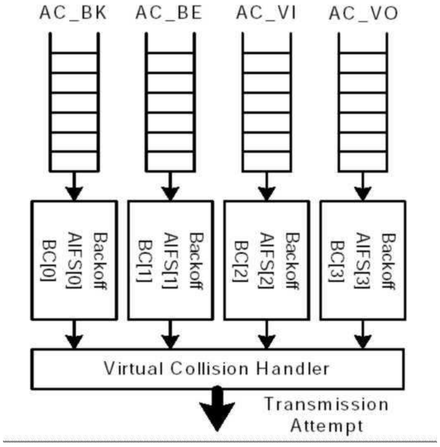

Figure 3.1: Four channel access functions for EDCA

The default values of $AIFS[AC]$, $CW_{min}[AC]$, $CW_{max}[AC]$ and $TXOP_{Limit}[AC]$ are fixed in [33]. However, how to adapt these parameters to the channel variations, is not defined by the standard and it remains an open research issue. To improve the throughput performance, EDCA packet bursting can be used in 802.11e, meaning that once a QSTA has gained an EDCF-TXOP, it can be allowed to send more than one frame without contending for the medium again. After getting access to the medium, the QSTA can send multiple frames as long as the total access time does not exceed the $TXOP_{Limit}[AC]$ bound. To ensure that no other QSTA interrupts the packet bursting, $SIFS$ is used between packet bursts. If a collision occurs, the EDCA bursting is terminated. This mechanism can reduce the network overhead and increase throughput by multiple transmissions using $SIFS$ and burst acknowledgments. However, bursting may also increase the jitter.

The two earlier works dealing with the transmission opportunity scheme in WLAN are presented in [66, 28]. Both works investigated, through analysis and simulation, the use of burst transmission with bloc Acknowledgement (Ack). They only consider infrastructure network where there is an access point which controls the $TXOP_{Limit}[AC]$ duration of the polled-stations. Moreover, the described analysis assumes that all the stations in the Basic Service Set generate the traffic with the same priority and thus no service differentiation

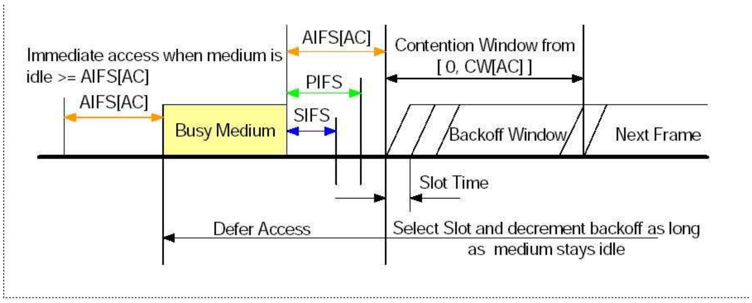

Figure 3.2: IEEE 802.11e EDCA channel access scheme

is considered. There are many analytical models that have been proposed to model the IEEE802.11e EDCA approach. Most of the presented analysis deal with service differentiation based on CW_{min}, $AIFS$, and persistent factor. In our work we adopt the default values of EDCA parameters introduced in [33] and we study the impact of varying $TXOP_{Limit}$ duration under different traffic load. We do not use block acknowledgments and we consider that each transmitted frame is immediately acknowledged.

3.2 EDCA analytical model

In this section, we propose a new analytical model of EDCA with packet bursts, in the saturation condition, i.e., the transmission queue of each AC is assumed to be always nonempty. Performance under saturation condition, which is not always the case in practice though, gives us fundamental bounds on system throughput, delay, and dropping packet rates. The proposed analytical model provides quantitative results of channel contention among prioritized traffic flows. It also gives us an insight on the different influence on service differentiation by each individual parameter (i.e.$TXOP_{Limit}$). Based on this result, one can configure traffic categories to achieve the desirable performance of service differentiation, whether moderate service differentiation or service separation between high and low traffic priorities.

The proposed model is based on the Markov chain shown in Figure 3.3 and introduced in [6, 73]. It extends the probability formulas to support differential $TXOP_{Limit}$ parameter in the different computed performance metrics. In the following, we denote by τ_i the probability that a station in the priority class i ($i \in \{0,1,2,3\}$) transmits during a generic slot time. Moreover, let p_i be the probability that AC_i senses the medium busy around it. According to [73]:

$$\begin{aligned} \tau_i &= \left(\sum_{j=0}^{L_{i,retry}} b_{i,j,0} \right) \\ &= \left(b_{i,0,0} * \frac{1 - p_i^{L_{i,retry+1}}}{1 - p_i} \right) \end{aligned} \tag{3.1}$$

where $b_{i,0,0}$ is the initial state of the AC_i in the corresponding Markov chain (see Figure 3.3), and $L_{i,retry}$ is the maximum number of retransmissions (retry limit). $b_{i,0,0}$ is expressed as follows

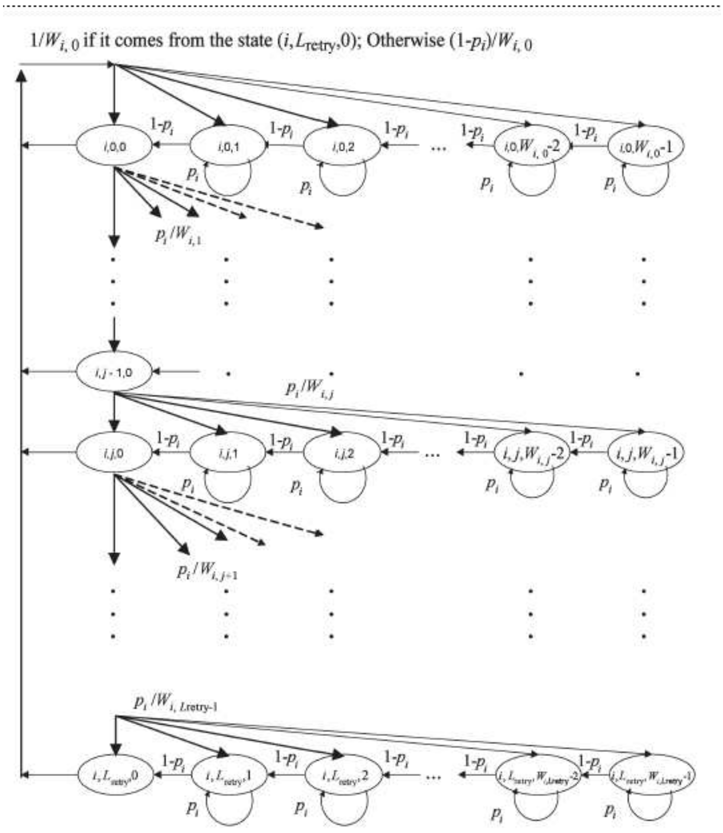

Figure 3.3: Marcov chain

$$b_{i,0,0} = \frac{1}{\sum_{j=0}^{L_{i,retry}} \left[1 + \frac{1}{1-p_i} \sum_{k=1}^{W_{i,j}-1} \frac{W_{i,j}-k}{W_{i,j}} \right] p_i^j} \tag{3.2}$$

Note that $W_{i,j}$ is the CW size at state (i, j). Thus, $W_{i,0} = CW_{min}$. We follow the basic EDCA backoff increase scheme:

$$W_{i,j} = min\left(2 * W_{i,j-1}, CW_{max}\right) \tag{3.3}$$

Let $p_{i,k}$ be the probability that an AC_i senses the medium busy. The p_i is produced when there is an internal or an external collision or both of them. Indeed, This implies that in the same node of any backoff of an AC_h with a higher priority than AC_i $(h < i)$, in the same node as the AC_i, transmits at the same time as AC_i. This is given by: $\prod_{h<i}\left(1 - \tau_h\right)$.

$$p_i = 1 - \left[\prod_{h<i} \left(1 - \tau_h\right) \right] * \left(1 - \tau_i\right)^{N-1} \left[\prod_{h=i+1}^{3} \left(1 - \tau_h\right)^{N} \right] \tag{3.4}$$

Where N is the total number of nodes in the network. From the point of view of a station, the probability τ that it accesses to the medium is:

$$\tau = 1 - \prod_{i=0}^{3} \left(1 - \tau_i\right) \tag{3.5}$$

Equations (3.4) and (3.1) form a set of a nonlinear equations. It can be solved by means of numerical methods.

We aim to derive for a given AC_i , the formulas of saturation throughput, delay, and queue dropping probability. We are also interested in studying the frame drop rate due to both buffer overflow and unsuccessful transmission after the maximum authorised number of transmission retries. We assume that the frame corruptions are only due to collisions, thus no channel error is considered.

For a given AC_i, the collision probability due to both internal and external collisions is:

$$p_{coll,i} = 1 - \tau_i \left(1 - \tau\right)^{N-1} \prod_{h<i} \left(1 - \tau_h\right) \tag{3.6}$$

and from the point of view of station is:

$$p_{coll} = \sum_{i=0}^{3} \tau_i \left(p_{coll,i}\right) \tag{3.7}$$

Where $h < i$ means that AC_h has higher priority than AC_i.

Let $p_{succ,i}$ be the probability to observe a successful AC_i transmission in the radio channel and p_{succ} the probability that a station achieves a successful transmission.

$$p_{succ,i} = N * \tau_i \left(1 - p_{coll,i}\right) \tag{3.8}$$

$$p_{succ} = \sum_{i=0}^{3} p_{succ,i} \tag{3.9}$$

The probability that a station is busy is

$$p_{busy} = 1 - (1 - \tau)^N \tag{3.10}$$

3.2.1 Throughput analysis

The saturation throughput represents the limit reached by the system throughput as the offered load increases, and is the maximum throughput the system can maintain under stable conditions. Therefore, we assume a fixed number of nodes, with each one of them having always a packet available for transmission in each AC_i. The saturation throughput of AC_i is then.

$$S_i = \frac{E\left[P_i\right]}{E\left[L\right]} \tag{3.11}$$

where $E[P_i]$ is the payload transmitted in a transmission period for a class i, and $E\left[L\right]$ is the length of a transmission period.

The average time T_{S_i} during which a bloc of frames are transmitted and the collision time T_{C_i} can be written, respectively, as for the RTS/CTS access mode:

$$T_{S_i} = T_{RTS} + SIFS + T_{CTS} + SIFS + AIFS_i + T_{data} + \delta + K_{TXOP_i} * (T_{data} + SIFS + T_{ACK} + 2\delta) \tag{3.12}$$

$$T_{C_i} = T_{RTS} + EIFS_i \tag{3.13}$$

Where δ is the propagation delay, K_{TXOP_i} is the maximum number of packets that an AC_i can transmit during $TXOP$ period and $EIFS_i = SIFS + T_{ACK} + AIFS_i$.

For the basic access mode:

$$T_{S_i} = AIFS_i + T_{data} + K_{TXOP_i} * (T_{data} + SIFS) \tag{3.14}$$

$$T_{C_i} = EIFS_i \tag{3.15}$$

For the IEEE 802.11a a function ceiling is used to handle the pad bits introduced in the PPDU format of IEEE 802.11a. The data transmission time T_{data} with size L_{DATA}, and the ack transmission delay T_{ACK} with size L_{ACK} are expressed as follows:

$$T_{data} = T_P + T_{PHY} + T_{SYM} * ceiling(\frac{16 + 6 + 8L_{H\_DATA} + 8L_{DATA}}{N_{DBPS}}) \tag{3.16}$$

$$T_{ACK} = T_P + T_{PHY} + T_{SYM} * ceiling(\frac{16 + 6 + 8L_{ACK}}{N_{DBPS}}) \tag{3.17}$$

Where, L_{H_DATA} is the length of data header, T_P is the transmission time of the physical preamble, T_{PYS} is the transmission time of the physical header, T_{SYM} is the transmission time of an OFDM symbol, and N_{DBPS} is the number of data bits per OFDM symbol.

So, the throughput is:

$$S_i = \frac{p_{succ,i}\left(K_{TXOP_i} + 1\right) E\left[length_{data}\right]}{\left(1 - p_{busy}\right)\theta + p_{succ} T_{S_i} + p_{coll} T_C} \tag{3.18}$$

Where $E\left[length_{data}\right]$ is the average data packet length and θ is the duration of the slot time.

3.2.2 Delay analysis

The frame delay time is defined as the time interval between two successive successful frame transmissions for a specific traffic category. It is possible that these two frames are not consecutive if a frame is dropped after exceeding the maximum retry count. Thus, all the flows with the same priority have the same average access delay. The average delay takes into account frames transmitted during a $TXOP$ since is computed from S_i:

$$E\left[D_i\right] = \frac{E\left[P_i\right]}{S_i} \tag{3.19}$$

3.2.3 Frame-dropping probability analysis

For the frame dropping probability analysis, we assume that there is no transmission error due to interference. Indeed, the non reception of an ACK is only due to collisions. Let $P_{i,drop}$ be the probability of packet drops.

$$P_{i,drop} = 1 - \left(1 - P_{i,drop,coll}\right) * \left(1 - p_{i,drop,queue}\right) \tag{3.20}$$

$$P_{i,drop,coll} = p_i^{L_{i,retry}+1} \tag{3.21}$$

where $P_{i,drop,queue}$ is the probability that a packet is dropped due to the queue overflow, and $P_{i,drop,coll}$ represents the probability of frame drops due to maximum retry limit [73]. Let K be the maximum size of the queue, and λ_i is the application rate of an AC_i. We assume an exponential arrivals and departures of packets in the queue in terms of packets. So the service rate is $\mu = \frac{1}{E[D_i]}$ and the traffic intensity ρ or the offered load is defined as:

$$\rho_i = \frac{\lambda_i}{\mu_i} \tag{3.22}$$

We consider the M/G/1/K state transition diagram. Let $P_{queue_{drop,i}}$ be the probability that there are K packets in the queue at an arbitrary time:

$$P_{queue_{drop,i}} = \frac{\rho_i^K\left(1-\rho_i\right)}{1-\rho_i^{K+1}} \tag{3.23}$$

3.3 Model extension to multihop networks

In this Subsection, we extend the above analytical model in order to evaluate the overall performance that can be achieved in multihop network. Each communication session from a source to a destination will be transmitted several times as it travels along the path through multiple hops. This increases the local traffic load, and also increases the collision probability. As a result these effects impact the final throughput. A packet transmitted in the network between two neighbouring nodes (at any hop throughout the transmission) will encounter a collision when either or both of the two following events happen in one time slot: (a) at least one of the transmitter's neighbours transmits; (b) at least one of the receiver's neighbours transmits. In the following, we denote by N_t the average number of a transmitter's neighbours except the receiver and by N_{rh} the average number of nodes located beyond a transmitter's transmission range but within the receiver's transmission range. The latter nodes N_{rh} represent additional contending stations to the receiver, which may also lead to collisions.

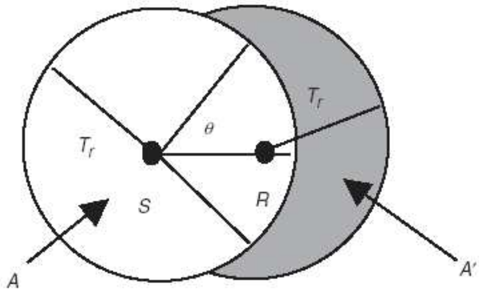

Figure 3.4: Contention area of the receiver

As mentioned before, p_{coll} is the probability of a collision seen by a packet being transmitted on the channel. The fundamental independence assumption given above regarding p_{coll} implies that each transmission sees the system in the same state. Then, at steady state each remaining station transmits a packet with probability τ. It should be noted that one of the basic assumptions of the throughput analysis presented in [6, 73] is the fact that there are no hidden terminals. The same approach does not apply directly due to its multihop transmission characteristics and the fact that neighbour terminals of the receiver are hidden from the transmitter (e.g. N_r nodes). However, under the saturation condition, terminals always have queued packets to send regardless of the source of the packet. Therefore, we use a similar approach to calculate the transmission probability $p'_{i,succ}$, that is the probability that a station transmits successfully in a randomly chosen slot time. The hidden terminal effect however is included in the calculation of the collision probability p'_{coll} by calculating the collisions around the receiver as well, as:

$$p'_{busy} = 1 - (1-\tau)^{N_{rh}+N_t} \tag{3.24}$$

$$p'_{i,coll} = 1 - (1-\tau)^{N_{rh}} (1-\tau)^{N_t-1} \prod_{h<i} (1-\tau_h) \tag{3.25}$$

and the probability that an AC_i success to transmit a packet as:

$$p'_{i,succ} = N_t * \tau_i \left(1 - p'_{i,coll}\right) \tag{3.26}$$

$$p'_{succ} = \sum_{i=0}^{3} p'_{i,succ} \tag{3.27}$$

3.3.1 Throughput analysis

If the total number of nodes in the network is n, then the total network integrated throughput, $S_{i,total}$, is expressed as:

$$S_{i,total} = \frac{n * p'_{i,succ}\left(K_{TXOP_i}+1\right)E\left[length_{data}\right]}{\left(1-p'_{busy}\right)\theta + p'_{succ}T_{S_i} + p'_{coll}T_C} \tag{3.28}$$

Taking in mind that the transmissions between source and destination considered go through multiple hops, the $S_{i,total}$ represents all these transmissions. Let now computing the average effective throughput. To this end, we have to divide the total amount of the integrated throughput $S_{i,total}$ over the average number of hop count $E[h]$ from the source node to the destination node. The $E[h]$ is defined as the mean number of hops between any pair of nodes. $E[h] = \frac{\log n}{\log E[d]}$, where $E[d]$ is the node density [31]. Let $P_{pathi,j}$ be the probability that exists a path between nodes i and j. Thus, we can write from [31]:

$$E[d] = \frac{2E[L]}{n} \tag{3.29}$$

$$E[L] = \frac{n\left(n-1\right)}{m\left(m-1\right)}\sum_{i=1}^{m}\sum_{j=i+1}^{m} P_{pathi,j}\left(t\right) \tag{3.30}$$

Where, $E[L]$ is the expected number of links in the network. In order to evaluate $E[L]$, an estimate for path availability is required. Among the few models available in the literature, we consider the model presented in [7].

If we consider h nodes between nodes i and j, the probability that packets from i reach j is:

$P_{pathi,j}\left(t\right) = [A(t)]^h$, where the $A(t)$ is the link availability between two neighbors at time t [7].

Let R_t be the transmission range of the sender and α is the network density. Thus, $N_{rh} = \alpha A'$ where the A' is the shaded area shown in Figure 3.4. The average size area can be calculated as:

$$A' = \int_{\Pi/3}^{\Pi/2} R_t\theta\left[\Pi - 2\theta + \sin\left(2\theta\right)\right]d\theta \cong \Pi * \left(0.8012R_t\right)^2 \tag{3.31}$$

Moreover, $\alpha = \frac{E[L]}{E_{max}}$, where E_{max} is the maximum number of bi-directional links in a network of n nodes, given by

$$E_{max} = \frac{n\left(n-1\right)}{2} \tag{3.32}$$

Thus, we can derive the effective saturation throughput as follows:

$$S_{effect,i} = \frac{S_{total,i}}{E[h]} \tag{3.33}$$

3.3.2 Delay analysis

The average delay in case of multihop networks is expressed as follows:

$$E'\left[D_i\right] = E\left[h\right] * \frac{E'\left[P_i\right]}{S_{effect;i}} \tag{3.34}$$

3.3.3 Frame-dropping probability analysis

We derive the probability of queue dropping by including the number of hops the packets have to cross. The aim is to study the effect of traffic load, queue length, and hidden terminals on the overall performance. In traditional analytical modeling of series of queues, the inter-arrival distribution of a queue is a function of service interval distribution at the previous queue and the arrivals due to internal queues at that node. Such series of queues are termed open-loop Jackson networks. From first principles, the service time distribution of a MAC queue in an open-loop Jackson network is identical to the inter-arrival time distribution with the mean arrival interval. Note that this result allows the modeling of each node individually as an independent M/M/1 queue. So, the queue drop probability is evaluated as described in Eq.3.23. The effect of hidden terminal is introduced in the equation since throughput and delay consider this effect. However, the probability of drops due to collision, is modified by including the transmission of terminals around the receiver and hidden to the transmitter:

$$P_{i,drop-coll} = \Big(p_i + \Big(1 - (1-\tau_i)^{N_r}\Big)\Big)^{L_{i,retry}+1} \tag{3.35}$$

3.4 Model validation and analysis of 802.11e performance

We compared numerical results with ns-2 simulations, using the TKN implementation of 802.11e [?] for the ns-2 simulator. The physical overhead of IEEE 802.11a is illustrated in Table 3.1. The data rate is $6\,Mbps$ and the control rate is $6\,Mbps$. Note that a data frame has 28 bytes as overhead including the MAC header and the FCS field and an ACK frame is 14 bytes in length [32]. Data rates of IEEE 802.11a are 6, 9, 12, 18, 24, 35, 48, and 54 Mbps and the corresponding N_{DBPS} are 24, 36, 48, 72, 96, 144, 192, and 216. The node topology of the simulation uses five and twenty different stations, contending for channel access. Each node uses all four ACs, and virtual collisions therefore occur. The EDCA parameters of each AC are presented in Table 3.2. Poisson distributed traffic consisting of 1024-bytes packets was generated at equal amounts to each AC. We further assume that the network operates in ideal channel conditions. The obtained results are presented in Table 3.3. We observe that the model corresponds relatively well with the outcome of the simulations. However, there are some differences that exceed the 97% confidence interval of the simulations. Note that the relative error is calculated by (Simulation Results-Analytic Results)/Analytic Results. The number of packets transmitted during *txop* period, considered in these results, is equal to 5.

SIFS	$16\mu s$	T_{PYS}	$4\mu s$
ACK size	14 bytes	T_{SYM}	$4\mu s$
PHY rate	6 Mbits/s	T_P	$16\mu s$
Slot-time	$9\ \mu s$	δ	$1\mu s$

Table 3.1: IEEE 802.11a PHY/MAC parameters used in simulation

In the following we investigate how the performance, in term of throughput, delay, and frame-dropping, of the EDCA is affected by the $TXOP_{Limit}$ value. To this end, we study mainly two effects on the performance. First, the variation of $TXOP_{Limit}$ and the number of nodes. Second, the variation of the offered load with the interface queue frame-dropping probability using different $TXOP_{Limit}$ values.

3.4.1 Effects of TXOP variations on overall performance

In this subsection, we aim to study the impact of the $TXOP_{Limit}$ value on the EDCA performance. To this end, we vary the number of packets that have to be transmitted during burst period (K_{txop}) for all ACs. Then, we observe the different metrics that have been derived by the above analytical model. Moreover, we give results for both 5 and 20 nodes, that are contending to gain the medium. By doing so, we also show the effect of the traffic load.

Figures from 1 to 8 show the obtained numerical results. We aim to pick out the more adequate length of $TXOP_{Limit}$.

Figures 3.5 and 3.6 show the throughput performance considering 5 and 20 nodes respectively. We observe that whenever the K_{txop} value increases, there is an evident step improvement in the overall performance especially for $AC1$ and $AC2$. This is due to the reduction of the per-MSDU overhead. Indeed, when the number of stations is equal to 5, the improvement in terms of throughput between the maximum and the minimum K_{txop} is about 23% for $AC1$ and about 12% for $AC2$. This improvement is less significant when K_{txop} is greater than 6 for all ACs. Therefore, when the number of nodes is equal to 20, the throughput gain reaches 30% for $AC1$ and about 20% for $AC2$. However, $AC3$ and $AC4$ keep almost a stable performance regardless of K_{txop} and the network load.

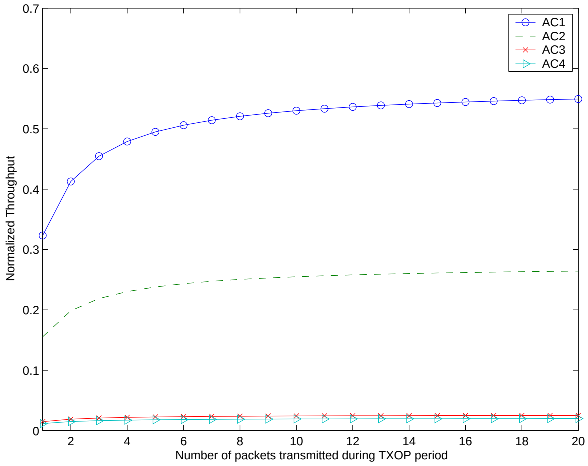

Figure 3.5: Throughput with 5 nodes

Parameters/ACi	**0**	**1**	**2**	**3**
CW_{min}	7	15	31	31
CW_{max}	15	31	1023	1023
AIFS[0,1,2,3](μs)	2	2	3	7
Max-retry limit[0,1,2,3]	7	7	7	4

Table 3.2: MAC parameters for the EDCA TCs

	Nodes	n=20	n=20	n=5	n=5
		throu.	delay	throu.	delay
AC1	Simul	0.33167	24.644	0.457	16.30
	Analy	0.31167	23.094	0.456	15.808
AC2	Simul	0.22693	27.574	0.209321	34.33
	Analy	0.23157	26.580	0.21991	32.862
AC3	Simul	0.03217	443.96	0.02174	347.1
	Analy	0.03366	440.67	0.022167	345.3
AC4	Simul	0.02331	554.88	0.01357	438.37
	Analy	0.024311	552.31	0.01937	435.28

Table 3.3: Simulation and analytical results

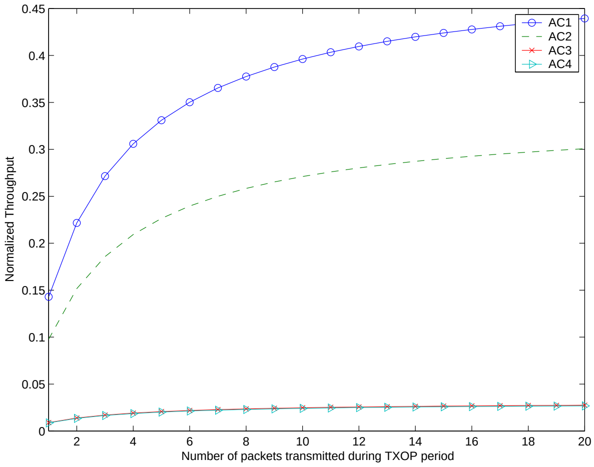

Figure 3.6: Throughput obtained with 20 nodes

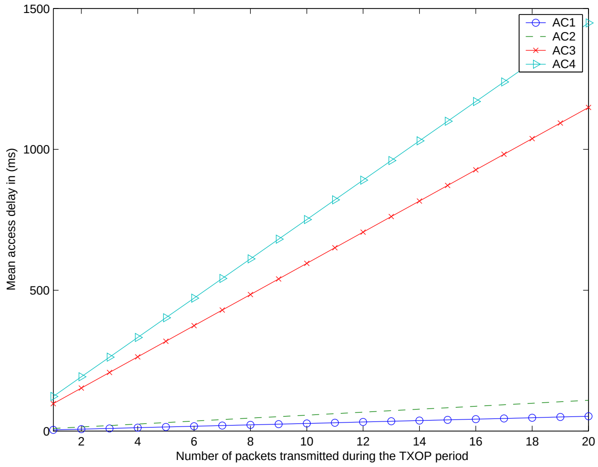

Figure 3.7: Mean access delay obtained with 5 nodes

According to Figures 3.7 and 3.8, the use of bursting still differentiates the media access delay for serving the high priority EDCA ACs. This affects mainly the delay of low priority $AC3$ and $AC4$ and generally it does not affect the others for both 5 and 20 nodes. Even when the value of K_{txop} is the same for all priority classes, the difference in the results between ACs increases rapidly. When the number of nodes is equal to 5, the mean access delay of $AC3$ and $AC4$ increases from $180\,ms$ when $K_{txop} = 2$ to $1050\,ms$ when $K_{txop} = 20$. Moreover, when the number of nodes is 20, $AC4$ suffers further and its mean access delay reaches $1500\,ms$ for the highest K_{txop}. However, $AC1$ and $AC2$ still keep the lowest and almost stable mean access delay even when the packet bursting increases. This means that the improvement in the throughput for high priority ACs has a great impact on the delay of lowest priority classes. The results indicate that these traffic classes suffer from starvation.

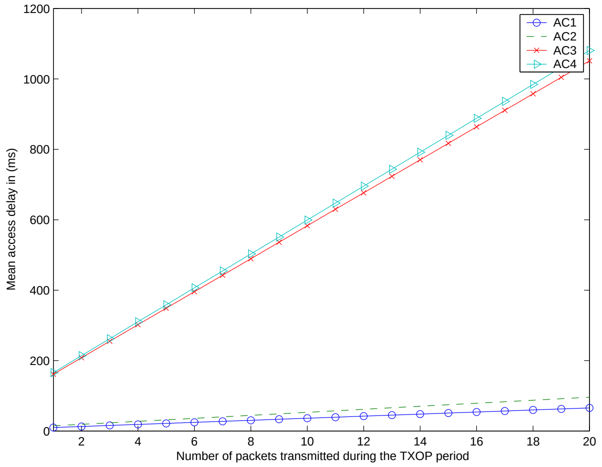

Figure 3.8: Mean access delay obtained with 20 nodes

We focus now on the impact of increasing packet bursting on frame dropping metric. We present results of drops due to IFQ overflow in Figures 3.9 and 3.10. Additionally, we show the results of the drop rates, due to maximum retry limit achieved and that is experienced by all ACs, in Figures 3.11 and 3.12.

From Figure 3.9, we observe that there is no IFQ drops for $AC1$ until $K_{txop} = 9$ and for $AC2$ until $K_{txop} = 4$. Then, the IFQ drop probabilities increase dramatically. As we can see, $AC3$ and $AC4$ experience high drop rates for all K_{txop} values. Moreover, when the number of nodes is equal to 20 (Figure 3.10), the results keep the same curves. The difference is that the drops for $AC1$ start from $K_{txop} = 6$ and for $AC2$ they start from $K_{txop} = 3$.

Figures 3.11 and 3.12 show that there is no important impact of the K_{txop} on dropped packets due to buffer overflow. However, the frame-dropping probability increases as the number of contending nodes does. Furthermore, since $AC4$ has the smaller retry limit, it has the highest frame dropping probability.

Our results show that whereas throughput is improved considerably when $TXOP$ is applied, care has to be taken not to increase the drops and delays above unacceptable rates and durations, respectively. Considering the setting of parameters we used in our evaluation, there is no benefits to use a K_{txop} value higher than 6.

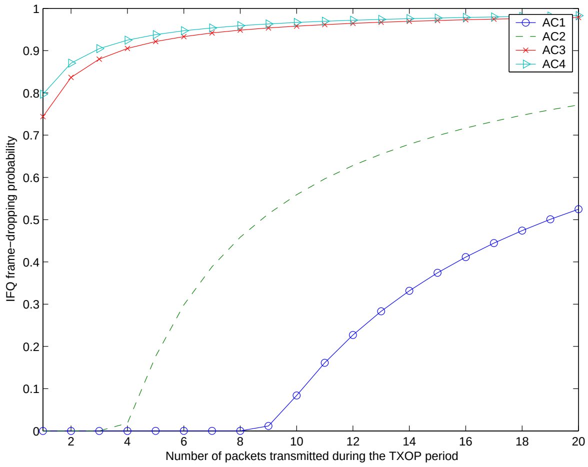

Figure 3.9: Interface queue frame-dropping probability with 5 nodes

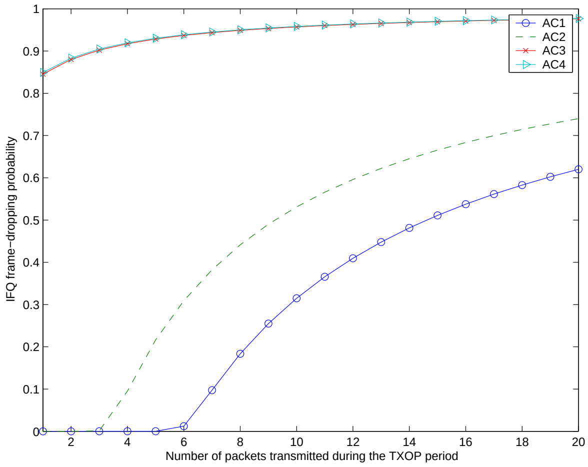

Figure 3.10: Interface queue frame-dropping probability with 20 nodes

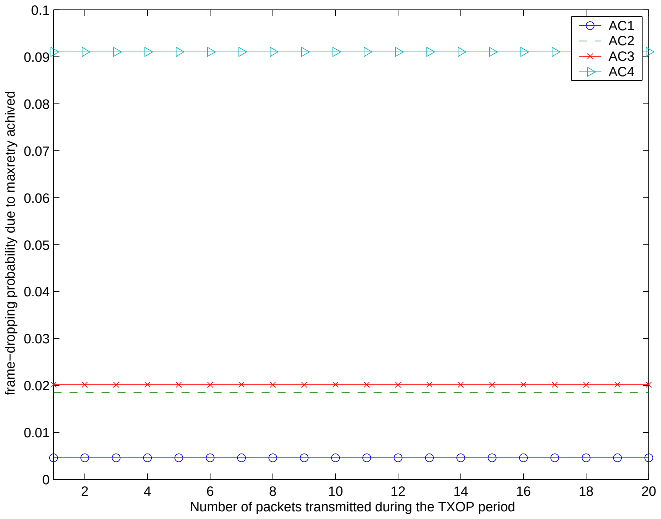

Figure 3.11: Frame-dropping probability due to retry limit achieved with 5 nodes

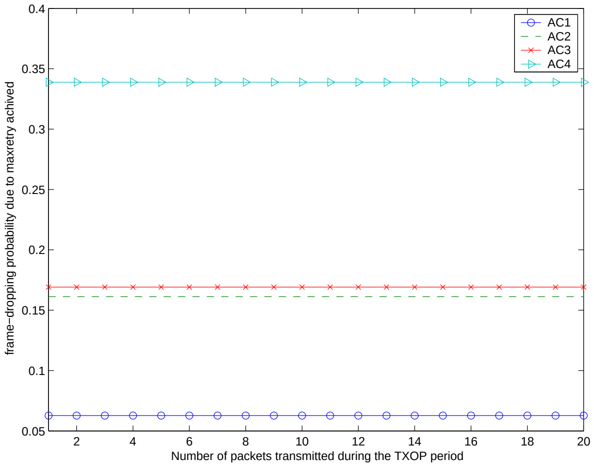

Figure 3.12: Frame-dropping due to retry limit achieved with 20 nodes

3.4.2 Effect of the sending rate variation on interface queue frame-dropping probability

In the following, we perform more experiments of specific settings and attempt to gain more understanding about the service differentiation effects of the $TXOP_{Limit}$ parameter. As we observed in the above subsection, the worst effect of increasing $TXOP$ value is the high frame-dropping rate especially due to IFQ drops. In this subsection, we further investigate this important issue. To this end, we vary the number of packets transmitted per second by each AC and we study the results obtained with different values of K_{txop}. In our experiment, each node send data traffic for all the four ACs using the same sending rate.

Figures from 3.13 to 3.16 show the analytical results for a number of stations equal to 20. The obtained curves indicate that when increasing the value of K_{txop}, increasing the sending rate do not overcome the frame dropping problem. Furthermore, for lower sending rate it is better to use a larger K_{txop} value.

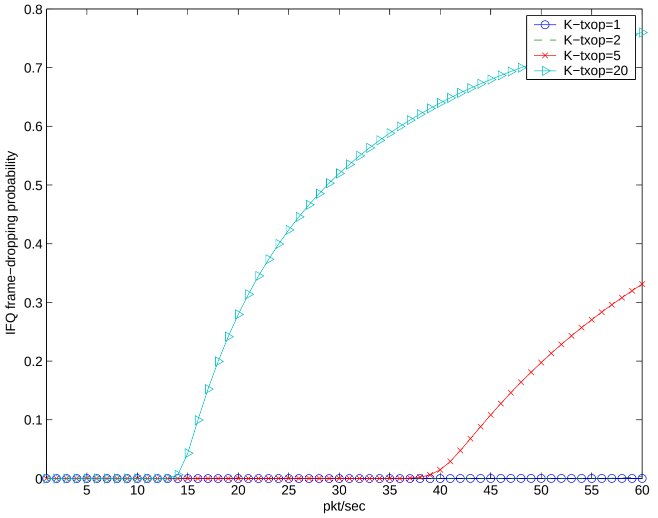

Figure 3.13: IFQ Frame-dropping probability of AC1

Based on the above analysis, we draw the following conclusion: if a wireless station carries a low-priority traffic flow, it should not use a larger value of K_{txop} to gain the wireless medium, thus favoring the fair performance enhancement. Note that applying a larger burst duration increases the frame dropping rates, even for the highest traffic flows. It the $TXOP_{limit}$ of all the ACs are carefully tuned to satisfy minimum packet drops and minimum delays, the desired fairness among traffic flows can be dynamically and accurately achieved.

In the next section, we review some works that have been introduced for EDCA mechanism enhancement. Then, we describe our scheme that proposes an adaptation of the $TXOP_{limit}$ parameter in order to provide a good application performance.

3.5 Related works on EDCA mechanism enhancements

Many works have addressed the enhancement of the EDCA scheme especially for highly loaded networks [47]. In [14] both the EDCA and the polling-based channel access modes are evalu-

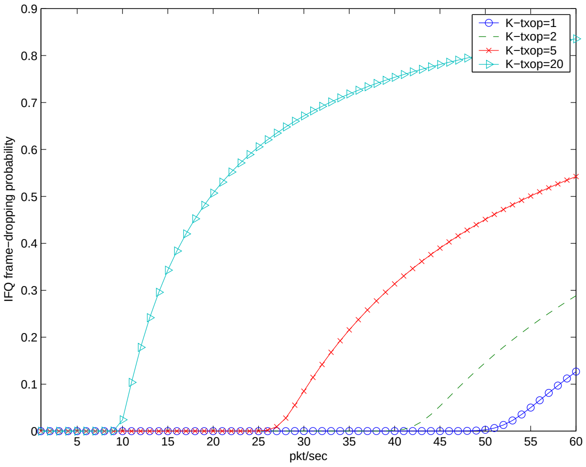

Figure 3.14: IFQ Frame-dropping probability of AC2

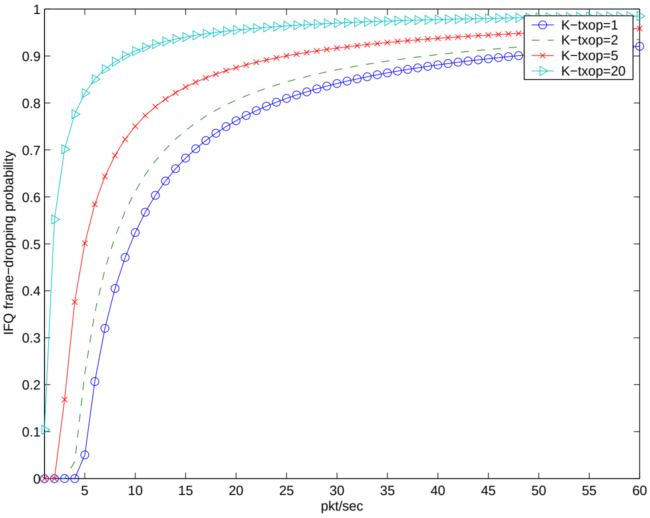

Figure 3.15: IFQ Frame-dropping probability of AC3

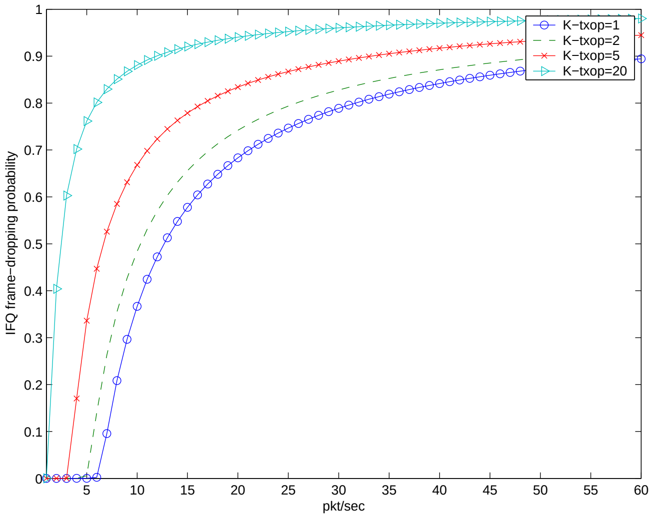

Figure 3.16: IFQ Frame-dropping probability of AC4

ated for QoS in IEEE 802.11 WLAN in carrying QoS applications. Through simulations, the authors show the performance under real time audio and video traffic. They find that EDCA provides satisfactory service differentiation among its four access categories. However, in the presence of heavy load traffic such as a high definition television (HDTV) signal transmission, it is more desirable to place such load under HCF polling mode to avoid the adverse impact of other traffic on this class of traffic. With a hybrid polling and EDCA protocol, network capacity is effectively increased to better support real-time audio and video transmissions in future home networks. Therefore, more centralized control at the AP is desired when heavy traffic load is expected in the network. However, if we consider a fully distributed network, at high load, EDCA cannot provide the minimum service guarantee for each class traffic.

In [9], the authors compare two approaches for Quality of Service support in WLAN-based ad hoc networks. The first approach is to use per-packet priorities, according to the IEEE 802.11e standard. The second approach is to allocate radio resources on the path between source and destination, according to the introduced protocol 'Distributed end-to-end Allocation of time slots for REal-time traffic' (DARE). For example, each node between source and destination allocates some dedicated time slots for this communication before the actual transmission starts. On the one hand, this removes uncertainties that come with a distributed random medium access. It thus has the potential to support applications demanding a non-varying end-to-end delay. On the other hand, such a reservation mechanism is typically much more complex than a priority mechanism. In particular, it adds signaling overhead to coordinate the nodes. All nodes between source and destination must agree in distributed manner on the reserved resources, the nonparticipating nodes must be informed so they abstain from transmission, and the reservation must be maintained and re-established when broken. Performance simulations show the following results: in case of low load, IEEE 802.11e has slightly lower end-to-end delay and higher packet loss rate than DARE, since it does not use any coordination among nodes for real-time packets. In case of medium load, DARE is superior in terms of jitter, delay, and packet loss. In case of high load, DARE clearly outperforms 802.11e. The results still hold if DARE has to repair the resource reservation path due to node failures.

In [39], the authors have presented an analytical model to analyze the performance of EDCA, the contention-based channel access mechanism in the forthcoming IEEE 802.11e protocol. All the important new features of the EDCA, as virtual collision, different AIFS, and CW have been taken into account. They also considered the difference of the count down procedure between the EDCA and the legacy DCF, as well as the retransmission limit. Based on the proposed model, the authors have studied the throughput performance for multiclass priority traffic and have proposed a recursive method to calculate the mean access delay. The model is validated via simulations. The effects of the CW and AIFS on the service differentiation ability of the protocol have been investigated. The results show that the number of ACs, or in other words, the traffic load, should be limited in order to provide a relatively satisfactory service level for both high priority and low-priority ACs. The model and analysis provide an in-depth understanding and insights into the EDCA mechanism. They also provide helpful and powerful tools for further study, such as parameterization for some types of traffic and development of call admission control schemes for further QoS improvement for WLANs. Any solution have been introduced to solve problem when the traffic load increases.

In [10], the authors exposed results relative to the interaction of reactive routing protocols for MANETs and the IEEE 802.11e MAC layer technology. This work shows the importance of using EDCA to achieve service differentiation comparing to the basic access (DCF). The study focused on the performance improvements in terms of TCP and UDP traffic in a typical MANET environment when uniquely routing packets are assigned to the highest priority access category under IEEE 802.11e. The difference in behavior of two reactive routing protocols (AODV and DSR) relating the traffic throughput and routing overhead results to their internal mechanisms, are detailed. Results show that when routing packets benefit from the prioritizing mechanism of IEEE 802.11e the performance is improved drastically. The authors found that this improvement is due to an increase in the responsiveness of the different routing protocols. In terms of TCP throughput gain achieves an increase of up to 150% with DSR and up to 300% with AODV. Maximum UDP throughput is also increased substantially, up to 200% for both routing protocols. Relatively to normalized routing overhead, which is the reference metric used in simulations to measure the performance of the routing protocols, the IEEE 802.11e allows achieving better results. The difference becomes more noticeable as the level of saturation in the network is increased, since saturation causes the malfunction of routing protocol mechanisms. Overall, the authors conclude that upgrading the MAC layer of MANET stations to IEEE 802.11e is very important not also for multimedia traffic support, but also to improve the efficiency of the routing mechanism used, especially if it is a reactive one.

To improve the performance under different load rates and to increase the service differentiation in EDCA-based networks, a new scheme called Adaptive Enhanced Distributed Coordination Function (AEDCF) has been proposed in [53] as an output of our previous work. This scheme extends the basic EDCA by making it more adaptive taking into account network conditions. Indeed, AEDCF uses a dynamic procedure to change the contention window value of each priority class differently. In fact, each class updates its contention window based on the estimated collision rate computed during a constant period. For further differentiation, each traffic category multiplies this collision rate by a priority factor [53].This mechanism offers to high priority traffic a higher probability to generate smaller CW value than low priority traffic and so they can access the medium first. Moreover, this scheme achieves a high medium utilization and it is much more efficient at high load. Furthermore, it improves total goodput, delay and delay-jitter. The TXOP is not considerd in this proposal.

In [54], a detailed evaluation of the EDCA protocol with the Contention Free Burst (CFB) option to quantify its performance gain is performed. The impact of the MAC transmit buffer size is also incorporated. Accordingly, the authors propose a suitable approach to guide the configuration of the burst limit ($TXOP_{limit}$). They have shown that the bursting option can be used to improve performance. However, it is shown that for an optimized operation, the proper configuration of the $TXOP_{limit}$ variable is crucial and could be associated with the MAC buffer size. The simulation results introduced in [54] show that, a limit proportional to at least 50% of buffer occupancy and not larger than 100% should be utilized. Moreover, the results prove that the bursting option can be used to improve performance. Indeed, it is shown that for an optimized operation, the proper configuration of the $TXOP_{limit}$ variable is crucial and be closely associated with the MAC buffer size. However, this consideration couldn't be efficient if the packet loss is not considered. Indeed, this work does not give any results about throughput performance and so the minimum service guarantees specially for low priority traffic that suffer from starvation at high traffic load.

It is clearly demonstrated, in all related works, that the basic EDCA operation fails to scale well. The tuning of the $TXOP$ duration can provide a big benefits to enhance application performance. Hereafter, we adapt this parameter to network conditions and average packet size in each priority queue in order to achieve a good QoS support while providing some minimum service levels for each traffic category.

3.6 ATXOP: Adaptive Transmission Opportunity scheme

In order to efficiently support time-bounded multimedia applications, we use a dynamic procedure to change the $TXOP$ duration. We believe that this adaptation will enhance medium utilization and so increase the total goodput of the traffic which becomes limited when using the basic EDCA, mainly for high traffic load.

During the TXOP, the station can send a burst of DATA frames separated by SIFS. The TXOP ends when there are no more frames to be transmitted or when the TXOP maximum duration expires. In a fully distributed network, and where there is no QoS Access Point (QAP) that adapts the $TXOP_{Limit}$ according to the traffic characteristics and the network conditions, the default TXOP maximum duration values could not be efficient when we have an heterogeneity of application characteristics. Indeed, we believe that the $TXOP_{Limit}$ value for each priority queue depends mainly on the average packet size to be transmitted in that duration and so on the medium utilization fraction during a controlled period T while maintaining service differentiation. Indeed, different packet size yielding different transmission time duration. Moreover, for high loaded networks it is better to maintain as lower as possible the overhead cost due to MAC control packets (RTS/CTS packets), in order to efficiently use the medium for data transmission. For this purpose, we develop a mechanism that adapts the $TXOP_{Limit}$ value according to the average packet size and the medium utilization. In the following, we give a detailed description of this proposal.

3.6.1 $TXOP_{Limit}$ calculation based on the average packet size in each queue

At the beginning of each control period T, we set each $TXOP$ duration according to the average packet size in each queue i and its priority level. We compute the average packet size in each queue i that we note $avgpkt[i]$. Our target is to ensure that the number of packets of each class served in every period is proportionally of its priority. In other words, let's

N be the total number of packets sent successfully by a given station during the period T, our objective is then to get the number of packets of class i as close as possible to $N_i = \left(\frac{\alpha_i}{\Sigma\alpha_j}\right)N$, where $\alpha_i$is a differentiation factor that gives a weight to the total transmitted packets for each class comparing to the total transmitted packets at each node. We can also write $N_i = TXOP_{duration}[i]\frac{rate}{avgpkt[i]}$. From the previous two equations we can write $TXOP_{duration}[i] = \left(\frac{\alpha_i}{\Sigma\alpha_j}\right)N\frac{avgpkt[i]}{rate}$. Then, we can write

$$TXOP_{duration}[i+1] = \frac{(\alpha_{i+1})}{(\alpha_i)}\frac{avgpkt[i+1]}{avgpkt[i]}TXOP_{duration}[i]. \tag{3.36}$$

This relationship between each two successive priority classes ensure a tightly differentiation between them. Therefore, if we set the value of TXOP for one AC, we are able to compute those of other ACs. We explain hereafter how TXOP of AC_0 is adapted dynamically according to the medium utilization.

3.6.2 Adaptive Transmission opportunity based on medium utilization

Assuming all ACs adopt different fixed TXOP durations according to their priority as described in the draft [33], they will have the same relative probability of obtaining a successful contention and hence the same average amount of throughput. However, this amount will decrease rapidly since the overhead cost will affect the effective throughput. Moreover, in saturation conditions, the success transmission probability is very low. Thus, it will be better to maximize the transmission opportunity duration of each AC while maintaining a good service differentiation.

At the end of each control period T, the station number s computes the TXOP of the highest AC (AC_0) as follows:

$$TXOP_{highest} = \max\left(TXOP_{min},\, f_s * TXOP_{max}\right) \tag{3.37}$$

where $f_s = \frac{Total_{BusTime}}{T}$, is the medium utilization parameter, $Total_{BusTime}$ is the total busy time around the node during the control period T, $TXOP_{min}$ is a minimum value that allows to avoid medium starvation, and $TXOP_{max}$ is a parameter used to prevent the medium being monopolized by a given class. The medium utilization parameter is computed dynamically in each period T expressed in time-slots. This period called update period should not be too long in order to obtain an accurate estimation and should not be too short in order to limit the complexity.

3.7 Performance evaluation of ATXOP

We have implemented ATXOP in the ns-2 simulator [49]. We have extended the EDCA scheme [33] to support our proposed algorithm detailed in section 3.6. We report in this section the set of simulations we have done with different traffic load and source characteristics. We also provide a performance analysis of our proposal based on the obtained simulation results and we compare it with the original scheme.

3.7.1 Scenario description

Our simulations use different types of traffics to evaluate service differentiation. Three queues are used in each station. The highest priority queue in each station generates packets with packet size equal to 160 bytes and inter-packet interval of 20 ms, which corresponds to 8 Kbps audio flow. The medium traffic queue generates packets of size equal to 1280 bytes each 10 ms which corresponds to an overall sending rate of 128 Kbps (video flow). The low priority queue in each station generates packets with sending rate equal to 120 Kbps, using a 1500 bytes packet size. The physical data rate is set to 36 Mbps. The nominal bit rate is 2 Mbps. For all the scenarios considered, we set the EDCA queue parameters based on the draft [33]. To increase the load of the system, we gradually increase the number of stations. We start simulations with two wireless stations, then we increase the load rate by increasing the number of stations by one every eight seconds. We increase the number of stations from 2 to 16 which corresponds to load rates from 9.5% to 100%.

For this purpose, we use the topology shown in Figure 3.17, which consists of n stations indexed from 1 to n. Each station generates the same traffic of three data streams, labeled with high, medium and low, according to their priorities. Station number n sends packets to station number 1. Station number i sends to station number $i+1$, for $1 \leq i \leq n-1$ three flows belonging to the three classes of service: Audio (high priority), Video (medium priority), and Background Traffic (denoted by BT for low priority). We use CBR sources to simulate BT, video, and audio traffics.

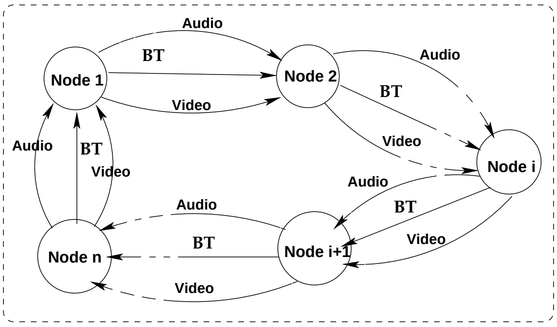

Figure 3.17: Simulation topology

In the following simulations, we assume that each wireless station operates at IEEE 802.11a PHY mode-6, see network parameters shown in Table 3.4.

Table 3.5 shows the network parameters selected for the three TCs.

3.7.2 Simulation metrics

We analyze several QoS metrics to evaluate the performance of our approach and we compare results with the basic EDCA mechanism protocol. The following metrics are considered:

- **Total goodput**: This metric computes the total amount of goodput delivered successfully by the MAC layer of all the station in the network.
- **Latency:** It is the average end-to-end delay of all traffics in the different stations. The average delay is used to evaluate how well the schemes can accommodate real-time flows.

SIFS	$16\mu s$
DIFS	$34\mu s$
ACK size	14 bytes
Data rate	36 Mbps
Slot_time	9 μs
CCA Time	$3\mu s$
MAC Header	28 bytes
Modulation	16-QAM
Preamble Length	$20\mu s$
RxTxTurnaround Time	$1\mu s$
PLCP header Length	$4\mu s$

Table 3.4: IEEE 802.11a PHY/MAC parameters used in simulation

However, real-time flows require both low average delay and bounded delay jitter. So we will also use the following metrics of latency distribution and delay variation.

- **Latency distribution**: Latency distribution allows to trace the percentage of packets that have experienced a latency less than the maximum delay required by the applications.

3.7.3 Results Analysis

To evaluate the performance of ATXOP, we investigate in this section the effect of the traffic load and compare it with the basic EDCA scheme.

We analyze throughput, delay, and delay distribution metrics to evaluate the performance of our approach and we compare results with the basic EDCA scheme. Figures 3.18, 3.19, and 3.20 show the significant improvement obtained by ATXOP scheme comparing to the basic EDCA.

Figure 3.18 shows that ATXOP provides significantly a better total throughput compared to the basic EDCA, mainly in high load situations (about 30% total goodput gain when the channel is fully loaded).

Parameters	**High**	**Medium**	**Low**
CW_{min}	7	31	31
CW_{max}	15	31	1023
AIFS(μs)	34	43	52
Packet Size$(bytes)$	160	1280	200
Packet Interval(ms)	20	10	12.5
Sending rate$(Kbit/s)$	64	1024	128
slot-time(us)	16	16	16
$TXOP_{max}(ms)$	0.003	0.006	0.003

Table 3.5: MAC parameters for the three TCs

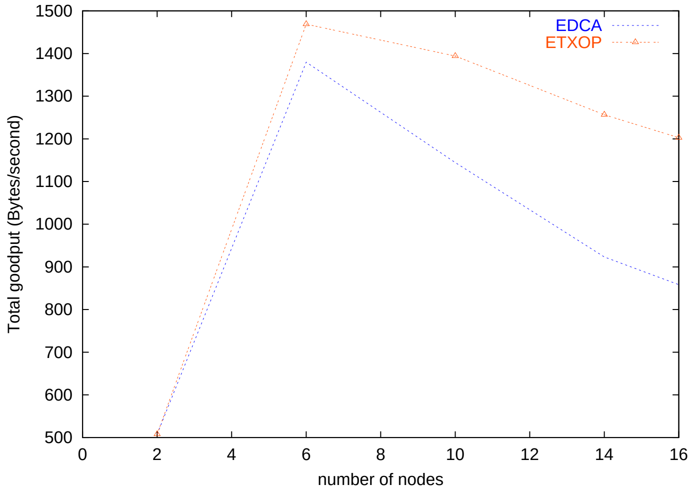

Figure 3.18: Mean goodput

Figure 3.19 shows the mean delay of all traffics. The ATXOP scheme is able to keep the delay lower than the basic EDCA even when the traffic load is very high. We can see that the mean delay for ATXOP is 41% smaller than that for the basic EDCA when the load rate is up to 100% (16 stations). Moreover, the mean delay of our approach is still similar or smaller than that of the basic access scheme when the load rate is low.

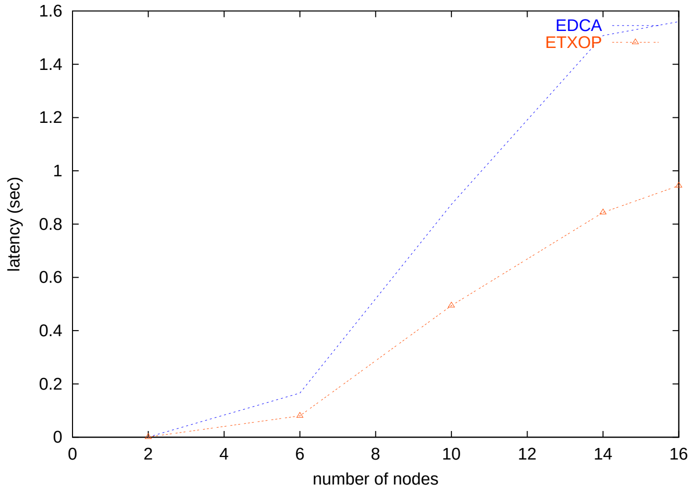

Figure 3.19: Latency

We show the latency distribution for video traffic in Figure 3.20, in which a fixed number of 16 stations is used to show the delay performance. There are considerable differences between them, i.e. more than 75% of video packets for ATXOP have delay less than $400ms$, whereas only 17% of video packets for EDCA have delay less than $400ms$.

All the obtained results show that ATXOP always outperforms EDCA. We believe that the adaptive transmission opportunity is very efficient to estimate the network status and reduce

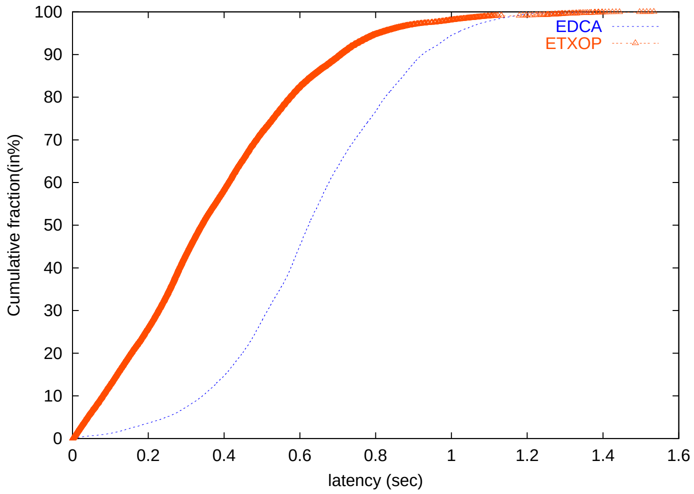

Figure 3.20: Latency distributions of video

collisions while providing service differentiation support. Indeed, it enhances the medium utilization and so increases the total goodput of the traffic which is much lower when using the basic EDCA, mainly for high traffic load. Indeed, in such conditions it is better to avoid adding overhead like RTS/CTS packets, and the backoff procedure for each transmission packet to efficiently use the medium for data transmission.

3.8 Chapter summary

In the first part of this chapter, we analyzed the performance of data burst transmissions, supported by the EDCA access mechanism. We developed an analytical framework to evaluate the impact of the $TXOP_{limit}$ on the overall performance. Though $TXOP$ parameter provides efficient service differentiation and preserves service to high priority traffic at high loads, we showed that it is especially prone to starving lower priority traffic. We suggested that the value of $TXOP_{limit}$ of the ACs, should be limited in order to guarantee a relatively satisfactory service level for both high priority and low priority classes. The model and the analysis provide an in-depth understanding and insights into the performance of the EDCA mechanism. It also provides helpful and powerful tools for further study, such as parameterization of the $TXOP_{limit}$ for some types of traffic characteristics for further QoS improvement for both WLANs and multihop networks. Indeed, EDCA parameters have to be properly set to provide prioritization of ACs and ensure minimum service guaranty to best effort traffics. Tuning them in order to meet specific QoS needs is a current research topic.

The second part of this chapter described a new adaptive transmission opportunity scheme for Quality of Service enhancement for IEEE 802.11 WLANs. To this end, we propose to adapt the TXOPLimit value as follow: at the beginning of each control period T, the TXOP duration is set according to the average packet size in each queue and its priority level. Then, we establish an analytic relationship between these values according to the medium utilization level while maintaining service differentiation. The goal is to enhance real-time applications and avoid starvation of low priority traffics.

We evaluate through simulations the performance of ATXOP scheme and compare it with the basic EDCA. The results show that our new approach outperforms the basic EDCA, especially at high traffic load conditions. Indeed, ATXOP increases efficiently the medium utilization ratio and it provides an overall goodput up to 25% higher than EDCA while achieving delay differentiation.

One of important extension of this scheme is to evaluate the performance of this proposal in multihop networks and provide interaction between MAC and routing protocols in order to tune the TXOP duration according to the network conditions that could be included in the routing packets. Therefore, the main key to QoS enhancement in wireless communications should be based on a coordination between all levels of the wireless protocol stack. Moreover, how to address an efficient cross layer QoS model based on energy conservation, stability, end-to-end delay, and other key metrics for real time applications. In other words, How to map between service differentiation and network QoS management.

In the next chapters, we consider a joint optimization of routing and medium access where the goal of the optimization is to achieve a good medium utulization, energy conservation, and provide a QoS enhancement of application performance.

Chapter 4

An energy consumption rate-based routing mechanism for MANETs

Some scenarios where MANET could be used are business associates sharing information during a meeting, military personnel relaying tactical and other types of information in a battlefield, and emergency disaster relief personnel coordinating efforts after a natural disaster such as a hurricane, earthquake or flooding. In such scenarios, maximizing the network lifetime by using the nodes with the maximum residual energy (lifetime) is a very important challenge since recharging battery is very difficult (hard) to do in such conditions.

This chapter describes an adaptive routing mechanism based on the energy consumption rate of nodes which is used to enhance on-demand and proactive routing protocols in MANET. The goal is to deal with energy exhaustion problem in order to extend network connectivity. Our algorithm allows a fair energy consumption during route establishment by building routes that are less congested than the others. To do this, the congestion information is obtained from a computed cost that depends mainly on the energy consumption rate which measures how fast a node is consuming its residual energy. The main features of our mechanism is that it is simple, efficient and it can be applied for any routing protocol. As an example, we mainly focus in this chapter on the enhancement of the AODV (Ad hoc On-demand Distance Vector) reactive routing protocol by adding the support of our proposed mechanism. At the end of this chapter, we also provide guidelines to incorporate our scheme into the OLSR reactive routing protocol. We believe that adding our proposed concept in other (on-demand and proactive) routing protocols is possible in a similar manner and it will enhance its performance.

We evaluate through simulations the performance of the modified AODV routing protocol including our scheme (resulting to a new routing protocol that we call **E-AODV**) and we compare it with the basic AODV routing protocol. Results show that our new concept outperforms the basic AODV. Indeed, it reduces for more than 20 % the total energy consumption and decreases the mean delay specially for high load networks while achieving a good packet delivery ratio. Moreover, the simplicity of the mechanism enables its implementation with a very reduced complexity.

The remainder of this chapter is organized as follows. In Section 4.1, we give the most important characteristics of AODV routing protocol and its main limitations. In Section 4.2, we describe our algorithm in detail. Simulation methodology and performance evaluation of our proposal are detailed in Section 4.3. The proposed guidelines to enhance the OLSR routing protocol by including the proposed mechanism in its two main operation phases (MPR list selection and route determination) are described in Section 4.4. We review some works related to our proposal in Section 4.5. Section 4.6 concludes the chapter by summarizing results and outlining future works.

4.1 The basic AODV routing protocol

We integrated our scheme in the reactive AODV routing protocol described in [43]. We choose AODV as one of the on-demand MANET routing protocols because, as shown in [74], it consumes less energy than other similar routing protocols such as Destination Sequence Distance Vector (DSDV) and Temporally Ordered Routing Algorithm (TORA). We studied the characteristics of the AODV protocol and enhanced the route establishment algorithm. The performance evaluation studied with different useful metrics and different scenarios, show the great benefits of this approach in terms of energy consumption, end-to-end delay, and route establishment delay.

Hereafter, we first give a short overview of the AODV routing protocol then we enumerate the limitations of this protocol by highlighting its disability to select the routes which have the best residual lifetime.

4.1.1 An overview of AODV

AODV is a reactive (on-demand) routing protocol. It is self-starting, enables multi-hop routing between participating mobile nodes wishing to establish and maintain an ad hoc network. This protocol builds routes between nodes only as desired by source nodes. It discovers routes quickly for new destinations, and does not require nodes to maintain routes to non-active destinations. AODV ensures link breakages and breakdowns are handled efficiently. The AODV protocol establishes routes using a Route REQuest (RREQ) / Route REPly (RREP) query cycle. So, when a node requires path to destination, it broadcasts RREQ message to its neighbors which includes latest known sequence number for that destination. This message is flooded until information required is complete by any means. Each node receiving the message creates a reverse route to the source. The destination sends back RREP message which includes number of hops traversed and the most recent sequence number for the destination of which the source node is aware. Note that if an intermediate node has a fresh route to the destination it doesn't forward the RREQ and it generates a RREP toward the source. Each node receiving the RREP message creates a forward route to the destination. Thus, each node remembers only the next hop required to reach any destinations, not the whole route.

Each node receiving a duplicate of the same RREQ, drops the packet. Moreover, AODV uses sequence numbers to ensure the freshness of routes. In fact, the routes to any destination are updated only if the new path toward that destination has greater sequence number than the old one or it has the same sequence number but with less number of hops.

From the AODV operations described above, we can conclude that it builds routes between nodes which have the less number of hops and does not take into account other QoS parameters such as the residual energy, the links stability, and the delay experienced at the MAC layer.

4.1.2 AODV Limitations

As stressed above, routes in AODV protocol are established based on the minimum hop count. Indeed, they are selected if it has the lower number of intermediate nodes between source and destination. Residual energy level and node congestion of these intermediate nodes are not considered in the route selection process. Consequently this might have a bad effect when the number of communications increases and so it is more likely to include other parameters that have a significant effect on network connectivity and lifetime. Furthermore, power is a very important constraint in wireless network. If a node, participating in a route establishment, has a very low energy, this route will break very soon. Thus, frequent link failures has a bad effect on the network lifetime: there are some nodes that will die sooner than other ones. So, this can affect network connectivity if the key nodes die very soon. The key nodes here are the nodes that routes cannot be established if they die when their energy returns to zero. To deal with these problems, the power should be taken into account in the route establishment algorithm. To this end, we propose an energy-aware routing establishment mechanism and we apply it to the AODV routing protocol. The main feature of our work compared to the related works in this area (that will be discussed in Section 4.5) is that is simple and efficient according to the obtained performance enhancement comparing to the results obtained with the basic protocol.

4.2 Our proposed mechanism

4.2.1 Motivation

Having in mind the different problems and constraints described above, that are in general related to the route selection procedure based on minimum hop count, we propose a simple energy consumption rate-based mechanism that aims to maximize the network lifetime and enhance the performance obtained by the basic AODV routing algorithm. The main goal of our mechanism is routing the packets through the nodes which we expect to have the better residual lifetime among all possibilities.

The idea of routing based on residual lifetime is not new as it has been used in some previous research works available in the open literature [27, 41, 68]. For example, in [41], the authors proposed a max-min algorithm that uses a parameter called $zPmin$ representing the minimum power consumption in the network. There is a well-known problem with these so called max-min (maximize the minimum residual energy/lifetime) schemes consisting on increasing the average path length when avoiding nodes with low residual energy (or lifetime). The obtained routes consume more energy per-packet and reduce the effective network lifetime. Indeed, the max-min strategy does not help if many nodes are running out of energy. The method described in [41] requires accurate power level information for all nodes in the network. For dynamic MANET characteristics this is not a feasible assumption.

In [27], routes are built based on a computed cost which takes into account the current energy level. Some of the ideas of this chapter could be considered as an extension of [27] proposal. However, we use a much more sophisticated way to compute the residual lifetime for each node. We aim to propose a simple extension of AODV that requires a very small additional overhead to get benefit from taking into account the energy considerations. Moreover, the cost metric that we use in building the routes is updated periodically according to real power consumption behavior. Another important feature of our proposal is its applicability to any kind of routing protocol. Furthermore, since reactive routing protocols have benefits for ad hoc networks, our scheme is suitable for on-demand concept and, as we will see in the performance evaluation section, it improves QoS performance metrics while achieving a good energy conservation.

4.2.2 Computing the expected residual lifetime

In our algorithm, we do not only consider the current energy level value of a node as the case of several mechanisms that have been proposed in this area. However, we take into account also the rate of energy consumption at each constant period of time. In fact, we believe that considering energy rate consumption allows us to get information about the energy exhausted in packet transmission and reception without doing complex computation of these values. Additionally, this also implicitly consider the data traffic load at each node. Then, using both the estimated rate consumption and the residual energy, we compute the *expected residual lifetime* assuming that the node continues to consume energy with that rate. By this way, we give more real information about the battery lifetime behavior in each node. Moreover, we try to differentiate between nodes that participate in communications more than other nodes even they have the same energy level. More precisely, at each period of time number j called "update period"[1] and for each node, we follow the following formula to compute the **energy consumption rate**:

$$E_{rate}(j) = \frac{E_{remain}(j) - E_{remain}(j-1)}{T_{Update}} \tag{4.1}$$

where $E_{remain}(j)$ is the estimated residual energy computed at update period number j and it is computed as follows:

$$E_{remain}(j) = \max\left\{ E_{current}(j) - \sum_{i=1}^{i=N_{pkts}} E_{Tx}(i), 0 \right\} \tag{4.2}$$

where $E_{current}(j)$ is the current energy value of the node. For more accurate estimation of this residual energy, we reduce the value of the power that will be consumed to transmit the remaining packets in the buffer noted by N_{pkts}. The parameter $E_{Tx}(i)$ quantifies the needed energy for transmitting the packet number i.

To minimize the bias against transient consumption rate, we use an estimator of Exponentially Weighted Moving Average (EWMA) to smoothen the estimated energy consumption rate values. Let $\overline{E_{rate}(j)}$ be the average energy consumption rate at step j (for each update period) computed according to the following equation:

$$\overline{E_{rate}(j)} = (1-\alpha) * E_{rate}(j) + \alpha * \overline{E_{rate}(j-1)} \tag{4.3}$$

where $\alpha \in [0,1]$[2], is the smoothing factor.

Then, we can estimate the **expected residual lifetime** $T_{lifetime}(j)$ in each node considering $E_{remain}(j)$ and $E_{rate}(j)$ values computed at each update period number j as follows:

$$T_{lifetime}(j) = \frac{E_{remain}(j)}{\overline{E_{rate}(j)}} \tag{4.4}$$

---

[1] The optimal value of this time period is out of the scope of this chapter.

[2] The value of α, we used is 0.25. This value is obtained from extensive simulation tests.

4.2.3 Computing the route establishment cost

Using the residual lifetime value computed using (4.4), each node number j computes a cost at each route request demand. This cost is defined as following:

$$cost_{node}^{(j)} = 1 - \frac{T_{lifetime}(j) * W_k}{T_{MaxLifetime}} \tag{4.5}$$

where $T_{MaxLifetime}$ is the maximum node's lifetime[3] and W_k is a multiplicative factor in the interval [0,1] defined for each energy level. Hence, we define four values of W_k referring to four energy intervals as shown in Table 4.1. The first one is from 50% to 100% of initial energy value. The nodes that have energy level in this interval are more favorite to participate in the route establishment. So they are assigned with the greatest weight which is equal to 1. The second interval is from 30% to 50%. The nodes in this interval are less favorite to participate in route establishment than the nodes in the first interval. They are assigned with a weight equal to 0.75. The third interval is from 10% to 30%. The nodes in this interval have a low energy level. The protocol should avoid the use of these nodes if there are other possibilities. The weight related to this interval is equal to 0.5. The last interval goes from 0% to 10%. In the route establishment scheme that we consider, these nodes are strongly avoided. Indeed, they are assigned with the lowest weight which is equal to 0.25. We aim by this way to extend the lifetime of these nodes if there are some other cost-effective alternatives. This yielding small value weight for small interval and a great one for the largest interval.

Normalized remain energy= $\frac{E_{remain}(j)}{E_{max}}$with E_{max} is the maximum energy level	W_k
[0, 0.1[	0.25
[0.1, 0.3[	0.5
[0.3, 0.5[	0.75
[0.5, 1]	1

Table 4.1: The mapping between the normalized remain energy and W_k

Now we will explain the benefits and the reasons behind the using of the weights in the cost computation. The computed lifetime is an estimated value based on the mean energy consumption rate in the previous periods. This generated residual lifetime of nodes varies over time depending on energy consumption rate. So there are some cases where nodes have less energy but generate a longer lifetime because they did not participate in data forwarding process in the previous periods. Hence, we should take about their energy level as a second criteria to generate their cost for route establishment. We give the following example to explain these cases. In Figure 4.1, node S would communicate with node D. It sends a RREQ to A and B. Nodes A and B have the following values for residual lifetime 65 sec and 62 sec, respectively and their energy levels are equal to 20W and 50W, respectively (assume for this example that the initial energy level is equal to 100W). Each node receiving the RREQ packet computes its own cost and includes it in the RREQ packet before forwarding it to S. If we do not use the weight in the cost computation and we only consider the residual lifetime value, the communication between nodes S and D will be through A. However, even A has a longer lifetime than B (which can be explained by the fact that A did not participate a lot in

---

[3] Without loss of generality the value of $T_{MaxLifetime}$is assumed to be the same for all nodes.

communications in the last periods), it has an energy level less than that of B. Hence, it is more fair to let B participating in route establishment than A.

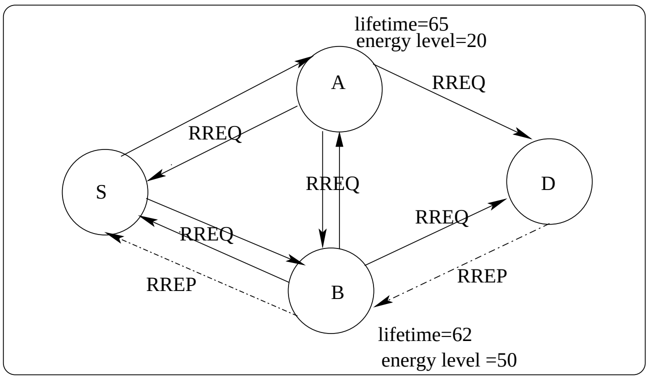

Figure 4.1: An example to show the benefits of introducing the weights when computing the cost of using a node in a route

4.2.4 Integration of our scheme in the basic AODV protocol

As described in Section 4.1, the basic AODV routing protocol uses minimum hop parameter to establish routes between sources and destinations. However, considering the new energy-constraint metric, we follow a new route discovery scheme. Indeed, routes are established based on the value of the cost of the residual lifetime ($cost_{node}^{(j)}$) defined in (4.5). Each RREQ packet includes the sum of costs of the traversed nodes in the used path. Then each node maintains for each reverse route the mean cost in the routing table entry and routes are built based on the minimum of the mean costs of all possible paths which is defined as follows:

$$\overline{cost_{path}} = \frac{\sum_{i=1}^{i=\#int\_nodes} cost_{node}^{(i)}}{\#int\_nodes} \tag{4.6}$$

where $\#int\_nodes$ is the number of intermediate nodes. Hence, each routing entry in the routing table is extended with the fields shown in bold style in Table 4.2.

Node	Next hop	Seq #	Hop count	**Mean cost**	**Min lifetime**	...

Table 4.2: New fields (in bold) of the AODV table entry

Furthermore, this new design requires that nodes act at all duplicate RREQ packets to select the maximum route lifetime that can be founded using 4.6. However, establishing routes without considering minimum hop count metric, as used in the basic algorithm, could increase the latency and possibly add more routing packets overhead. Therefore, our algorithm should design the energy conserving concept while achieving a trade-off between extending route lifetime and data-delivery quality in terms of the end-to-end delay. To this end, routes are updated regarding the difference on the number of hops between the new route and the previous one. This means that the route is changed when a new valid one which has a lower $\overline{cost_{path}}$ is detected and the extra number of hops in the new path could not be more than a given number of hops comparing to the number of hops in the old one in order to not affect other QoS metrics like the end-to-end delay. We fixed this number to 3 to take into account the exposed and hidden problems that may occur when selecting a longest path with a better residual lifetime. This condition is set in order to deal with the following case that is more likely to happen: for example, given two routes $path_i$ and $path_j$ that can connect the source S to the destination D, where $\overline{cost_{path_i}} < \overline{cost_{path_j}}$. The number of intermediate nodes between S and D using $path_i$ and $path_j$ are respectively in_i and in_j, where $in_i + 4 < in_j$. When applying the route selection procedure described above, route $path_j$ will be selected without taking into account the difference between the number of hops, which is clearly wrong. Indeed, adding lot of nodes to establish a new route because the new route will have a longer expected lifetime (lower $\overline{cost_{path}}$), may have a bad effect on the total network lifetime as well as the end-to-end delay.

Moreover, nodes that receive an RREP routing control packet includes the minimum residual lifetime in the message before sending it along the reversed path. This allows the source node to avoid when possible route failure and so it observes during data communication a timer that is compared with the minimum residual lifetime in the route as depicted in Table 4.2. The source sends a RREQ packet before this minimum residual lifetime of the route expires. Furthermore, when we have two routes toward the same destination with the same cost, we consider the route with maximum minimum residual lifetime. By this way, we ensure that the connectivity in a network is maintained as long as possible.

4.3 Performance evaluation

We have implemented our mechanism in the ns-2 simulator by extending the AODV protocol to support our energy consumption rate-based algorithm. We report in this section the large set of simulations we have done for various network topologies and scenarios. We also provide an analysis of the obtained performance results.

The energy model used bears similarities to earlier studies [24]. It is assumed that the radio interface, when powered on, consumes 1.15W when listening to the channel for any incoming packet, 1.2W while actually receiving a packet and 1.6W while transmitting a packet.

4.3.1 Simulation scenarios

The 50 nodes used in our simulations move in an area of 1500x300 according to a random waypoint mobility model as described in [76]. The radio model is very similar to the first generation WaveLAN radios with nominal radio range of 250m. The nominal bit rate is 2 Mbps. In this mobility model each node moves toward a random destination and pauses for certain time after reaching the destination before moving again. In our simulations, the nodes move at a maximum speed of 20m/sec. The pause times are varied to simulate different degrees of mobility. The sources start sending traffics at random times toward the beginning of the simulation and stay active throughout. The sources are CBR (Constant Bit Rate) and generate UDP packets at 4 packets/sec, each packet being 512 bytes. Each simulation is running for 900 seconds simulated time. Each point in the plotted results represents an average of ten simulation runs with different random mobility scenarios, while the error bars represent a 95 % confidence interval.

4.3.2 Simulation metrics

We analyze several QoS metrics to evaluate the performance of our approach and we compare results with the basic AODV protocol. The following metrics are used for the performance evaluation study:

- **Gain on Packet Delivery Ratio (GPDR):** This metric measures the gain (in %) on the packet delivery fraction at the end of simulations of our new mechanism E-AODV, compared with the basic AODV protocol. Note that, the delivery fraction is measured as the ratio of the number of data packets delivered to the destination and the number of data packets sent by the source.
- **Routing overhead:** It is the total number of bytes of transmitted routing control packets.
- **Average delay:** It is the average delay of all the flows computed at different stations. The average delay is used to evaluate how well the schemes can accommodate real-time flows.

In order to show the gain on energy and the effect on network connectivity and thus the useful lifetime, we evaluate the following metrics:

- **Gain on remaining energy:** This metric stands for the gain (in %) on the total remain energy at the end of simulations of our new mechanism E-AODV, compared with the basic AODV protocol.
- **The number of dead nodes during simulation**: This metric allows to have an idea on how soon nodes are dying out of power and how many nodes are dead (i.e., have zero energy) during simulation. It also measures implicitly *the time to first dead node* which provides an information about the starting time after which the nodes start to die.

4.3.3 Simulation results and analysis

We present in this subsection the performance of the basic AODV and our energy consumption speed-based routing algorithm applied to AODV (E-AODV) for the various metrics presented above. We vary the number of traffic sources and pause times to reflect various loads and mobility[4].

In Figures 4.2, 4.3, and 4.4, we plot the mean delay of our new mechanism and the basic AODV routing protocol. This metric is improved with 20 and 30 sources which demonstrates the efficiency of re-routing based on energy consumption rate. More precisely, a low energy consumption rate means that the node does not participate heavily in data transmissions (sending, receiving, and forwarding packets). This lets packets follow routes having a better expected lifetime and so are less congested which leads to lower delay comparing to that obtained with minimum hop count-based routing. Additionally, the basic AODV does not take into account energy behavior and so it does not avoid nodes contributing more than other on data forwarding. Moreover, we remark that the improvement on delay increases with low network mobility as the basic AODV does not change routes frequently in the stationary network case. However, in such scenarios, our algorithm allows re-routing and refresh routes

---

[4] Note that pause time = 0 means constant movement and pause time = 900 sec means stationary network.

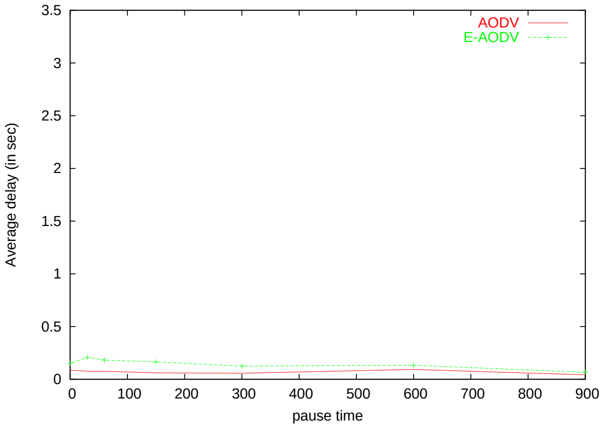

Figure 4.2: Average delay for 10 sources

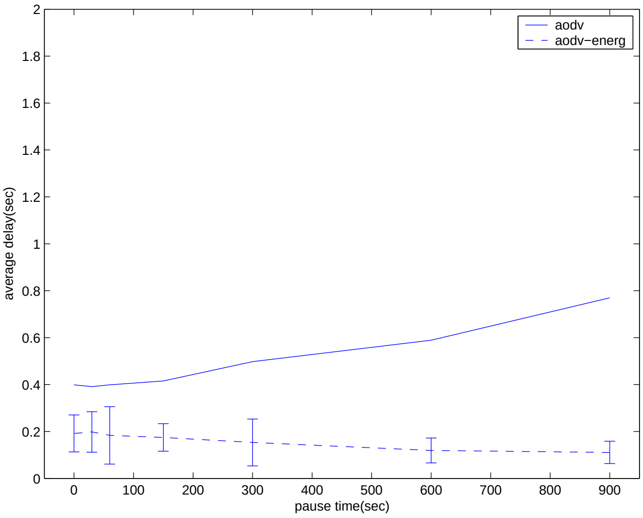

Figure 4.3: Average delay for 20 sources

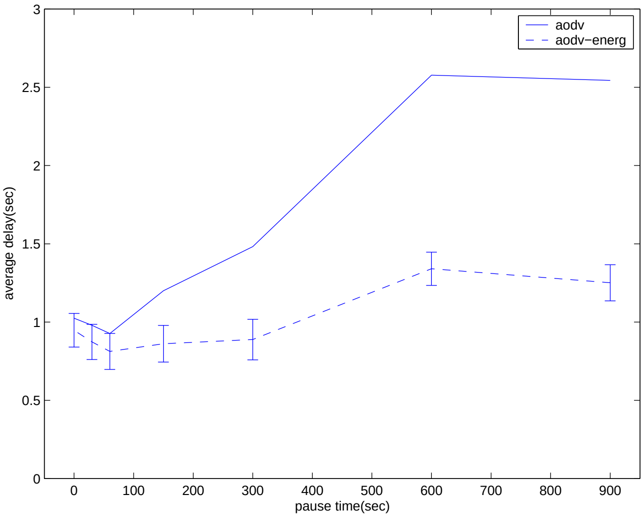

Figure 4.4: Average delay for 30 sources

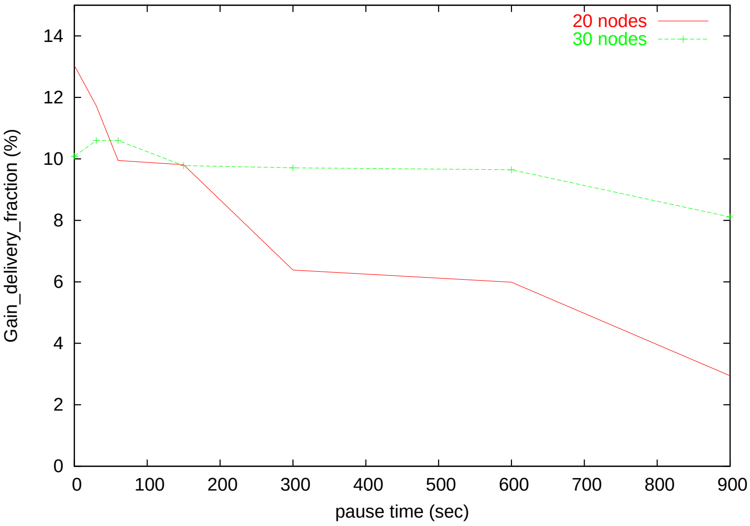

Figure 4.5: Gain on delivery fraction for 20 and 30 sources

including new nodes that have better quality than in the old routes which improves the end to end delay. Note that we mean by a good quality node, the node that is less loaded and has more energy than other nodes in the some other alternative route possibilities. Furthermore, we follow the RREQ broadcasting mechanism by source nodes since the minimum residual lifetime in the route will turn to zero soon. By this way, route failures are more avoided than in the basic AODV protocol. There are no improvement in the obtained average delay with low loads (10 sources). Indeed, the obtained results with our mechanism are a little greater than those of the basic AODV which could be explained by the fact that we use routing packet broadcasting in the research of other alternative routes more than in the basic AODV which might be high comparing to the low total load. The plot shown in Figure 4.6 enforces this claim given that the obtained route overhead of our mechanism is larger than in the original protocol. Moreover, this affects the packet delivery fraction which presents the same little difference between our algorithm and the basic one for the same reasons presented above. However, we obtain a good performance enhancement using 20 and 30 sources as shown in Figure 4.5[5]. The improvement attempts more than 10 % for 30 sources and 13 % for 20 sources. These results are proven by the fact that routing overhead is low with our new mechanism as shown in Figures 4.7 and 4.8.

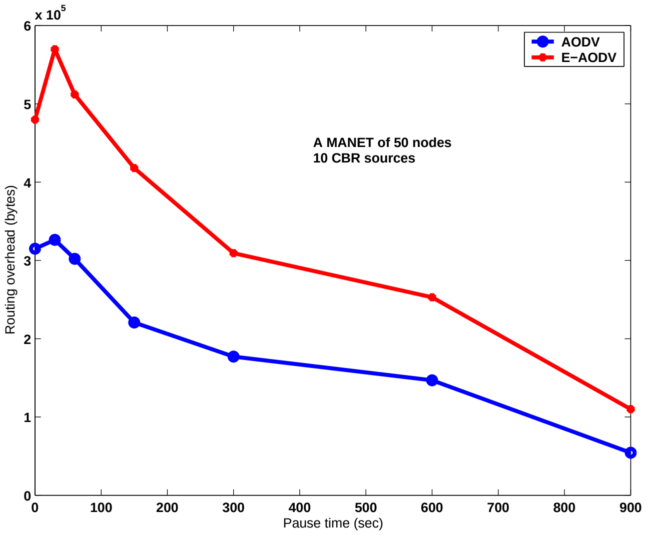

Figure 4.6: Routing overhead for 10 CBR sources

To demonstrate the efficiency of our scheme regarding to the energy consumption, we plot in Figure 4.9 the gain in the total remaining energy that is obtained at the end of simulation. Our algorithm presents an improvement for more than 26 % for 30 sources, more than 17 % for 20 sources, and for more than 10 % for 10 sources. We observe that the gain on energy increases with high load. This shows that our addressed mechanism give more benefits when communications increase. Moreover, this metric demonstrate how network longevity can be extended.

In Figure 4.10, we investigate the behavior of the number of death nodes during simulation. The plot presents how many nodes are dying as a function of simulation time for 20 sources and

---

[5] We did not show the results for 10 sources because it has no gain value.

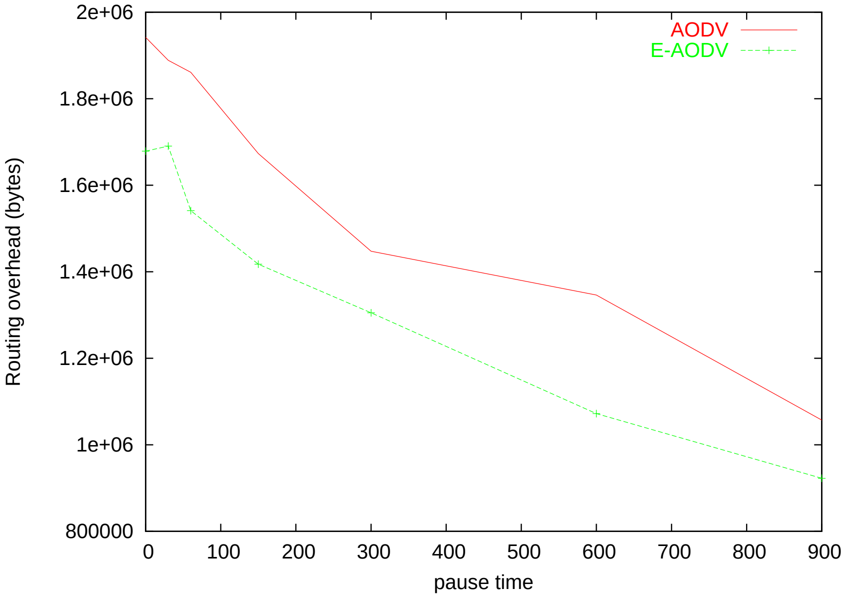

Figure 4.7: Routing overhead for 20 CBR sources

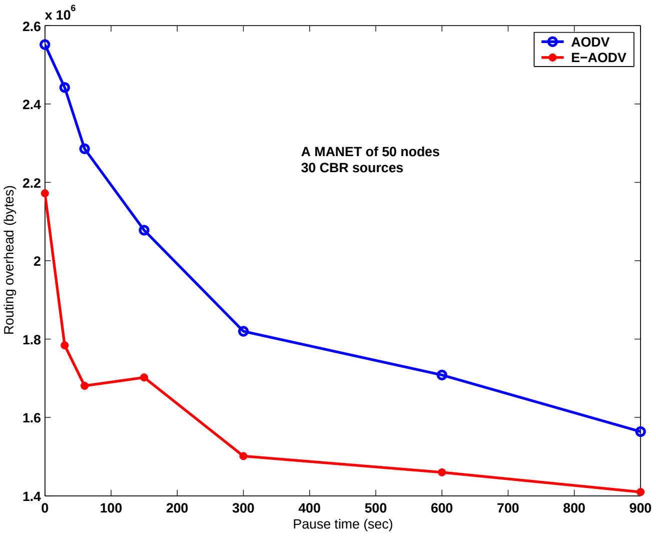

Figure 4.8: Routing overhead for 30 CBR sources

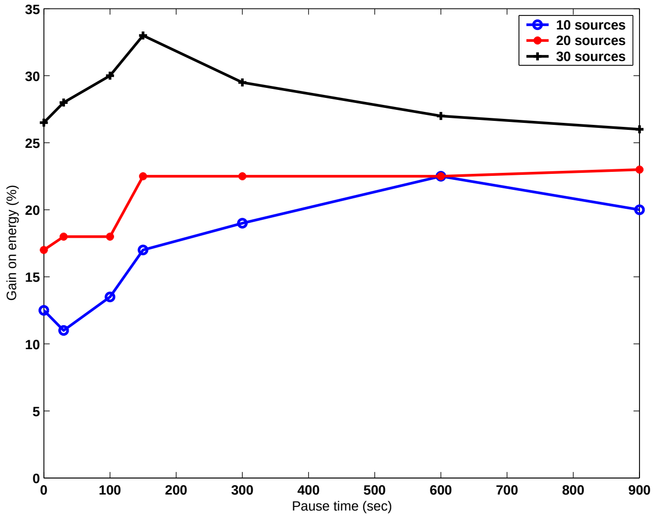

Figure 4.9: Gain on total remaining energy

pause time equal to zero. Indeed, this number has a great importance since it informs about network connectivity. On one hand, the more the large number of survived nodes is, the more routes would be established. On the other hand, when this number is small, communications between some nodes might be impossible. This is due to the fact that the nodes that could participate in the route built between these nodes (sources and destinations) are dead and so there no possibility that they communicate between each other even their battery allow a long lifetime. The experiment that we have done shows that our mechanism is able to keep more survived nodes than the basic AODV protocol. Indeed, there is about 100 second difference between the first died node in the two mechanisms. Moreover, we can see that our algorithm always outperforms the basic protocol during all simulation time. Indeed, the number of survived nodes in our mechanism is usually greater than in the basic protocol. For example, at 500 second, there are only five nodes that are dead for our mechanism, however the basic AODV protocol leads to more than 15 dead nodes.

This result shows the ability of our scheme to keep as smaller as possible the number of congested nodes participating in data forwarding so that to enhance the load balancing in the network. This objective is reached thanks to the addition of the energy consumption rate consideration in the routing process.

4.4 Extensions of OLSR to support the rate-based energy consumption mechanism

In this section we shortly provide an overview of the OLSR (Optimized Link State Routing) routing protocol. Then, we describe how our energy-based routing scheme can be implemented in OLSR in order to enhance the selection of the MPR (MultiPoint Relay) list and the determination of the routes.

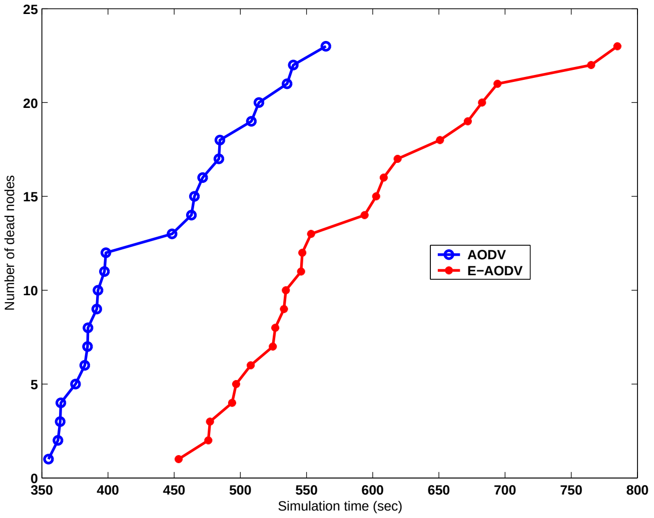

Figure 4.10: Number of dead nodes for 20 sources (pause time = 0)

4.4.1 An overview of OLSR

OLSR is a table-driven (proactive) routing protocol detailed in the IETF RFC 3626 [16]. Each node in OLSR periodically constructs and maintains the set of neighbors that can be reached in 1-hop and 2-hops. Based on this, the dedicated MPR algorithm minimizes the number of active relays needed to cover all 2-hops neighbors of each node.

In OLSR, a node forwards a packet if it has been elected as MPR by the sender node. In order to maintain its routing table, OLSR periodically transmit link state information over the MPR backbone (or MPR graph) which consists on the network of nodes participating on the broadcasting of control packets to the two neighbourhood. The link state information is included in the Traffic Control message which contains the list of neighbours of each node. Upon a convergence, an active route is created at each node to reach every potential destination node in the network. The convergence is reached when every node has a global view about the topology of the network.

4.4.2 Enhancement of the MPR election phase

We extend the algorithm used to select MPR nodes. Hence, the energy consumption rate is considered in the MPR selection process. Indeed, MPR nodes that have the better metric computed according to expressions in Section 4.2, are elected. This procedure is done by including the energy information in the periodic Hello message exchanged between nodes in the same neighborhood. By selecting the nodes in the MPR list those with the highest expected residual lifetime, we ensure that the MPR graph is the stable one among all possibilities.

4.4.3 Enhancement of the route determination phase

The selection of a path between each pair of nodes is done through the exchange of Traffic Control (TC) messages. In order to take into account the energy level of nodes, we include this information in the TC message. Each node can then build a weighted graph representing the network. The weights correspond to the residual lifetime of the link between each two nodes in the transmission range of each other. A node is then capable to determine the path which has the highest expected lifetime by processing all possible paths toward the destination.

4.5 Related work

Energy coming to be constraining factor for many mobile systems. There are several works that have been addressed to deal with battery usage conservation problem. As mentioned above the works that have been presented in this area can be divided into three main classes.

The first class of approaches, in which our work can be classified, focuses on power aware routing. There are many routing algorithms based on energy conserving mechanisms, that have been described for ad hoc networks. Routing based on expected node lifetime (i.e. "rate based") is a very frequently used idea and there are many papers about this [27, 41, 68]. In [68], the authors propose a max-min algorithm and present some theoretical properties to show how to choose routes while running minimum energy. Despite of the great benefits of that mechanism, it requires periodically an exchange of the energy consumption value information of all the nodes in the network. In one hand, this adds more overhead to obtain this accurate information and consumes more energy. In the other hand, this design will be hard to address in a large network with high number of nodes. However, our algorithm is designed for on demand and doesn't require to have a global view of the network. In [61], different metrics as energy consumed per-packet, time to network partition, variance in battery life of nodes, are presented to introduce power-aware routing. Maximizing network lifetime based on the knowledge of the message rate has been also investigated in [12]. The mechanism describes an hierarchical routing methods. The authors propose an algorithmic approach that aims to reduce the battery consumption. This approach defines a class of flow augmentation algorithms coupled with flow redirection. Unlike the conventional approach of minimizing the cost of the route from a source to a given destination, the strategy here was geared toward balancing the battery usage among the nodes in the network in proportion to their energy reserves. The presented algorithms are only centralized and there is no solution for distributed network. In [27], an energy-aware routing algorithm that follows a new route discovery scheme, is described. This scheme is based on the remaining energy value, number of neighbors and mainly a sleep-active node model. Using current energy level in the route establishment algorithm, it is not sufficient to achieve a maximum route lifetime since the behavior of energy consumption is very important to a better estimation of the residual lifetime of nodes. This sleep-active node model, switches off the radio interfaces when nodes are idle. Our algorithm does not apply this model. That is, switching off radio interface can interfere with routing, as routes cannot be formed via sleeping nodes. This also can result in longer routes or in failure of route discovery. The latter may result in longer delays or lost packets due to buffer overflows at the source. To find a good trade off between energy conservation and good performance, it requires more energy consumption in order to compute the requested paths. Moreover, improvement performance remains restricted to some limited scenarios. The algorithm described in [68], considers the remaining battery capacity of each node as metric. It chooses the route with minimal total transmission power if all nodes in the route have remaining battery capacities higher than a threshold; otherwise routes including nodes with the lowest remaining battery capacities are avoided. This requires that the lowest energy node information should be provided. No previous knowledge of the information generation rate is required. However, it is necessary to evaluate alternative routes for each destination and to know the minimum value of the battery capacities, dynamically changing, of the hosts belonging to each route. This design is different from our work. That is we consider in the cost computation, the expected residual lifetime of each node, according to the communication experiences of a node in the previous period. Indeed, a giving two nodes that have the same energy level, might not have the same residual lifetime if they do not participate in communication at the same rate. Moreover, [68] considers battery operating devices with adjustable transmission energy that we do not consider in our work.

The proposal in [13], called “span”, presents a conserving energy scheme based on sleep-active model. A local election is used to elect coordinators. The elected coordinators are connected by paths in such communication can be possible between large number of nodes. Each node has at least one coordinator neighbor. The coordinators remain awake at all times and therefore form a low latency routing backbone for the network. The Span coordinator election algorithm is intended to approximate a minimal capacity-preserving set of coordinators. One of the design goals of our proposal is to be distributed without relying on a coordinator which could be a single point of failure.

The second class of energy conservation that we introduced above, proposes some modifications in 802.11 MAC layer to achieve power saving. In [21], the authors propose the Power Controlled Multiple Access (PCMA) mechanism, in order to enhance medium utilization by reducing the transmit power of nodes as well as possible. Moreover, the described mechanisms in [22, 60]suggest some strategies to power off nodes during idle time in order to save power. Several parameters are used to decide when a node should transit into a specific mode as location information [22], or number of neighbors around the node obtained by the feedback of broadcasting [43]. In [60], nodes can be in two modes: active mode or sleep mode to enable energy conservation depending on communication behavior. A mobile should power itself off in two cases: the first one is when it has no packets to transmit and a neighbor begins transmitting a packet not destined for it. The second case is when it does have packets to transmit but at least one neighbor-pair is communicating. Each mobile determines the length of time that it should be powered off through the use of a probe protocol, the details of which are available in [60]. However, the difficulties of considering the sleep-active mode is what is the best way to transit between the two modes (sleep active) without affecting the performance, as explained above. Furthermore, combining power optimization and efficient power aware routing is very interesting.

The third class deals with the issues related to power control. They address some techniques of tuning the transmit level of every packets according to many giving parameters as neighbor number, end-to-end network throughput, topology information, and other techniques based on routing concepts. In [50], the protocol proposes that a node can adjust its transmit power based on a limit number of its neighbors that it should has. In [22], the authors suggest some methods based on power control that aim to optimize the global throughput neighbor a long the path by observing the degree of each node participating in route establishment. Comparing to our work, we did not consider optimizing power while sending messages. The goal of power control mechanisms addressed in MAC layer is to minimize as possible as, the amount of contentions between nodes and so the interference. While these schemes have some benefits that concern the enhancement of medium utilization, reducing collision, and improving the total throughput, they have also some disadvantages. Indeed, making the control power at MAC layer to control the number of neighbors in order to reduce the interference, can affect the role of routing protocol. Hence, routing protocols couldn’t have the possibility to establish routes with the optimal nodes since every time the next hop is depending on MAC layer transmit power. To solve this problem, power control is coupled with routing to provide the control of topology [35]. The authors propose three algorithms based on power control, routing and clustering in Ad hoc network. That is, a solution for implementing power control at network layer is described. The results of that work provide a good energy conservation. However, some problems are addressed for future works as how to achieve a trade-off between power saving and QoS issues.

There are other parameters in the other layers that we can take into account, to adjust the transmit power, conserve energy, and provide QoS guaranties specially for real time applications. This will let the power saving a cross-layer design that aims to enhance the performance and let the energy consumption close to the minimum. The challenge here is to achieve a compromise between complexity and QoS enhancement. Indeed, the complexity should to be reduced because it consumes power as well as communications, some times much more.

4.6 Chapter summary

On-demand routing protocols are useful for mobile ad hoc network environment for their low routing overheads. However, they require to consider the reasons for link failure to improve its performance. Link failure stems from node mobility and lack of network resources. Therefore, it is essential to capture the aforesaid characteristics to identify the quality of links. Furthermore, the routing protocols that support QoS must be adaptive to cope with the time-varying topology and time-varying network resources. For instance, it is possible that a route that was earlier found to meet certain QoS requirements no longer does so, due to the dynamic nature of the topology. In such a case, it is important that the network intelligently adapts the session to its new and changed conditions. Indeed, it is not enough to find a shortest path but also with available resources as battery. If battery energy is not taken into consideration in their design, it may lead to premature depletion of some nodes' battery leading to early network partitioning. Since computing complexity consumes power as well as communications, we proposed in this chapter a simple and efficient energy consumption rate based algorithm to establish routes between sources and destinations.

Performance evaluation using ns-2 simulator shows that the longevity of the network can be extended by a significant amount. Overall, we conclude that our mechanism demonstrates significant benefits at high traffic and high mobility scenarios. We expect that these scenarios will be common in ad hoc networking applications. Even though we implemented the algorithm on AODV, the technique used is very generic and can be used with any on-demand protocol. Furthermore, we have also shown how our proposed mechanism can be added in the OLSR proactive MANET routing protocol.

Indeed, a large literature presents several works that have been done to optimize power consumption in ad hoc network. That considerable research has been devoted to low-power design of the entire network protocol stack of wireless networks in an effort to enhance energy efficiency. Therefore, the key to energy conservation in wireless communications should be based on a coordination between all levels of the wireless protocol stack. One of main problems of wireless links is that all the nodes compete for the resources and channel access without taking into account knowledge about neighbor communications. Moreover, how to address an efficient QoS model based on MAC layer mode access and other key metrics for real time applications.

The proposed mechanism described in this chapter can be classified as a *source-initiated and network-aided* approach because is the source which selects the path to use and initiate the route determination/change procedure but it needs information from intermediate nodes along the path toward the destination. The next chapter describes another enhancement for routing protocols in MANETs that also consider the energy consumption of nodes but it gives the possibility to intermediate nodes to enhance the data forwarding and not only the source. The proposal, which can be classified as a *source and network-initiated and assisted* approach, is based on a joint optimization of the routing and MAC protocols where the goal of the optimization is to achieve a good medium utilization, energy conservation, and provide a QoS enhancement of application performance.

Chapter 5

A cross-layer approach for efficient data forwarding in MANETs

In recent years, the research focused on adapting existing algorithms and protocols with respect to the layered architecture. However, the proposed solutions often do not provide adequate support for real-time applications and wireless links constraints. Indeed, there is a tight interdependence between layers in wireless networks, that promotes adaptability at all layers based on information exchange across communication layers.

In this chapter, we present a Cross-Layer Forwarding Strategy (CLFS). This strategy suggests to collaborate the routing and the MAC modules in order to optimize the data forwarding in MANETs. We apply this strategy to the AODV routing protocol resulting to a new routing protocol called F-AODV (Forwarding-AODV). This proposal aims to minimize the number of Forwarding Nodes (FNs) in the network. As a result, we decrease the contention amount and we improve the medium utilization. In addition, we extend the basic routing discovery procedure to include new features. The selection of the FNs is based on maximum battery level and queue occupancy. This information is injected into routing request and reply messages. Then, each node is able to select the FN that will participate in path establishment. In order to maintain a fair collaboration between wireless nodes, the forwarding procedure is dynamically distributed and assigned to nodes in the network. Through ns-2 simulations, we demonstrate that F-AODV has not only an efficient throughput, that could be further improved, but also achieves a high degree of fairness among applications.

The remainder of this chapter is organized as follows. In Section 5.1, we introduce the motivations behind our work by stressing the necessity of optimizing data forwarding in MANETs. The description of the proposed F-AODV cross-layer routing strategy is given in Section 5.2. Simulation methodology and performance evaluation of our proposal are detailed in Section 5.3. We devote Section 5.4 to the comparison between the E-AODV proposal described in the previous chapter and the F-AODV mechanism where cooperation between routing and MAC protocols is considered. Section 5.5 concludes this chapter by summarizing the outcomes and outlining future works.

5.1 Optimizing data forwarding in MANETs: why it is so important?

In MANET, network services are delivered by the cooperation of all nodes instead of predeployed facilities. Due to limited radio transmission range, data packets are usually forwarded by multiple intermediate nodes before they reach the destination. Therefore, packet transmission does not come for free. In addition to the bandwidth and computational cost, energy is spent by each Forwarding Node (FN). Consequently, when the number of FN increases in the same hop, the contention amount increases too, which affects applications' performances. We strongly believe that minimizing the number of FNs and optimizing their choice for each active session enhances the medium utilization, reduces the number of collisions, and increases the network lifetime. When a node is selected as a FN, it will certainly constantly consume more energy, compared to non forwarding nodes. If this trend continues, the nodes in the FN list will turn-off much earlier than the others and causes the disconnection of the network. To overcome this problem, we allow intermediate nodes to redirect the selected route on each hop and form cooperative coalitions on the fly. Hence, the forwarding activity is balanced between nodes which provides a fair resource consumption. An intermediate node between the source and the destination chooses one node among its neighbors toward the destination. This selection is based on battery level and queue occupancy of intermediate node between the source and the destination as well as the weight values (W_i) of these neighbors, which is assigned to each node i in the network according to their data load as we will detail later in the following section. These weights are used to tune and adapt the Contention Window (CW) and $TXOP$ duration MAC layer parameters. The main objective of this tuning process is to give a high medium access probability for the nodes in the FN set.

At this stage we would like to first stress that the proposed medium access control (MAC)-based performance studies revealed that battery capacity may not be efficient for achieving energy-based fairness and system longevity for wireless mobile multi-hop ad-hoc and sensor networks. However, energy conservation may be attained only if valuable MAC (and PHY) input is passed to the network layer. Indeed, most of the proposals compute global or local metrics which are used to make decisions for route establishment, scheduling, tuning transmission rate, tuning power transmission, etc. In the work presented in this chapter, we investigate a new MAC layer adaptation scheme using both TXOP (Transmission Opportunity) and CW (Contention Window) parameters. As we will show at the end of this chapter, the presented approach allows to select minimum forwarding nodes in the network and maintain a good packet delivery rate while still adapting to local characteristics variations.

5.2 Detailed description of F-AODV

5.2.1 Overview of F-AODV

We propose a forwarding scheme and adaptive transmission opportunity in wireless ad-hoc networks. The goal of our proposal is to achieve a good medium utilization while providing a good application performance. In order to decrease the contention amount, the routing layer selects a minimum number of FNs to set up communications. Then, the functionality of FNs is balanced among candidate nodes to achieve good resource management and fairness.

A critical issue in the selection of the proper broadcast and routing strategy in mobile ad hoc networks is optimizing energy consumption and extending network longevity while maintaining connectivity and satisfying latency constraints. We propose to consider both MAC queue length fraction and battery level to select the best FN among all possibilities. Each node incorporates the value of these two metrics in AODV periodic Hello messages. By this way, nodes are able to select the appropriate FN in a distributed way. Moreover, these accurate information allow nodes to manage intelligently their resource in terms of energy consumption and data processing complexity.

As the energy conservation and the network lifetime are critical issues in wireless ad-hoc networks, we take the amount of energy left at neighbor nodes into consideration when selecting one route from multiple paths. The selected node is then chosen for all possible communication set up that have to traverse that hop, while its queue level does not reach 90% of its maximum level. To this end, each node needs to report its energy and queue levels to its neighbors. The current (residual) energy level is normalized to the maximum battery capacity and scaled to 100. The normalized residual energy level helps to handle heterogeneous nodes with various battery capacities. An additional two bytes recording the energy and queue levels are piggybacked onto the Hello messages. Consequently, all the neighbors receive the packet and record the corresponding values. Then, the multi-path selection takes all the next hops from available paths, and checks the associated normalized remaining energy levels known to the node. The next hop with the highest energy level is selected. Then, the other hop possibilities are classified according to their battery occupancy. When a FN does not still have the larger energy level or its queue occupancy reach 90 % of its capacity, the packet forwarding process is balanced to the node of best quality in term of energy and queue occupancy. It is possible that the energy information collected at a node is not accurate. However, the promiscuous nature of wireless channel provides for each node a great opportunity of overhearing neighbor information up-to-date and accurate. The energy usage of a node indicates the amount of radio activities. Thus it could be regarded as an indication of traffic load at the node. While selecting the next hop according to the energy levels, load balancing among the neighbors is also achieved.

The throughput of data traversing a given node (called also the cross throughput), indicates how much a node participate to the data forwarding process. Indeed, loaded nodes suffer from high queue drops and so rapid resource degradation when they cannot access to the medium for a minimum duration. To overcome this problem, we adapt MAC layer parameters with respect to node's traffic load. Thus, TXOP duration and CW settings are tuned regarding to an assigned weight which is computed according to node's forwarding activities.

5.2.2 Computing useful metrics

In order to minimize the complexity of selection optimal forwarding nodes, each node has to periodically compute two useful parameters:

- **Normalized battery level**: The normalized battery level $E[E]$ is defined as the ratio of the remaining energy E_{remain} to the maximum energy level E_{max}: $E[E] = \frac{E_{remain}}{E_{max}} * 100$. To avoid frequent FN changes, we just consider absolute values. That is mean a node having 80.1 % or 80.5% of energy level values are considered 80 %.

- **Normalized queue occupancy**: The normalized buffer size $E[B]$ is defined as the ratio of the current buffer size B_{remain} to the maximum queue occupancy B_{max}: $E[B] = \frac{B_{remain}}{B_{max}} * 100$

To minimize the bias against transient behaviours, we use an estimator of Exponentially Weighted Moving Average (EWMA) to smooth the estimated values of both $E[E]$ and $E[B]$.

5.2.3 Computing nodes' weights

For each user session between a source and a destination, a path is established thanks to a (on-demand or proactive) routing protocol running at the network layer. Each intermediate node can construct the list of potential Forwarding Node (FNs) (called FNL - Forwarding Node List) toward the destination in its neighborhood by observing all received packets at the MAC layer. Indeed, each node checks whether the destination and source addresses included in the packet header, are already recorded in the FNL or not. Then, each node updates its FNL accordingly. There might be another way that could be used to determine the FNL by observing AODV RREP control packets. However, this method can not be accurate as intermediate nodes can reply to RREQ messages. As a result, nodes toward the destination could not have knowledge about this new communication set up.

The analysis of captured traffic by a node is used to estimate the traffic load generated by each neighbour and to evaluate its contribution to the total traffic load around the node number i. We define the traffic weight W_i of node i as follows

$$W_i = \frac{load_i}{\sum_{j=1}^{j=N_i} load_j}$$

where $load_j$is the traffic load generated by the neighbor node number j.

5.2.4 Route establishment scheme

We extend AODV routing protocol [43] to support our proposal of enhancing the data forwarding process. When a source node tries to build a path from itself to a certain destination, it generates a Route REQuest packet (RREQ). Besides those fields specified in AODV, it also sets the QoS indicators $E[E]$ and $E[B]$ that we have defined above. Then, each node maintains for each reverse route the costs in the routing table entry. The routes are built based on maximum energy level. The route establishment can be done with different paths. Among the next hop possibilities, we select the nodes which are recorded in the FNL and have the highest weight. Nodes in FNL, which have a queue level that exceeds 90% of its initial maximum value, are not chosen because they are overloaded.

We enhance the route establishment procedure by including role retention between all potential nodes in the FNL. Periodically, every node broadcasts Hello messages, which include $E[E]$ and $E[B]$ parameters. A node compares its own values with those of its neighbors. If it is among the top richest nodes in its neighborhood, it may claim its ability to forward packets and updates its routing table accordingly. Thus, we allow intermediate nodes to redirect the route on each hop and form cooperative coalitions on the fly.

For compatibility reason, all other routing operations remain similar to AODV. Under different circumstances, the network may choose either the original or the extended version of AODV without any additional cost.

5.2.5 Adaptation of MAC layer parameters

Many results have highlighted the importance of considering the impact of setting MAC layer's parameters on the application performance [47]. We aim to provide better and efficient medium utilization for loaded nodes while achieving a good QoS support and fair channel access.

5.2.5.1 Adapting the TXOP duration

We believe that service differentiation based on adaptive packet bursting can provide a good performance in a distributed network where contention is costly. To efficiently use the $TXOP$ option, we propose that nodes in the network adopt an adaptive transmission opportunity duration based on the node congestion level. Under some ergodicity assumption, the throughput of a node i for one transmission can be expressed as: $th[i] = \frac{Psize[i]}{Tx[i]}$, where:

$$\begin{aligned} Tx[i] &= cE[coll] * (E[backoff] * T_{slot} + T_{data}[i] + AIFS[i] + T_{slot}) + T_{rts} \\ &+ T_{cts} + AIFS[i] + T_{data}[i] + T_{ack} + E[backoff] * T_{slot} + 3 * SIFS + T_{slot} \end{aligned} \tag{5.1}$$

is the average virtual transmission time of a packet by the MAC layer. $T_{data}[i]$ is the transmission time over the wireless channel of a packet of a size equal to $Psize[i]$, $E[coll]$ is the average number of collisions in a virtual transmission time (or a Virtual Transmission Cycle VTC), $E[backoff]$ is the average number of idle slots resulting from the queue's backoff for each contention period, T_{ack} is the acknowledgment's transmission time, and T_{slot} is the SlotTime, which depends on the physical layer type. The above expression of the throughput (Eq. (5.1)) illustrates that the optimal case is reached when a successful packet transmission is followed by another successful packet transmission without any collisions or idle time loss, i.e. $E[coll] = E[backoff] = 0$. Then, we can write:

$$Tx[i]_{(E[coll]=0)} = T_{rts} + T_{cts} + AIFS[i] + T_{data}[i] + T_{ack} + 3 * SIFS + T_{slot} \tag{5.2}$$

The maximum throughput of node number i is then:

$$th_{max}[i] = \frac{Psize[i]}{Tx[i]_{(E[coll]=0)}} \tag{5.3}$$

We observe from (5.3), that the throughput depends usually on the time elapsed to send RTS/CTS packets. The CFB (Collision Free Burst) is introduced to send packets without using RTS/CTS frames as shown in the following equation:

$$TXOP_{dur}[i] = T_{rts} + T_{cts} + 2 * SIFS + (K_i + 1)\,(T_{data}[i] + T_{ack} + 2 * SIFS) \tag{5.4}$$

Where K_i is the maximum number of packets that can be sent without using RTS/CTS frames. In the basic EDCA scheme, the $TXOP$ duration is initially fixed for all priority queues. This could not be efficient when node's traffic load and medium characteristics vary dynamically. In this chapter, we aim to provide medium access ability for FN nodes as they are generally very loaded. For this purpose, we consider the weight parameter W_i introduced in Subsection 5.2.3 to tune the maximum CFB duration for each node during a VTC. As a result:

$$TXOP_{dur}[i] = \frac{W_i}{\sum_{j=1}^{j=n} W_j} * VTC \tag{5.5}$$

where n is the number of neighbors of node number i. From (5.4) and (5.5), we can write:

$$VTC = \frac{\sum_{j=1}^{j=n} W_j}{W_i} * \left(T_{rts/cts} + K_i * T_{data}[i]\right) \tag{5.6}$$

To determine the value of VTC, we have only to set the value of K_i packets that the most congested FN (that has the highest W_i) has to sent in order to avoid packet drops. Furthermore, we can set VTC as follows:

$$VTC = (\sum_{j=1}^{j=n} K_j T_{data}[j]) + E[backoff]$$

where n is the number of nodes in the same neighborhood, and $E[backoff]$ parameter is the average backoff value estimated as follows:

$$E[backoff] = \frac{CW + 2}{2} \tag{5.7}$$

5.2.5.2 Adapting the contention window max and min values

In the previous QoS enhancement studies, the contention window size was used as the main differentiation mechanism [53]. Authors have considered the medium utilization and the amount of collision, as the main parameters to enable differentiation and enhance the application performance. In our work, we follow another method that is based on the $TXOP$ duration value obtained by the above analysis for each neighbor node. If one node successes to transmit packets, it adjusts its CW_{min} as follow:

$$CW_{min}[i] = \frac{\sum_{j \neq i} K_j * Tx[j]}{2} + 1 \tag{5.8}$$

$K_j * Tx[j]$ is dynamically settled according to traffic load of each node neighbors. Moreover, it can achieve differentiation because we allow more transmission opportunity duration for the most loaded nodes which involves to less CW_{min}. Hence, if we have node f, which is most loaded, $K_f * Tx[f]$ is the maximum of all $K_j * Tx[j]$ for $j \neq f$. Then, for all neighbors p of node f, with $p \neq f$, $CW_{min}[f] \leq CW_{min}[p]$.

Following the same reasons, we set the CW_{max}as follow:

$$CW_{max}[i] = (\sum_{j \neq i} K_i * Tx[j]) + 1 \tag{5.9}$$

5.3 Simulation and performance analysis

We implemented our proposal in ns-2 network simulator [49]. We have extended the AODV protocol and EDCA scheme to support our cross-layer algorithm. We report in this section the simulation results. We also provide an analysis of the obtained performance.

5.3.1 Scenario description

We simulated Local Area Network (LAN) and multihop network scenarios to show the performance of our approach. Our simulation uses video traffics to evaluate QoS support. Each active station generates packets of size equal to 1280 bytes each 10 ms which corresponds to an overall sending rate of 1024 Kbps. To increase the load of the system, we increase the number of parallel flows. Moreover, we consider an arbitrary start and end times of communications to show how the proposed model could be adapted to the dynamic network load. The radio model is very similar to the first generation WaveLAN radios with nominal radio range of 250m.

In the following simulations, we assume that each wireless station operates at IEEE 802.11a PHY mode-6, see network parameters shown in Table 5.1.

SIFS	$16\mu s$
DIFS	$34\mu s$
TXOP	$0.0015ms$
$CWmin$	15
$CWmax$	**31**
Data rate	36 Mbps
Slot_time	$9\mu s$
CCA Time	$3\mu s$
Preamble Length	$20\mu s$
RxTxTurnaround Time	$1\mu s$
PLCP header Length	$4\mu s$

Table 5.1: IEEE 802.11a PHY/MAC parameters used in simulation

5.3.2 Performance metrics

To evaluate the performance of the different schemes, the following metrics are used:

- **Goodput:** this metric indicates the total bytes that have been successfully delivered to the destination nodes.
- **Average end-to-end delay:** it is the mean end-to-end delay of all competing flows. The average delay is used to evaluate how well is the global delay performance of different flows.
- **Latency distribution:** latency distribution allows to measure the percentage of packets that have latency less than the maximum delay required by the applications. Obviously, real-time flows require both low average delay and bounded delay jitter.

5.3.3 Performance study in wireless LANs

We consider a simple scenario to show how to enhance the performance when considering the adaptive MAC layer parametrization that we have proposed in Section 5.2.5 (called hereafter Adaptive MAC Parametrization - AMP). The scenario consists of three nodes in the same transmission range of each other. The first and the second node send respectively 10 and 5 video flows to node 3 through the access point. Figure 5.1 and Figure 5.2 show the goodput and the delay distribution results. Indeed, our proposal AMP provides significantly more than 10% of total throughput compared to the basic EDCA. Moreover, the coordinates (0.08, 80) and (0.08, 50) imply that 80% of packets have a delay lower than 0.08ms. However, for the original EDCA scheme, only 50% have delay less than 0.08ms. The adaptive setting of CW values and $TXOP$ durations for all flows reduce the number of collisions that allows a good medium utilization. Hence, the throughput is improved and the delay is well minimized.

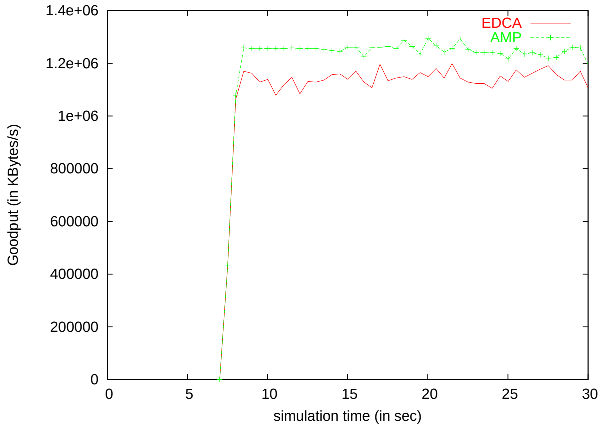

Figure 5.1: Goodput results for WLANs

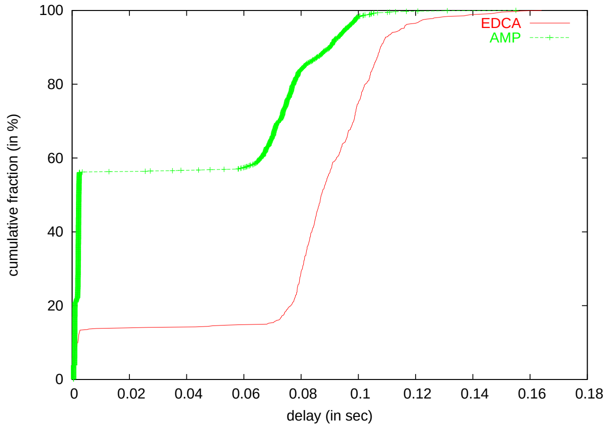

Figure 5.2: Delay distribution results for WLANs

5.3.4 Performance study in multihop wireless networks

The simulated scenarios consist of 9 nodes located in a uniform distribution within an area of 1500x300 forming a multi-hop network. The nodes send video traffics to each other.

Figure 5.3 and Figure 5.4 show the throughput and the average delay performance, respectively for our proposed model and the basic protocol. We observe that F-AODV (the modified

AODV protocol enhanced with the AMP scheme) provides significantly more total throughput compared to the basic EDCA-AODV (Figure 5.3). There is more than 50% throughput gain when using the layer cooperation. This result demonstrates the benefits of reducing the FN number while adapting the MAC layer parameter to the traffic load. Indeed, as the load becomes heavier, the level of contention will not increase proportionally. By this way, the medium utilization is improved. In the EDCA scheme, the high contention level affects the throughput performance due to the fact that its parameter setting is static and cannot be adaptive to the traffic load.

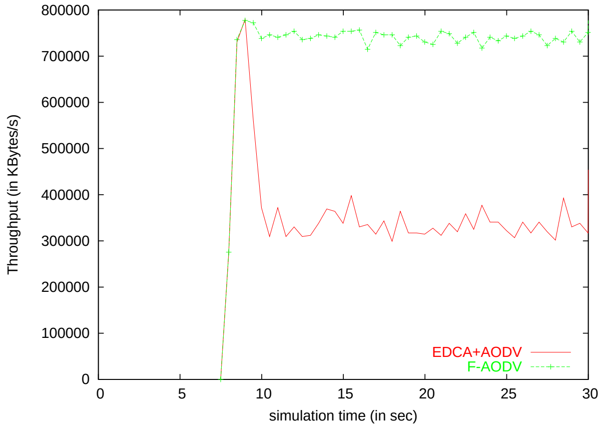

Figure 5.3: Total throughput results for multihop wireless networks

As Figure 5.4 illustrates at the beginning of simulation, F-AODV achieves higher delay than the basic approach. The tight between the two curves is minimized when all nodes starts the transmission. One reason behind this phenomenon is that when routes are not yet settled, the selection of FN takes some delay before route establishment. However, the average delay of F-AODV decreases over the time and therefore achieves the same performance as EDCA-AODV combination. This behavior could not affect the performance of our scheme as the delay tends to be improved once the selection of FN is done.

The main advantage of our proposal is that, it does not only enhance the global throughput of video applications, but also it can ensure a good fairness degree among all the flows.

We define the **Fairness Index** (FI) as follows:

$$FI = \frac{\left(\sum_{i=1}^{i=n} T_i\right)^2}{n \sum_{i=1}^{i=n} (T_i)^2} \tag{5.10}$$

where n is the number of flows, and T_i is the throughput of flow i. Note that $FI \leq 1$, and it is equal to 1 if all T_i are equal, which corresponds to the highest degree of fairness between the different users. As shown in Figure 5.5, our scheme is always fairer than EDCA-AODV basic protocols. The main reason comes from the fact that the AMP scheme adjusts the CW value regarding to the traffic load of neighbors, which allows a fair medium access probability. Moreover, the adaptation of the $TXOP$ duration allocates to nodes a relative medium occupation according to their load. This provides better fairness between different users since the queues of the different users will be transmitting almost all the time when efficiently adapting

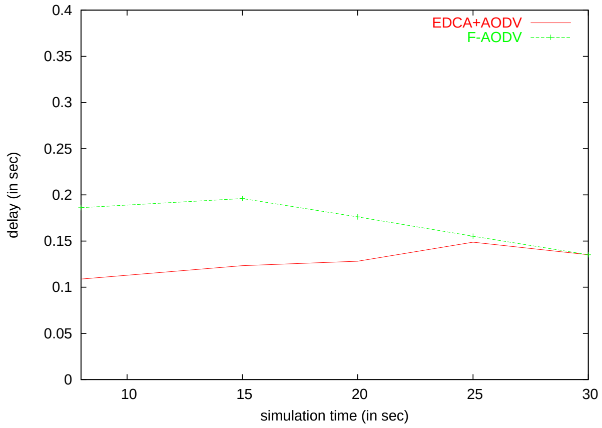

Figure 5.4: Average end-to-end delay results for multihop wireless networks

MAC parameters. This is not the case when using basic EDCA-AODV protocols because there is no consideration of node congestion, which leads to an unfair medium access.

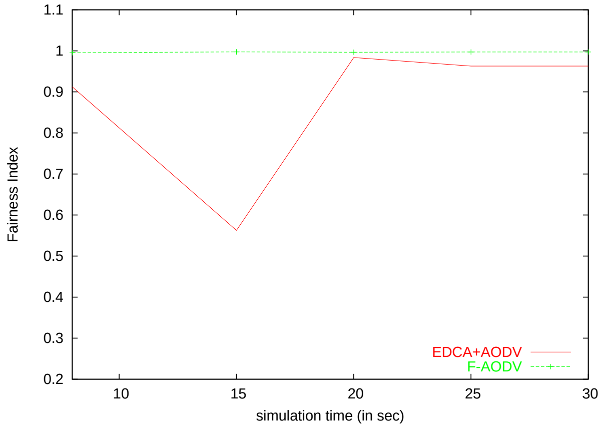

Figure 5.5: Fairness Index

5.4 An extensive quantitative comparison of basic AODV, E-AODV and F-AODV

The objective of the next set of simulations is to compare the performance of E-AODV, proposed in the previous chapter, F-AODV, presented in this chapter, and the basic AODV protocol. We aim to evaluate the benefits of considering inter-layer cooperation and adaptation using several network scenarios. Recall that E-AODV considers only energy rate consumption metric in route establishment scheme. However, F-AODV ensures further, MAC layer adaptation for congested nodes.

We consider squared area of 1000m x 1000m. The different simulation parameters are summarised in Table 5.2. Each plotted point is the average of 10 simulation iterations, while the error bars represent a 95% confidence interval.

We measured several significant metrics for MANETs: Packet Delivery Ratio (PDR), Routing Overhead (RO), Average Delay (AD), and Route Error Rate (RER).

We study the effect of the node density, the influence of the initial speed variation and the data traffic rate on the performance of the E-AODV, F-AODV, and the basic AODV protocols.

Simulation time	$900s$
traffic	CBR, 4pkt/s
Packet size	512 bytes
Mac rate	2 Mbps
Initial speed	$Sp_{min} = 5m/s$, $Sp_{max} = 25m/s$
Speed	Uniform
Density	#nodes $*\frac{\Pi * range^2}{X_{dim} * Y_{dim}}$
Range	$250m$
Simulation area	$1000*1000m$
#nodes	40, 50, 60, 70, 80
Confidence Interval	95%

Table 5.2: Simulation parameters

5.4.1 Impact of network density

We illustrate, on the first set of simulations, the influence of node density (in terms of average number of neighbors per node) computed as shown in Table 5.2, on E-AODV, F-AODV, and the basic AODV performance. The corresponding results are presented in Figures 5.6, 5.8, 5.7, and 5.9.

Figure 5.6, shows the obtained packet delivery ratio results. The general trend of all curves is a decrease in PDR with high node density. This is mainly due to higher probability of collisions and channel contention. We observe that F-AODV outperforms both AODV, and E-AODV especially at high node density. The improvement achieved by F-AODV, compared to AODV, is about 9% at low node density and about 14% when the node density increases. E-AODV and F-AODV exhibit similar trends at low node density. However, the obtained performance by F-AODV becomes higher than E-AODV when the node density increases. This behaviour is explained by the fact that F-AODV minimizes the number of nodes that participate in communications used by F-AODV which in turn causes a low probability of contention. Thus, F-AODV can accommodate more packet delivery in this case by reducing the number of collisions using a low number of FNs. Moreover, this is a direct consequence of adapting the MAC layer parameters (thanks to the AMP scheme described above) incorporated in F-AODV. Indeed, giving more access ability to FNs by allowing them more transmission opportunity duration (high $TXOP$ length) and assigning them minimum $CWmin$ and $CWmax$ to increase the access probability to the channel. Furthermore, due to load balancing effect triggered by the features of the algorithms that use E-AODV and F-AODV, their associated performance remain significantly high compared to the basic AODV protocol. This indicates the robust nature of the protocols and their ability to adapt themselves to increasing load.

The AODV protocol uses minimum hop count as metric. These results are an inherent bias toward the same routes involving to congestion.

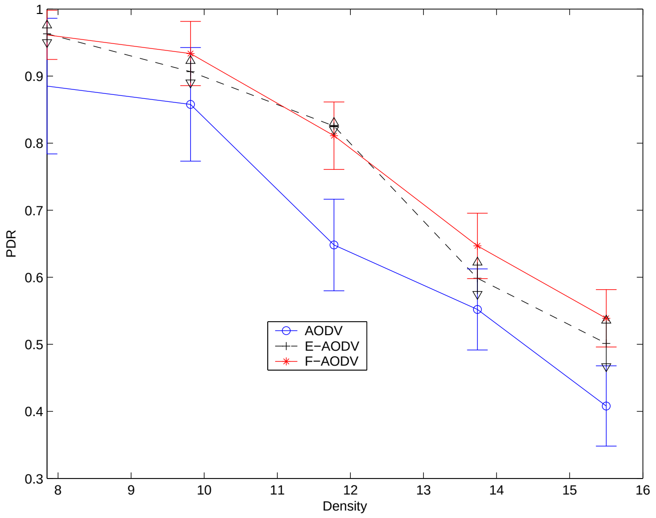

Figure 5.6: The effect of increasing the node density on the Packet Delivery Ratio (PDR)

A similar observation can be done in Figure 5.7, where we depict the Route Error Rate (RER) results obtained by F-AODV, E-AODV, and the basic AODV protocol. We observe that F-AODV has the minimum RER compared to AODV and E-AODV.

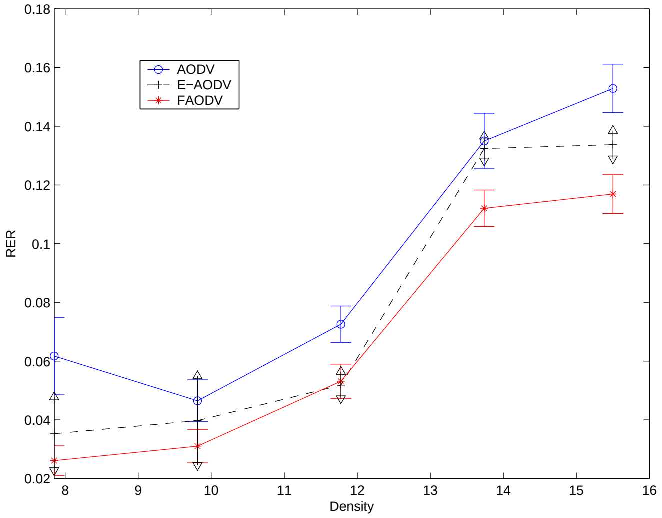

Figure 5.7: The effect of increasing the node density on the Route Error Ratio (RER)

In Figure 5.8 illustrates the routing overhead incurred by different routing protocols under different node density and load conditions. Routing overhead is an important metric to compare these protocols, since it measures the scalability of a protocol, the degree to which it will function in a congested or a low-bandwidth environment, and it has a direct impact

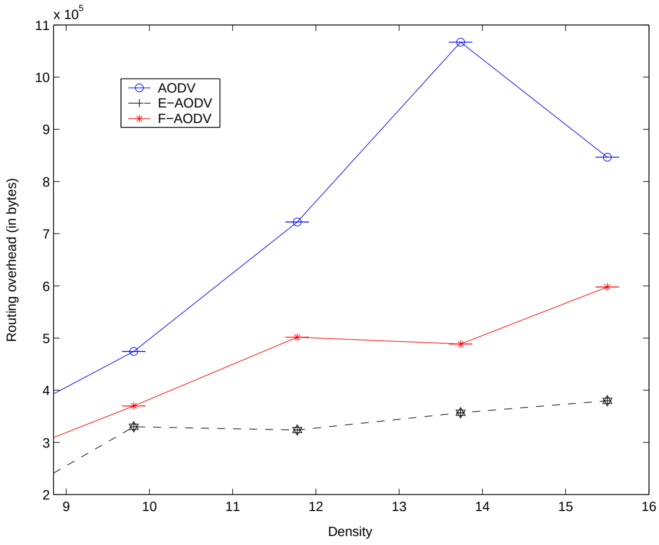

Figure 5.8: Routing Overhead

on network utilization efficiency. Protocols that generate large numbers of routing packets and/or bytes can also increase the probability of packet collisions and may delay data packets in network interface transmission queues. In Figure 5.8, we observe that both F-AODV and E-AODV have a lower overhead in terms of bytes compared to AODV protocol. Once again, this is due to high reactiveness of F-AODV and E-AODV to link changes compared to AODV, induced by congestion and energy exhaustion. Although F-AODV provides better PDR than E-AODV, E-AODV has minimum routing overhead in terms of bytes. In F-AODV, a large amount of packets are used for the role rotation of the forwarding process, which allows a distributed selection of the FNs and increase overhead (see Section 5.2). Moreover, F-AODV carries $E[E]$ and $E[B]$ parameters in control packets and hence packet size is higher.

The average end-to-end delay includes all possible delays caused by buffering during route discovery latency, queuing at the interface queue, retransmission delays at the MAC, propagation, and transfer times. Generally, there are three factors affecting end-to-end delay of a packet: first the route discovery time, which causes packets waiting in the queue before a route is found, second congestion state of the network, which causes packets waiting in the queue before they can be sent, and finally the path length. The more number of hops a packet has to go through, the longest time it takes to reach its destination.

Figure 5.9 depicts the variation of the average delay as a function of node density. The delay increases with load for all protocols. With a low node density, the lower delay is incurred by AODV protocol. However, when the node density increases, E-AODV performs slightly better than F-AODV. It is important to note that E-AODV and F-AODV still show significantly lower delay compared to AODV in high congested network.

5.4.2 Impact of traffic Load

In this set of simulations, we investigate the influence of data traffic rate on the performance of the studied protocols. We fix the number of nodes to 40 and we increase the inter-packet arrival time from.

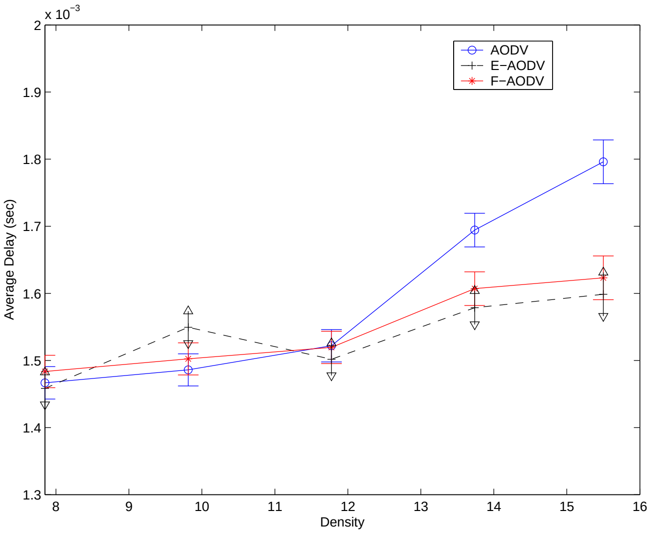

Figure 5.9: The effect of increasing the node density on the average delay

Figure 5.10 illustrates the PDR results. With low inter-packet arrival time, which corresponds to high data rate, E-AODV and F-AODV perform better than AODV. Indeed, the improvement is about 40% for E-AODV and 30% for F-AODV, compared to AODV. When we decrease the data rate, the three protocols have approximatively the same performance. However, F-AODV provides the minimum delay at high data rate compared to AODV and E-AODV. Its performance becomes similar to AODV when we increase the inter-packet arrival time. Contrarily, E-AODV has a high delay.

Figure 5.12 depicts the RER results. At high load, E-AODV has the lower RER compared to F-AODV and AODV. However, the results on RER of the three protocols are similar when we increase the inter-packet arrival time. The RO results shown in Figure 5.13, remain quite similar to those presented for the effect of node density.

5.4.3 Impact of node speed

In this set of simulations, we investigate the influence of node mobility on the performance of the studied protocols. Thus, we varied the initial speed. Indeed, the increase of initial speed leads to an increase on the average speed. In return, the mobility of the network becomes high.

As nodes become highly mobile, the probability of link failure increases. Consequently, the route error rate also increases. However, due to the consideration of energy metric and node load in route establishment scheme, E-AODV and F-AODV have the minimum route error rate compared to AODV as shown in Figure 5.14. In Figure 5.16, we illustrate the results of routing overhead. E-AODV has the minimum routing overhead compared to F-AODV and AODV. Figure 5.15 shows that E-AODV ans F-AODV have higher packet delivery ratio (see Figure) as a consequence of load balancing effect triggered by both node mobility and the use of the adaptive cross-layer mechanisms. Indeed, route failure due to power exhaustion and node congestion are avoided using our proposals. We observe that F-AODV has the lower RER and the higher PDR compared to E-AODV and AODV. This is due to the fact that F-

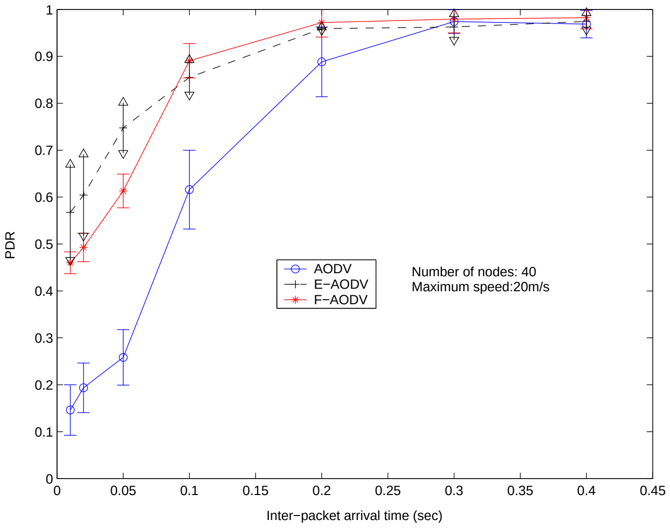

Figure 5.10: The effect of increasing the data rate on the PDR

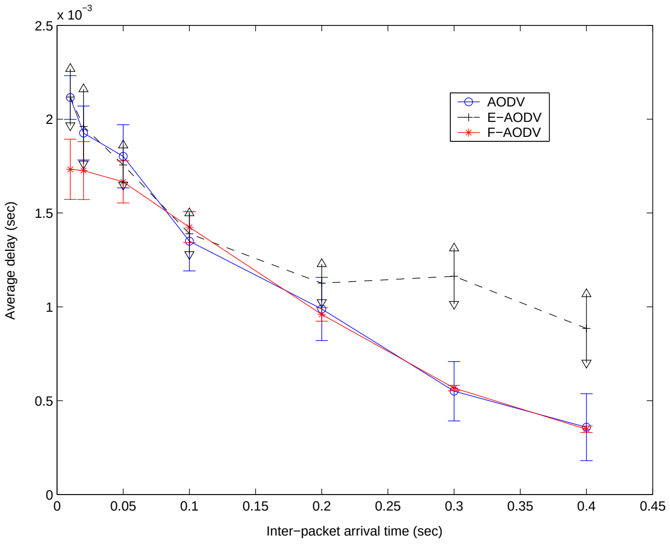

Figure 5.11: The effect of increasing the data rate on the average delay

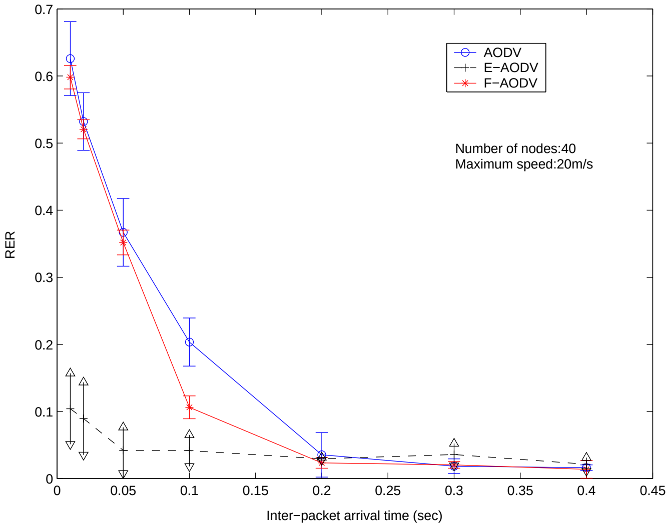

Figure 5.12: The effect of increasing the data rate on the RER

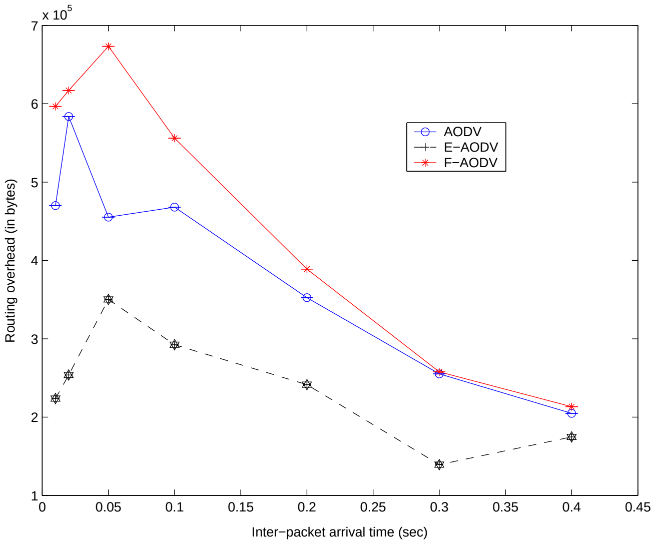

Figure 5.13: The effect of increasing the data rate on the routing overhead

AODV employs FNL, allowing nodes to use other route possibilities in case of routing failure. In return, this avoids re-starting the route discovery process (as described in Section 5.2).

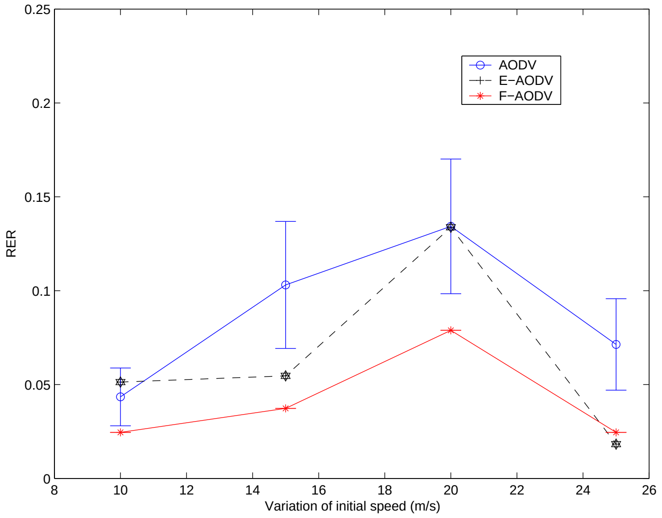

Figure 5.14: The effect of increasing the initial speed on the RER

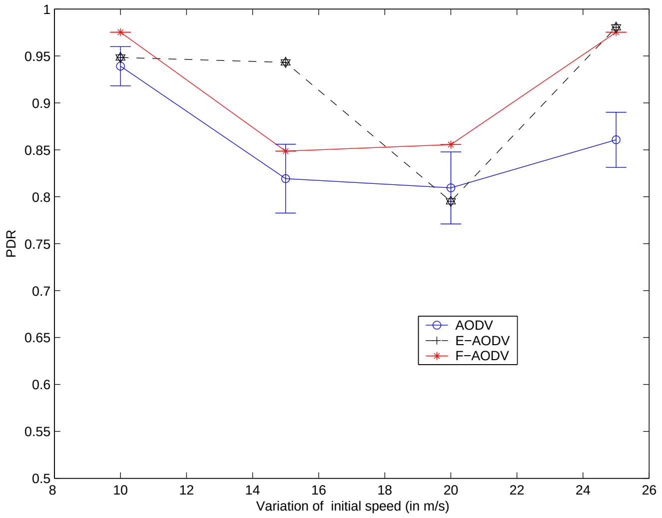

Figure 5.15: The effect of increasing the initial speed on the PDR

Another interesting observation is that for the most protocols the end-to-end average delay uniformly increases from low mobility rate to medium mobility rate (see Figure 5.17).

5.4.4 Discussion

The immediate remark that we can observe from the analysis presented above, is that both E-AODV and F-AODV which are based on MAC and network layers' cooperation provide a QoS

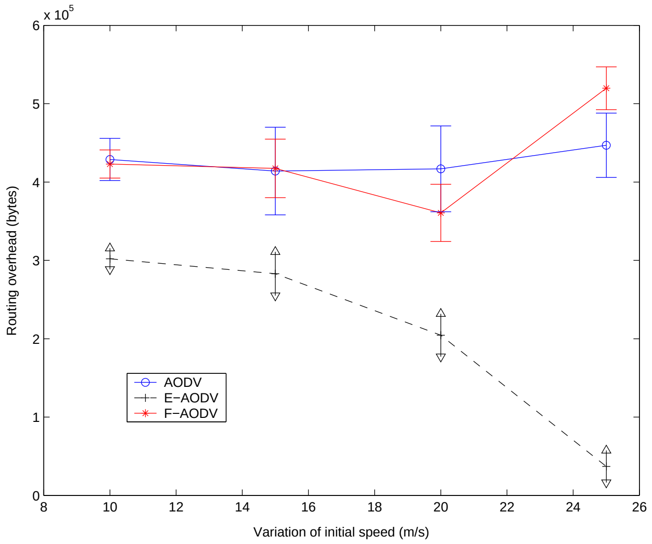

Figure 5.16: The effect of increasing the initial speed on the routing overhead

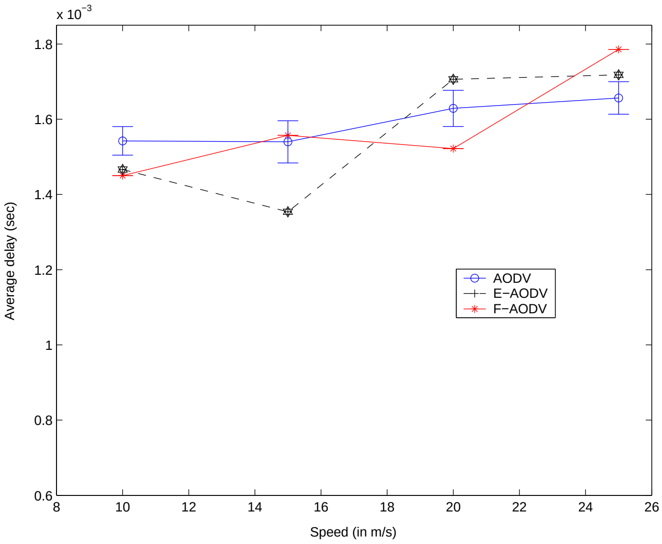

Figure 5.17: The effect of increasing the initial speed on the average delay

enhancement in terms of packet delivery ratio (PDR) especially at high node density. On the other hand, the MAC layer adaptation scheme used by F-AODV enables a higher improvement of the PDR compared to E-AODV. Furthermore, we see a significant enhancement in terms of the average end-to-end delay at high node density. However, a slight improvement of this metric is observed at high data rate. However, the E-AODV mechanism has the higher average delay compared to the basic AODV and F-AODV at low data rate.

Although the mobility causes frequent link failure, it allows diversity and load balancing. Moreover, our proposals enable nodes with better characteristics (nodes that are less congested and have high energy level) to participate in the data forwarding process. Consequently, the probability of route breaks is reduced and the routing overhead is minimized. It is also notable that the results in terms of average delay as function of node mobility, for the three protocols, are almost similar.

Overall, we can conclude that we have to take into account the application QoS requirements as well as the network characteristics in order to select the appropriate routing scheme that leads to better performance. One can also learn from the simulation results that in some cases (for example at low data rate), it is inefficient to count on the inter-layer parameters in route establishment scheme. Thus, we have to consider the accuracy level of the inter-layer parameters used in route establishment in order to achieve the overall performance enhancement objectives.

5.5 Chapter summary

This chapter presented a cross-layer approach called F-AODV, based on the cooperation between MAC and routing protocols. The objective of our proposal is to reduce the number of forwarding nodes (FN) in the network by considering their battery level and queue occupancy. On the other hand, it dynamically adapts MAC layer parameters of intermediate nodes, which allows them to achieve high probability of transmission success. The simulation results we obtained show that our model provides a total throughput significantly higher than the basic scheme. Besides, it provides a higher degree of fairness than the basic AODV protocol between competing user sessions. Furthermore, we compared the simulation results obtained with and without considering inter-layer cooperation. To this end, we investigate the performance of AODV, E-AODV and F-AODV with different scenarios and network mobility patterns. The results that we got, showed that the performance of the inter-layer cooperation paradigm depends on the network characteristics and the application constraints.

Globally, we showed that E-AODV and F-AODV protocols perform better than the basic AODV protocol because they ensure a good routing management scheme for MANETs and prevent (or at least delay) data congestion in the network by traffic load balancing.

Obviously, the energy is one of the important metrics to consider when designing communication protocols for MANETs. However, there are other parameters which have similar importance but they have been included only in limited number of new proposals in the open literature. One of them is the stability of wireless links which measures how long each wireless link between two pair of nodes will be alive. In the next chapter, we propose a new routing scheme which takes into account the stability metric when establishing routes between wireless devices. We also develop the required algorithms to estimate in an accurate way this metric based on the measurements done at the MAC and network layers.

Chapter 6

A cross-layer stability-based routing mechanism for MANETs

It is widely known that having neighborhood information can help on optimizing the operations of several protocols including routing and medium access protocols. This work presents a new routing scheme for mobile ad hoc protocols that effectively determine and use the information of link stability to select the most stable path between each pair of nodes in the network. This scheme could be applied for both reactive (on-demand) and proactive routing protocols.

Unlike existing similar protocols, our proposal has two main specific features. First, it is designed for both on-demand and proactive routing protocols and second, and most importantly, it uses the cross-layer paradigm to gather useful measurement from the MAC layer in order to estimate the stability rate of links in an accurate way. Hence, we develop adaptive stability metrics to identify stable links in a mobile wireless networking environment based not only on the analysis of periodic routing messages, but also on the MAC measurements. Our stability metrics then only rely on on-line statistical evaluation of observed link durations. Neither do they require information on signal strength, nodes speeds, nodes directions, radio conditions, or spacing of the mobile devices, nor do they depend on the availability of additional hardware such as GPS receivers or a synchronization of the devices.

Without loss of generality, we integrate our proposal in the AODV routing protocol and we demonstrate the ability of the selected metrics to determine stable links with a high probability in a wide range of scenarios using ns-2 simulations. The simulation results also demonstrate that our proposed protocol is able to overcome transient network characteristics due to mobility, and extend the longevity of established routes.

The remainder of this chapter, is organized as follows. Section 6.1 provides the motivation behind this work and explain the importance of considering the stability metric in MANETs. The description of the proposed cross-layer routing protocol is given in Section 6.2. Simulation methodology and performance evaluation of our proposal are detailed in Section 6.3. In Section 6.4, we present a brief and up-to-date overview of the most close related works that have previously addressed the stability issue in mobile ad hoc networks. Section 6.5 concludes the chapter by summarizing results and outlining future works.

6.1 Importance of considering links stability in MANETs

One of the key factors, which makes difficult the development of QoS routing protocols for wireless ad hoc networks, is the link breakage as a result of the mobility of its mobile devices. Hence, every self-organizing system capable of change has several characteristics that can take different values according to the mobility pattern as well as the energy level of mobile devices. For example, a node can have different positions, move with different speeds, and have different directions. All these variable factors can determine temporary the characteristics of the system. However, even though it is very useful that a node has these knowledges of its neighbors, it is very difficult and costly to collect all these information. Indeed, it requires a complex computations and so consumes many vital resources that will involve a serious impact on mobile ad hoc network performance. Thus, instead of selecting weak links which will break soon and introduce more maintenance overhead, we can select stable links, i.e. having longer expected lifetime, at the user session starting's time. By taking into consideration the link stability metric in routing protocols, the routing overhead can be significantly reduced. Moreover, the end-to-end QoS performance can be improved drastically.

Published research works that have addressed the stability-based routing problem usually require dedicated equipments to measure the stability degree of wireless links. However, our proposal described in this chapter does not require a GPS-like system and the exchange of positions information. We describe a stability-aware routing protocol that uses the cross-layer paradigm by making the MAC and routing layers interacting to efficiently estimate the stability of wireless links. The proposed scheme operates at both the MAC and the network layers.

At the MAC layer, each node estimates the average load of all its neighbors by observing and analysing all data/control packets sent in the neighborhood even if it is not the destination of these packets. Hence, for each node a *load vector* is computed where each element represents the traffic load observed from one of its neighbors. Obviously, the size of this vector is the number of neighbours which may vary dynamically.

At the network layer, observing the rate of received routing periodic *Hello* messages of neighbor nodes for each measurement period (called hereafter the *update period*), a *stability vector* is computed where each element stores the stability rate of one of its neighbours. In order to carry the stability information about the traversed wireless links between the source and the destination, we propose to add a *stability information* field to the Route REPly (RREP) and Route REQuest (RREQ) messages. The RREP message goes via the reverse path through the intermediate nodes till reaching the source which selects the most stable route using one of two possible approaches: path-stability-based and neighborhood stability-based.

Hereafter, we propose an inter-layer cooperation mechanism called Cross-layer Stability Routing Protocol (CSRP) that aims to use stable routes during sessions lifetime. We incorporate the proposal in the AODV routing protocol resulting on a new version of AODV called hereafter S-AODV (AODV with Stability metric consideration). The main features of our proposal is that it considers accurate parameters that are shared between different layers to maintain topology stability for routing data packets. The stability process is simple and efficient and can be easily applied to real world scenarios. Results provide an enhancement in terms of several QoS metrics such as the delay, the bandwidth, and the energy consumption.

6.2 Proposal description

A neighbor is considered stable with regard to a given node when both of them are stable or they are moving to the same direction with the same speed. However, this cannot be known without monitoring frames sent by each node's neighbor in absence of GPS-based localization system. This section details the operations of our proposal as well as how the needed information to quantitatively evaluate the stability of wireless links is gathered.

6.2.1 Preliminaries and assumptions

The format of the route request (RREQ) and route reply (RREP) packets in the extended AODV described in this chapter is slightly different from those used in the basic AODV protocol. Indeed, in addition to the hop count, we introduced a new field to store the *stability information* of all traversed links from a node to another one either for RREQ or RREP message. In the following sub-sections, we will detail how the content of this field is computed for each kind of message.

In addition of the new field added in the routing messages, a node has to add a new field, called *path stability rate* in each routing entry it maintains. This field informs about the stability of the path to the corresponding destination. More information on how to compute this field will be given later.

6.2.2 A short overview of our proposal

Before embarking into the proposal details, let's give a short overview of its main operations. We assume that a *stability vector* which measures the stability of neighbors is maintained and updated every specific period of time by each node in the mobile ad hoc network. On one hand, when a node i receives a route request message from its neighbors j, it updates its next hop to the source according to the stability rate link (i, j) in order to use the most stable next hop for the reverse path. On the other hand, when a node receives a route reply message, it updates the *stability information* field by taking into account the stability rate of the node from which it received this message. The **minimum stability** rate value of intermediate nodes is maintained for each route. At the MAC layer, the node computes the average load of all the neighbors by observing all packets sent in the wireless channel regardless the source and the destination nodes. Hence, an element in a vector, called the *load vector*, is maintained for each neighbor. The traffic load information will be applied as a metric to select two routes having the same *path stability rate* .

The source selects the best path when it receives more than one route reply by choosing the **maximum** of the minimum stability rate obtained values. Two methods are proposed for the path selection: path-based selection and neighbor stability selection. In the first method, the source decides which path to use by taking into account only the stability rate of available paths toward the destination. However, in the second approach the source considers the stability of neighbours of nodes belonging to the available paths. We believe that the former method is adequate for networks with heterogeneous movement patterns while the latter one is interesting for networks with homogeneous movement patters. These methods as well as the motivations behind considering them will be detailed later in the following subsections.

6.2.3 Computing the stability vector

Each node monitors the control packets received from its neighbors. These control packets include not only routing period packets but also each frame that the MAC layer receives.

A stability vector is maintained by each node which is indexed by the neighbors link layer 2 identity. Hereafter, we detail how this vector is computed and updated and explain its effects on route lifetime as well as the end-to-end QoS guarantees in mobile ad hoc networks.

Let's first define some useful parameters. Denote by $\triangle_{up}$ the update period after which the stability vector V_i at node i will be updated and let n_i be the number of neighbors of node i.

Assume $n_{ij}(\triangle_{up})$ be the number of AODV Hello messages received from node j at node i during $\triangle_{up}$. Note that

$$n_{ij}(\triangle_{up}) \leq n_{HELLO}^{max}, \tag{6.1}$$

where n_{HELLO}^{max} is the expected maximum number of Hello messages that can be sent during $\triangle_{up}$. Indeed, as Hello messages are broadcasted periodically, we can know the maximum number of these messages that could be sent by any node in the network in a given period. These messages can be lost as they are broadcasted and not retransmitted if they are lost due to frame collisions or data corruptions.

Given the parameters defined above, the ***stability rate*** of node j at node i (or the stability rate of the link $i \rightarrow j$) is then

$$V_i(j) = s_{ij} = \frac{n_{ij}(\triangle_{up})}{n_{HELLO}^{max}} \tag{6.2}$$

Obviously we have $0 \leq s_{ij} \leq 1$. Note that, because of the risk of collisions, node i might not receive an expected Hello message sent by node j at an expected time especially because it is sent in broadcast. In order to overcome this problem and to increase the accuracy of the estimated stability value, we also count any received frame from neighbors if an expected Hello message does not arrive at the expected time. So, if from an expected time to receive that message to the current time the node i receives another control packet or a data packet from j, it increments $n_{ij}(\triangle_{up})$.

6.2.4 Selecting the stable reverse path by intermediate nodes

First of all recall that in the basic AODV, when a node receives a RREQ send by a source, it has to create a reverse path to that source in order to be able to send back eventually received RREP messages. This path is known as the *reverse path*.

As we have mentioned above, associated with each routing entry to node s at node i, a stability rate that we denote $SR(s,i)$. Before forwarding a RREQ message, a node has to include in the ***stability information*** field of RREQ, its stability rate to the source. When a node i receives a route request message RREQ(s,d) from a neighbor j, it computes a new stability rate for the source as follows:

$$SR_{new}(s,i) = \min\left(SR(s,j),\, s_{ij}\right) \tag{6.3}$$

If the node has already a routing entry to the source node s through its neighbors k ($k \neq j$), it has to consider the one that has the **maximum stability rate** and to update the routing table according to the obtained result. If the next hop to the source s has been changed, node i has to broadcast the RREQ packet to its neighbors. These operations are done at each intermediate node until the route request reaches the destination d.

6.2.5 Selecting the most stable path by the source

When the routing protocol at the source node receives the first packet of an user application to send to a destination node, it checks whether it has a valid route to the destination or not. If not, it broadcasts a route request in order to compute a valid path. Each intermediate node sets up a reverse path to the source node and if it has a valid route to the destination, it sends back a route reply message otherwise it broadcasts the request to its neighbors as described above. When the destination receives a route request, it sends back to the source a route reply message.

Before sending/forwarding a route reply message, each node has to update the stability information contained in the message header by taking into account the stability of the link from which it receives this message. The node generating this message sets this field to 0 if it is the destination, otherwise it sets this field to the stability information it has in its routing entry to reach the destination.

After a period of time following the sending of its route request (RREQ), a source receives a first route reply (RREP) message which includes the stability information about the path toward the destination. Because the source node cannot know in advance how many route reply messages it will receive, it starts using the first received path. If it receives a new route reply (RREP) message form another neighbor that has a stability information better than that of the current selected path, it switches to the new path otherwise, it continues using the previous selected path.

The way how the compute the stability information depends on the method that the source wants to adopt to select the best path. Hereafter, we explore two methods to select the best route by the source: path stability-based selection and neighborhood stability-based selection and we describe how the scalability information field is computed for each method.

6.2.6 Path stability-based selection

The stability information field in this case is updated as follows. Assume that a node i receives a RREP(m,s) (generated by node m) message to forward from a node j where the stability information field is set to s_{RREP}. Node i has to set the stability information field to $min(s_{i,j},\, s_{RREP})$, so that the source s retrieves in each received route reply the stability rate of the corresponding path.

This method ensures that the chosen path remains stable as long as possible given that the links it contains are the most stable among the possible links to reach the destination. In other words, it has the longest residual lifetime.

The motivation behind this method is that there are some network configuration where the shortest path is not always the most stable one. Using this method avoids having unstable paths especially in mobile networks where some nodes are moving very fast while some others are more less fixed or are moving together (group moving).

6.2.7 Neighborhood stability-based selection

We believe that the path stability-based selection is not adequate when the nodes have similar movement patterns which means that the values in the stability vector is quite similar in all nodes. That's why we propose a second approach based of the neighborhood stability which reflects how stable the neighbors of each node. To measure the neighborhood stability, we use an entropy-based technique.

Entropy [56, 5] presents the uncertainty and a measure of the disorder in a system. There are some common characteristics among self-organization, entropy, and the location uncertainty in mobile ad hoc wireless networks. These common characteristics have motivated our work in developing an analytical modeling framework using entropy concepts and utilizing mobility information as the corresponding variable features. In our context, we define the entropy at node i as

$$H_i(\triangle_{up}) = \frac{-\sum p_{ij}\log p_{ij}}{\log n_i}, \tag{6.4}$$

where $p_{ij} = \frac{sij}{\sum s_{ik}}$, where k belongs to the n_i the neighbors set of node i. Hence, the stability information in the route requests will contain the minimum of entropy function values of traversed links. So, when node i receives a route reply from its neighbor node j, it compares its neighborhood stability $H_i(\triangle_{up})$ with the stability information s_{RREP}. It then sets this value to $\min(H_i(\triangle_{up}), s_{RREP})$ and forward the route request to the next hop toward the source.

In the following section, we report the simulation results of this proposal. We apply the above method since it can collect more information about next hop and its neighbor density comparing to the simple stability rate introduced in Subsection 6.2.6.

6.2.8 Analytic results: entropy vs. number of neighbors

Hereafter, we study the effect of the variation of the number of neighbors on the Entropy function using analytic results. Figure 6.1 shows that the entropy function increases when the number of neighbors increases. This result proves the utility of using this feature to capture the stability level of a given node based on the number of neighbors around it.

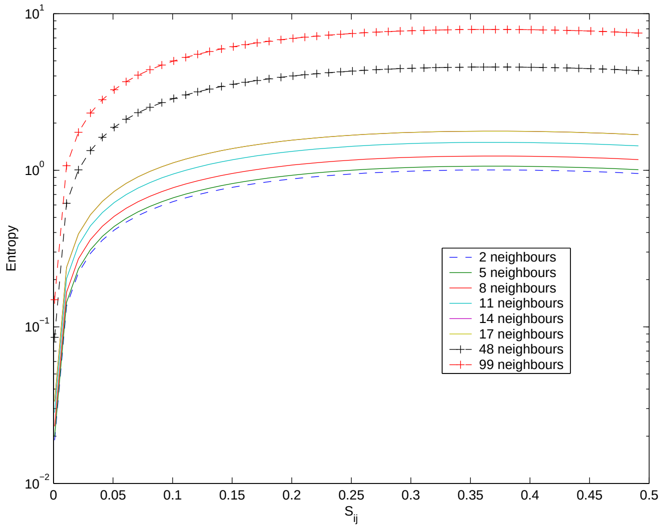

Figure 6.1: Entropy vs number of neighbors

6.3 Performance evaluation

We have implemented our proposal in ns-2 network simulator [49]. We have extended the AODV [43] protocol and DCF [32] scheme to support our cross-layer algorithm. We report in this section the results of simulations we have done for various network scenarios. We also provide a performance analysis of our proposal based on the obtained simulation results.

6.3.1 Scenario description

The simulated scenarios consists of 50 nodes located in a uniform distribution within an area of 1500x300 forming a multi-hop network. These scenarios are generated by the enhanced random waypoint mobility model described in [76].

In this mobility model each node moves toward a random destination and pauses for a certain time after reaching the destination before moving again. In our simulations, the nodes move at an average speed of 15 m/sec. The pause times are varied to simulate different degrees of mobility. The traffic sources start at random times after the beginning of the simulation and stay active during the remaining simulation time. The sources are CBR (Constant Bit Rate) and generate UDP packets at 4packets/sec, each packet being 512 bytes. Each simulation is run for 900 seconds simulated time. Each point in the plotted results represents an **average of ten simulation** runs with **different random mobility scenarios**. Note that the number of source nodes is 30 and the radio model is very similar to the first generation WaveLAN radios with nominal radio range of 250m. The nominal bit rate is 2Mbps.

6.3.2 Simulation metrics

We analyze several QoS metrics to evaluate the performance of our approach and we compare the obtained results with the AODV basic mechanism protocol. The following metrics are defined:

- **Packet delivery fraction**: the delivery fraction is measured as the ratio of the number of data packets delivered to the destination and the number of data packets sent by the source.
- **Routing overhead**: it is the routing overhead which is measured as the total number of bytes of transmitted routing packets.
- **Average delay**: it is the average delay of all competing flows. The average delay is used to evaluate how well the schemes can accommodate real-time flows.
- **Gain on remaining energy**: this metric evaluates the gain (in %) on the total remain energy obtained by our new proposal (S-AODV) compared to the basic mechanism. It shows the gain on energy and the effect on the connectivity of the network and thus the useful lifetime.

We present in this subsection the performance of the basic **AODV protocol** and our **S-AODV** proposal for the various metrics presented above.

In Figures 6.3, we plot the mean delay of our new mechanism and the basic AODV routing protocol. It's obvious from the curves that the mean delay is improved well using S-AODV. Indeed, in such scenarios, our algorithm allows re-routing and refresh routes including new

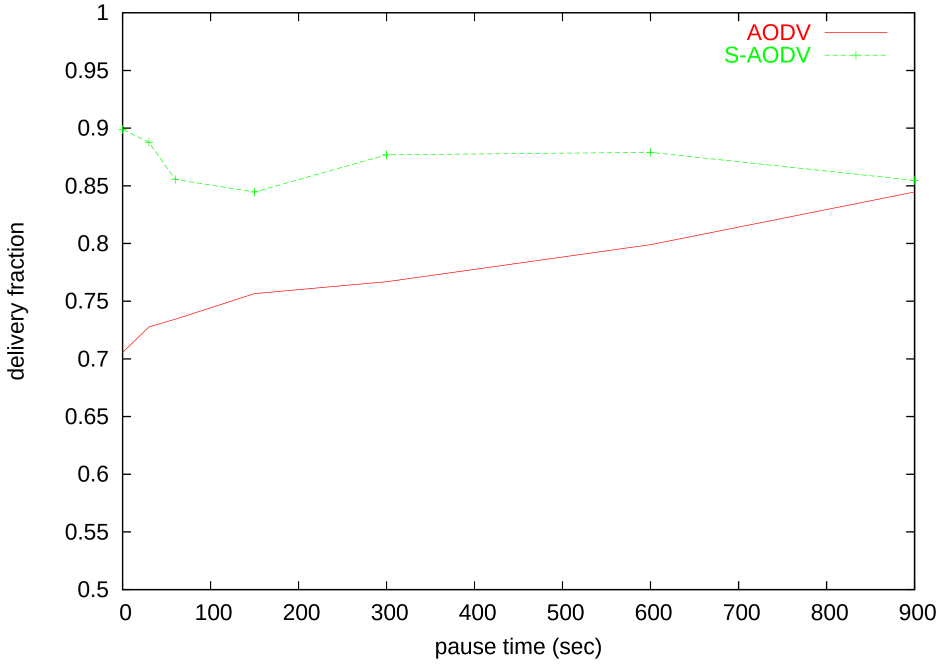

Figure 6.2: Total packet delivery fraction

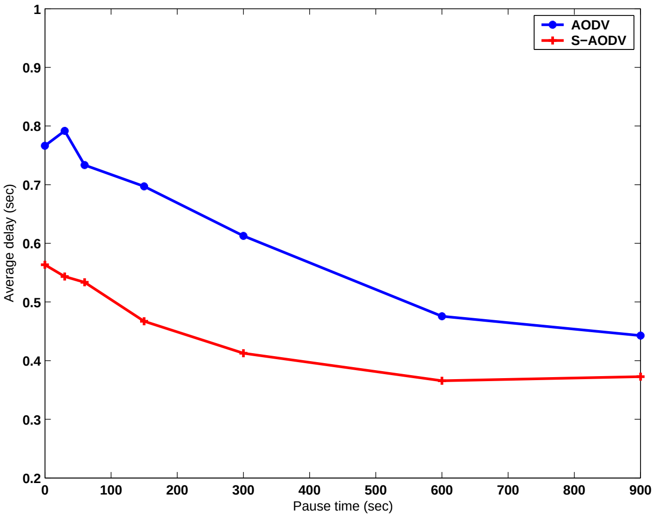

Figure 6.3: Average delay

nodes that have better quality than in the old routes which improves the end to end delay. Note that we mean by a good quality node, the node that is more stable and less busy regarding the other nodes in the some other alternative route possibilities. More specifically, high stability rate informs that the node does not change frequently its neighbors. Moreover, differentiate between nodes that participate heavily in communications including sending, receiving, and forwarding packets, and other nodes helps to avoid **at the same time selecting unstable and high congested nodes.** Indeed, this lets packets follow routes that generate a high cost and so are less congested which yielding to lower delay comparing to the obtained delay with routing based on minimum hops count (Figure6.3). This later, does not take into account node stability regarding to its neighbors during the simulation time and so it ignores links with lowest stability rate and furthermore busy nodes. Furthermore, we remark that the improvement on delay increases with high network mobility. The basic AODV change routes frequently which increases routing overhead consumed to re-establish broken routes (see Figure6.4). However, our proposal can select stable route even with mobile nodes but they follow the same movement direction. By this way, route failures are more avoided than in the basic AODV protocol and so the routing packet broadcasts decrease. Indeed, in such scenarios, our algorithm allows re-routing and refresh routes including new nodes that have better quality than in the old routes which improves application performance. Hence, the improvement on packet delivery ratio attempts more than 16% for high mobile network as shown in Figure 6.2. However, no significant improvement for stable network (pause time = 900). These results are proved by the fact that routing overhead is low with our new mechanism as shown in Figure 6.4.

To demonstrate the efficiency of our scheme regarding to the energy consumption, we plot in Figure 6.5 the gain in the total remaining energy that is obtained at the end of simulation. Our algorithm presents an improvement for more than 18% with comparing to the basic AODV. We observe that the gain on energy increases with mobility. This show that our protocol gives more benefits when mobility increase. Moreover, this metric demonstrates how network longevity can be extended using our proposal.

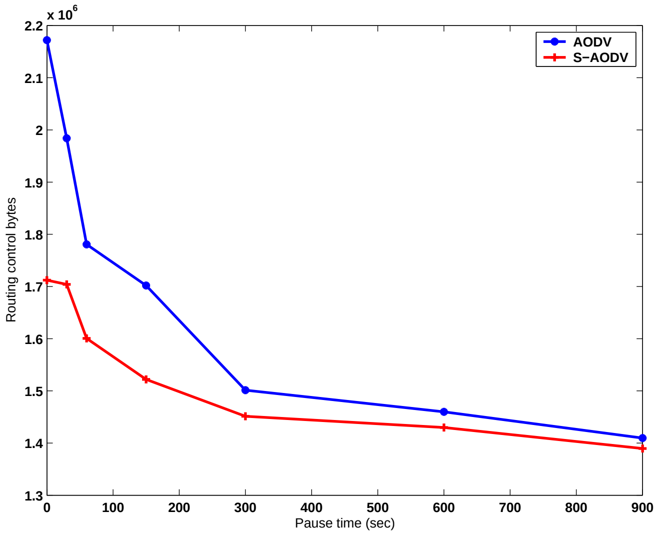

Figure 6.4: Routing overhead

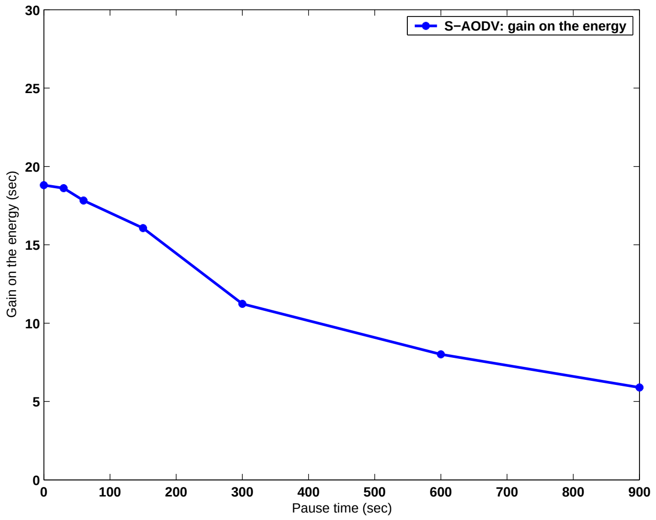

Figure 6.5: Gain on the remaining energy at the end of each simulation

6.4 Related works

Many works have investigated the stability issue in mobile ad-networks. Hereafter, we review the most related works close to our mechanism and that address the use of the stability information in the routing protocols.

The stability of a link is given by its probability to persist for a certain time span, which is not necessarily linked with its probability to reach a very high age. One of the earliest works in the context of link stability is the development of Associativity Based Routing (ABR) [67]. The idea behind this ad hoc routing protocol is to prefer stable links over transient links. A link is considered to be stable if it exists for a time of at least $A_{thresh} = 2rtx/v$, where rtx is the transmission range and v denotes the relative speed of two devices. It is left open how to determine the relative speed v among the mobiles which in turn determines A_{thresh}. ABR measures the lifetime of a link periodically. The motivation behind this approach has been found in assuming an implicit grouping for links that reach a certain age. After a time of A_{thresh}, it is assumed that the nodes move with a similar speed in a similar direction and thus are likely to stay together for a relatively long period of time. However, this assumption is justified only intuitively for dynamic scenarios.

In a dynamic environment, we may combine another metric, path longevity, with the metrics above to avoid frequent route switching and to reduce routing overhead. After data transmission starts, some decision will be made when the metric drops to a predefined threshold. There already exists some research on selecting stable routes. [25, 55] provide some metrics to find stable paths. [26] defines a parameter called the stability of the route r, which is: $Stability(r) = (Associativity(r)/RelayingLoad(r))$, The term Associativity is the same as defined in the Associativity Based Routing (ABR) protocol [67]; the RelayingLoad is the number of routing entries in the routing table of that node. On the other hand, due to the node mobility, some links may finally break during the transmission even considering the route stability at the route discovery process. Link breakage will cause packet delay and more overhead to find a new route. One possible way is to predict the link status and switch to a new route before the link breaks. Some work is required to decide which parameter should be chosen to better predict the link failure and with low overhead before we can apply the prediction algorithm. Ideally, after predicting the link failure, a node has alternative routes to the destination to avoid packet drop and delay. For some routing protocols, such as DSR, nodes store alternative routes in their route caches. These cached routes should be maintained to follow the topology changes.

Signal Stability-Based Adaptive Routing protocol (SSR) presented in [20] is an on-demand routing protocol that selects routes based on the signal strength between nodes and a node's location stability. This route selection criterion has the effect of choosing routes that have "stronger" connectivity. It distinguishes strongly connected from weakly connected. However, this concept is considered only as a supplement to SSA signal strength based approach and has been found to perform poorly in [20]. links where a link is considered to be strongly connected, if it has been active for a certain predefined amount of time. SSR comprises of two cooperative protocols: the Dynamic Routing Protocol (DRP) and the Static Routing Protocol (SRP). Also based on signal strength measurements is the Route lifetime Assessment Based Routing (RABR) [4]. It tries to predict the time when the received signal strength falls below a critical threshold using a measured value of average change in received signal strength. Another prediction method for link durations is presented in [30]. The method is based on distance measurements between mobile devices. A refinement in [7] takes possible changes in speed or direction of motions into account. The distance between connection peers may be acquired with the help of GPS receivers or signal strength measurements. Apart from the shortcomings of these two methods, the problem with this approach is that the distance of a receiver is only a very vague hint on link availability. In realistic environments the coverage area of a radio transmission hardly ever has a circular shape and is subject to strong fluctuations. A further approach based on the availability of GPS measurements has been suggested in [62]. The Flow Oriented Routing Protocol (FORP) follows a similar approach of calculating a link s residual lifetime from a mobile s own speed and the speed and distance of the connected party. However, this method strongly depends on the assumption of a free space propagation model and on having GPS equipment available for distance measurements and time synchronization. These requirements can hardly be presumed in a realistic environment.

One possible way is to predict the link status and switch to a new route before the link breaks. Some work is required to decide which parameter should be chosen to better predict the link failure and with low overhead before we can apply the prediction algorithm. Ideally, after predicting the link failure, a node has alternative routes to the destination to avoid packet drop and delay. For some routing protocols, such as DSR, nodes store alternative routes in their route caches. These cached routes should be maintained to follow the topology changes. [77] introduces such a mechanism based on a route caching validation probability Pv and local search radius k. The source node will attach a threshold pt with the RREQ message so that the intermediate nodes can compare their Pv values to decide if the cached routes are freshed enough to use.

The co-operation between layers to enable performance enhancement is very important and useful in wireless ad-hoc networks. Numerous works have been presented in the open literature that introduce several coupling ways and solutions between different communication layers [52]. The global objective of such co-operation is to achieve a reliable communication-on-the-move in highly dynamic environments as well as QoS provisioning.

6.5 Chapter summary

In this chapter we proposed a cross-layer stability-based routing in mobile ad hoc networks. We described a new routing algorithm based on accurate stability parameters in dynamic network characteristics. In order to measure the stability of wireless links, each received frame at the MAC layer is processed even if the node is not the final destination. Information about the neighbor sending this frame is recorded and used by the routing protocol to increase the accuracy of the estimation of the stability vector which contains the stability rate of each link with the neighbors. This stability vector is then used by the routing protocol to select the best next-hop to a given destination.

Performance evaluation using ns-2 simulator show the importance of considering the stability information in route selection process. Overall, we conclude that our mechanism demonstrates significant benefits at high and unstable traffic scenarios. Even though we implemented the model in AODV, the technique used is very generic and can be used with any routing protocol such as DSR and OLSR. Furthermore, this proposal can be applied to single channel and multi-channel based medium access protocols, and there is no need for synchronization.

The stability metric is very important to select the best path between a source and a destination in MANETs. This can help also to reduce the packet losses due to link breakage. Another important metric which is important mainly for delay-sensitive applications is the end-to-end delay. In the next chapter, we tackle the problem of the determination of routes by taking into account this metric and we develop a cross-layer routing scheme which implements an intelligent algorithm operating between MAC and routing layer to select the path with the minimum end-to-end delay.

Chapter 7

A cross-Layer routing mechanism for delay-sensitive applications in MANETs

Quality of Service (QoS) support is critical to wireless home networking, video on demand, audio on demand, and real-time voice IP applications. Time-bounded services such as audio and video conference typically require specified bandwidth, delay and jitter guarantees, but can tolerate minor losses. One of the main problems of wireless links is that all the nodes compete for the resources and the channel access without considering the knowledge about the quality and the requirement of neighbours' communications. There is no consideration to guarantee packet delay and jitter to flows supporting time-bounded multimedia services. Indeed, the mechanisms on how to access the radio channel are extremely important in order to guarantee QoS and improve application performance.

In this chapter, we present a cross-layer routing mechanism, which is based on the cooperation between the On-Demand **A**ODV routing protocol and the EDCA MAC protocol. This proposal, called D-AODV (AODV with delay metric consideration) aims to find the best path according to the application requirements in terms of delay, bandwidth, route stability, etc. Without loss of generality, this chapter focuses only on determining the path with the lowest delay. Each node periodically estimates the average transmission delay for each class of service defined by 802.11e. This information is injected into routing request and reply messages crossing each node. The sender is then able to select the best path which fits with its delay requirement. In order to overcome transient network characteristics due to new communications set up and mobility, we develop a new buffer management scheme for the audio class of service. The primary aim of this scheme is to discriminate audio packets according to their tolerated end-to-end transfer delay and their current experienced delay. The simulation results demonstrate that our proposal improves the performance of delay-sensitive applications while maintaining a good packet delivery ratio of other traffics.

The remainder of this chapter is organized as follows. In Section 7.1, we introduce our motivations for the work presented in this chapter. The description of the proposed D-AODV cross-layer protocol is provided in Section 7.2. Simulation methodology and performance evaluation of our proposal are detailed in Section 7.3. Section 7.4 concludes the chapter.

7.1 Motivations

As we have demonstrated in Chapter 3, despite many enhancement mechanisms that have been introduced based on the EDCA to achieve QoS support, performance evaluation results in multi-hop networks proved that EDCA still suffers from significant throughput degradation and high delay at high load conditions. These limitations are caused by the increasing time used for channel access negotiation and transit network characteristics. In this context, the quality of the routing path plays a key role in the success of application delivery and QoS support. Consequently, the minimum hop count algorithm, applied to establish routes between sources and destinations, could not satisfy application requirements.

Our work considers another metric in the route establishment process. This metric, which is included in route cost computation, incorporates the MAC layer average delay transmission of all nodes participating in the route. We believe that by doing this, we are able to provide a good distributed load balancing among nodes by choosing nodes that are less congested and hence have lower delay. Indeed, the first received RREP in the basic AODV does not correspond usually to the one experiencing the lowest delay, as routing control packets have always the highest priority and are enqueued at the head of the queue of the MAC layer. However, mobile ad hoc networks have dynamic quality characteristics. So, the cost of the established route could change fast, which is not useful for delay-sensitive applications as we can provide a bounded end-to-end delay. To this end, our proposal aims to select the best route according to the estimated end-to-end delay and not to the number of hops. Moreover, our proposal suggests that the MAC layer of each intermediate node observes the experiencing delay of each ongoing packet and gets from the routing layer the required remaining time after which this packet has to reach the destination. Considering these parameters, we suggest a buffer management scheme which allows packets to be delivered within the maximum tolerated end-to-end delay.

Briefly, our mechanism is based on cross-layer interaction between MAC and routing protocols. Indeed, even EDCA is introduced to support link layer service differentiation, we show that it is insufficient to provide QoS guarantees for real-time applications especially in dynamic networks. Hence, sharing information between MAC and routing protocols, is essential to achieve a good performance and overcome a low packet delivery ratio in a network with highly changing characteristics.

7.2 Proposal description

As we mentioned above, our cross-layer proposal is based on the interaction between MAC and routing layers. Hereafter, we detail the proposed mechanisms that have been considered in our architecture.

7.2.1 Estimating node's transmission delay

Several research works [52] have showed the importance of considering the impact of MAC-layer on the application performance. In our scheme, when the MAC layer receives an unicast packet from the routing layer to be sent to the next-hop, it saves the current time. Then, when it receives the acknowledgment (ACK) for this packet, it computes the transmission delay, which is simply the ACK's reception time by the MAC layer minus the packet's reception time from the routing layer. This delay takes into account queuing, transmission, and propagation delays. These delays are computed for all transmitted packets during a configurable *period* called T. If there is no traffic, only propagation delay is considered. At the end of each period T, the node updates the average transmission delay for the corresponding priority. The obtained time also takes into account the eventual retransmission retries at the MAC layer. Furthermore, only successful transmitted packets are considered for estimating the average transmission time in intermediate nodes. We note by $D^{j}_{curr}(k)$ the current computed average packet delay at step j for node k.

To minimize the bias against transient delay, we use an estimator of Exponentially Weighted Moving Average (EWMA) to smooth the estimated values. Let $D^{j}_{avg}(k)$ be the average delay at step j (for each update period T) computed according to the following iterative relationship:

$$D^{j}_{avg}(k) = (1-\alpha) * D^{j}_{curr}(k) + \alpha * D^{j-1}_{avg}(k) \tag{7.1}$$

where j refers to the j^{th} update period T and $D^{j}_{curr}(k)$ is the instantaneous delay, α is a smoothing factor and it effectively determines the memory size used in the averaging process. In order to take into account the network dynamics, we choose α to be the current Medium Utilization (MU) measured around the node during the previous period. The MU metric is computed as follows:

$$\alpha = MU = \frac{T - IdleTime}{T} \tag{7.2}$$

where *IdleTime* is the portion of time when the medium is idle. This leads to α in the interval $[0, 1]$. We choose the using of the MU metric, since it quantifies the up-to-date state of the wireless medium and takes into account neighbour's transmission and deferring backoff times.

7.2.2 Including delay information in routing control packets

The routes are established based on the end-to-end estimated delay transmission cost as computed above. Each RREQ packet between a source s and a destination d includes the traversed path cost. Then each node maintains for each reverse route this cost in the routing table so that routes are built based on the route's cost, defined as follows:

$$Cost_{(s,d)} = \sum_{k \in path(s,d)} D_{avg}(k) \tag{7.3}$$

Before sending a route reply, the destination has to include the computed route's cost which is mentioned in the routing request that it has received. Note that in the basic AODV, intermediate nodes can send back to the sender a routing reply message when they have already stored information about routes to each the destination. In our proposal, we disable this feature as we want the sender to receive an up-to-date information about the estimated end-to-end delay. Hence, all route replies are sent by the destination.

7.2.3 Selecting the best path

As described above, the basic AODV routing protocol uses the minimum hop count criteria to establish routes between sources and destinations. However, considering the new cross-layer model, we follow a new route discovery scheme. Indeed, the source uses the first path retrieved while still accepting other routing replies during the user session lifetime. When a new routing reply arrives, the source observes the estimated delay included in this reply. If it is less than the stored delay of the current available route, it updates the metric of that route entry and uses it for the next packets. This policy allows to reduce the time that the source's waiting time to send the data packets as it cannot know in advance how many route replies it will receive.

7.2.4 MAC layer buffer management interaction with routing layer

The dynamic characteristics of mobile ad-hoc networks decrease the efficiency of the chosen route. In order to overcome transit network characteristics, using these metrics in a cross-layer model might be inefficient because they are based on inaccurate values which do not reflect the real situation around a given node. Moreover, as a node moves with an arbitrary speed toward an arbitrary destination, the computed metrics (according to the average MAC layer delay, or to the participation of the node in communication and the traffic load level around it) could change over time. Therefore, other nodes that consider the metrics of that node to build routes, could have an inaccurate information since they change according to mobility, traffic, and capacity. To overcome this problem, we develop a new buffer management scheme for the audio class of service that aims to discriminate audio data packets regarding their current experienced delay and their tolerated end-to-end transfer delay (*400ms in this context* [65]). This mechanism consists of reordering audio packet transmissions in the queue according to two parameters. The first parameter is the total delay that packet experienced during the route until arriving at the current node. To avoid synchronization problems, we incorporated in the packet header the sum of times that the packet has experienced in the previous intermediate nodes. The second parameter is the estimated delay that the packet is expected to encounter before it reaches the destination. This parameter is available in the routing table and it is shared with the MAC layer. If the sum of the two delays reaches the maximum tolerated end-to-end delay, the packet has to be inserted at the head of the queue of the audio class. Audio packets are then inversely ordered according to their remaining lifetime. Moreover, packets that exceed their maximum tolerate end-to-end delay are dropped. By this way we alleviate the network congestion since these packets will not be considered at the destinations.

7.3 Performance evaluation

We implemented our proposal in the ns-2 network simulator [49]. We have extended the AODV protocol and EDCA scheme to support our cross-layer algorithm. We report in this section the results of simulations we have done for various network scenarios. We also provide an analysis of the obtained performance results.

7.3.1 Scenario description

The simulated scenarios consist of 50 nodes located in a uniform distribution within an area of 1500x300m forming a multi-hop network. These scenarios are generated by the enhanced random way-point mobility model [76]. Our simulation uses different types of traffic to evaluate service differentiation. Three queues are used in each active station. The highest priority queue in each station generates packets with packet size equal to 160 bytes and inter-packet interval of 20 ms, which corresponds to 64 Kbps PCM audio flow (**high sensitive-delay applications**). The medium traffic queue generates packets of size equal to 1280 bytes each 10 ms which corresponds to an overall sending rate of 1024 Kbps. The low priority queue in each station generates packets with sending rate equal to 260 Kbps and using a 200 bytes packet size. To increase the load of the system, we gradually increase the number of flows until 84 flows: 30 audio flows, 25 video flows, and 29 best effort flows. Note that the number of source nodes is 30 sources. Moreover, we consider an arbitrary starting and end time of communications to show how the proposed model could be adapted to the dynamic network load. The radio model is very similar to the first generation WaveLAN radios with nominal radio range of 250m. The nominal bit rate is 2 Mbps. In our simulation the nodes move at an average speed of $20\,m/s$.

SIFS	$16\mu s$
DIFS	$34\mu s$
ACK size	14 bytes
Data rate	36 Mbps
Slot_time	9 μs
CCA Time	$3\mu s$
MAC Header	28 bytes
Modulation	16-QAM
Preamble Length	$20\mu s$
RxTxTurnaround Time	$1\mu s$
PLCP header Length	$4\mu s$

Table 7.1: IEEE 802.11a PHY/MAC parameters used in simulation

In the following simulations, we assume that each wireless station operates at IEEE 802.11a PHY mode-6, see network parameters shown in Table 7.1.

Table 7.2 shows the network parameters selected for the three TCs.

We have done an extensive set of simulations to observe the effect of the update period T on the delay performance. In the following simulations we have chosen T equal to 2 seconds, which provides us a good latency.

7.3.2 Performance metrics

To evaluate the performance of the different schemes, the following two metrics are used:

- **Latency distribution:** latency distribution allows to trace the percentage of packets

Parameters	**High**	**Medium**	**Low**
CW_{min}	7	31	31
CW_{max}	15	31	1023
AIFS(μs)	34	43	52
Packet Size($bytes$)	160	1280	200
Packet Interval(ms)	20	10	12.5
Sending rate($Kbit/s$)	64	1024	128
slot-time(ms)	16us	16us	16us

Table 7.2: MAC parameters for the three TCs

that have latency less than the maximum delay required by the applications. As real-time flows require both low average delay and bounded delay jitter, we propose to use the two metrics of latency distribution and delay variation.

- **Throughput**: this metric shows the total number of bytes that have been successfully delivered to the destination nodes.

7.3.3 Performance results and analysis

We present in this subsection the performance of the basic **EDCA-AODV** and **D-AODV** protocols for the various metrics presented above. We simulate 10 random scenarios and then take the average of the performance values.

Hereafter, we present results of different network mobility scenarios. In Figure 7.1 and Figure 7.2 we show the delay distribution of audio traffic and the variation of throughput over the simulation time respectively. The simulated scenario corresponds to a pause time of node equal to zero (in the random waypoint model), that leads to a highly mobile network. However, Figure 7.3 and Figure 7.4 are obtained in case of a static ad hoc network (no mobility - infinite pause time).

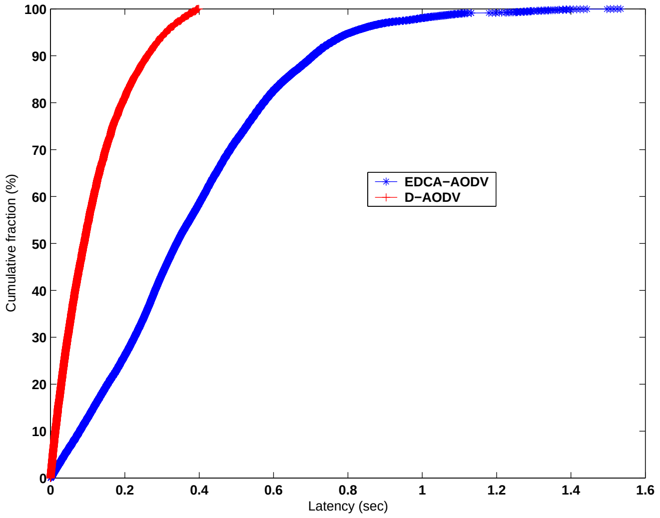

Figure 7.1: Delay distribution of D-AODV and EDCA-AODV for a highly mobile ad hoc network (pause time = 0)

We show the latency distribution for audio traffic in Figure 7.1 and Figure 7.3 in which a number of 30 voice flows is used to show the delay performance. On a cumulative distribution plot, an ideal result would coincide with the y-axis, representing 100% of results with zero latency. Although we cannot reasonably expect zero latency, we would like to obtain a consistent performance, corresponding to a vertical line.

Observing Figure 7.1 and Figure 7.3, we can conclude that mobility affects the delay performance of both D-AODV and EDCA-AODV due to the often route failures. However, D-AODV always outperforms the basic EDCA-AODV. There are considerable differences between them, i.e. more than 98% of audio packets have a delay value less than 400ms, whereas only 60% of audio packets for the layered protocols have a delay less than 400ms.

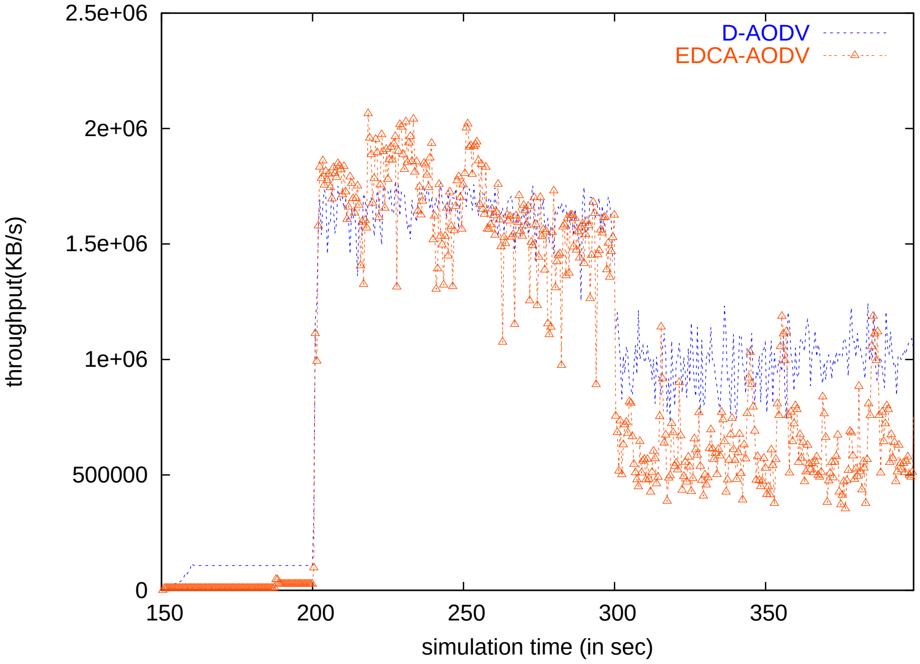

Figure 7.2: Throughput of D-AODV and EDCA-AODV for a highly mobile network (pause time =0)

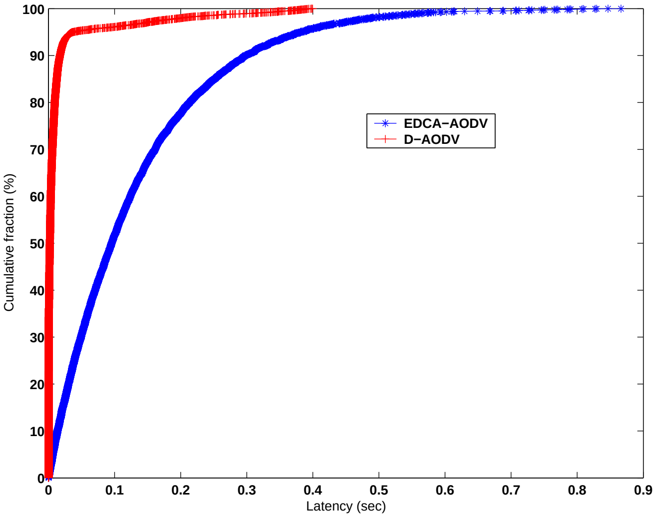

Figure 7.3: Delay distribution of D-AODV and EDCA-AODV for a static ad hoc network (no mobility)

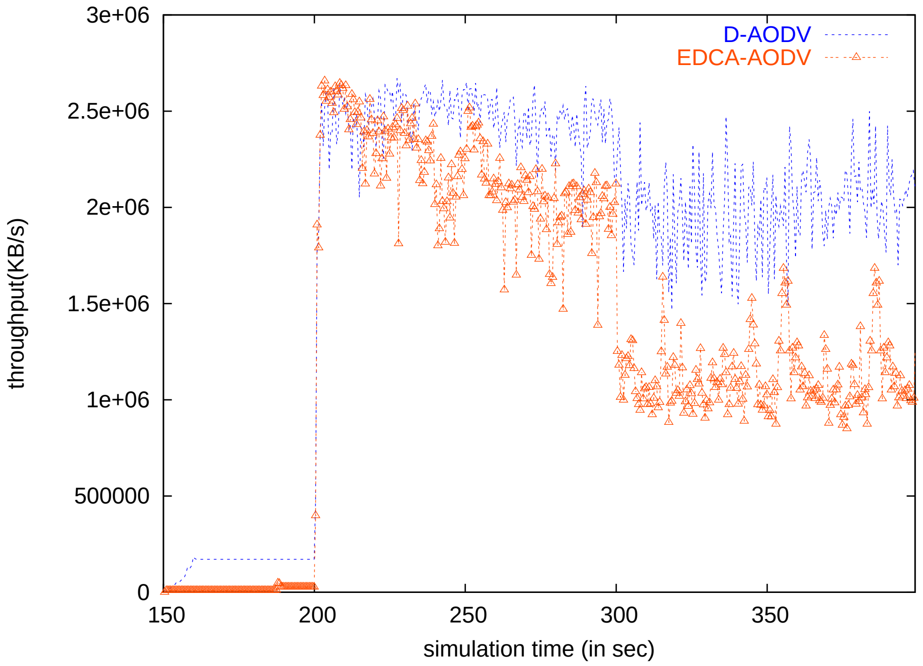

Figure 7.4: Throughput of D-AODV and EDCA-AODV for a static ad hoc network (no mobility)

These results are due to the good effect of including MAC layer delay in routing process and the importance of introducing the buffer management mechanism to satisfy application requirements. Note that the maximum delay of audio packets of D-AODV is less than $400ms$, whereas for the basic mechanism, the maximum value is much more than $800ms$. Indeed, getting a delay more than $400ms$ degrade the quality of audio reception. This is the effect of our proposal that drops packets which have experienced more than the maximum tolerated delay before reaching destination.

Choosing routes with minimum delay will decrease the number of congested nodes that build a QoS route. This allows load balancing in the network. Furthermore, the dropped packets in the intermediate nodes decrease the congestion of the network. Indeed, there are less competing packets to gain the medium in the next hops, that these packets expect to traverse, and so more successful transmissions. Figures 7.2 and 7.4 demonstrate the good effects of the considered schemes. Indeed, the total delivered throughput obtained by D-AODV is higher than EDCA-AODV. This ensures a good packet delivery for video and best effort traffics.

Obviously, when nodes are static, the improvement in term of throughput (Figure 7.4) is higher than that obtained in mobile network (Figure 7.2). The results are proved by the fact that routes are more often re-established due the the high mobility. So, less data packets are transmitted.

From the simulations, we can conclude that D-AODV can provide a good enhancement in term of delay for delay-sensitive application while achieving a good delivery packets for other traffic categories. Due to space limitations, we cannot include more simulations that we have conducted but we believe that performance results presented here are quite representative of what we have obtained.

7.4 Chapter summary

This chapter presented a new MAC and routing cross-layer approach called D-AODV. The cooperation goal aims to provide a good delay and delay jitter for delay-sensitive applications but it could also be applied for other QoS parameters such as bandwidth, loss rate, and a route stability. The approach is an adaptive service differentiation based on buffer management and route establishment strategy. Performance evaluation using ns-2 simulator shows the importance of considering the MAC-layer delays in route selection process. Overall, we conclude that our mechanism demonstrates significant benefits at high and unstable traffic scenarios. Even though we implemented the model in AODV, the technique used is very generic. Thus, it can be used with any kind of routing protocol such as OLSR by including of the required information in the TC (Traffic Control) control messages. Furthermore, this proposal can be applied to single channel and multi-channel based medium access protocols, and there is no need for synchronization. Additionally, the scheme could also be applied for the basic 802.11 protocol that does not support service differentiation at the link layer.

The cross-layer model is introduced mainly to enhance the performance of the real time applications and achieve better QoS support. However, the proposed cooperative algorithms and parameters have to be rigorously selected, compared, and optimized regarding network characteristics and application requirements.

In the next chapter, we will compare the performance obtained when we consider the delay, described in this chapter, in addition to the stability, introduced in the previous chapter, as the routing optimization metric for several scenarios. The aim of this comparison is to demonstrate the importance of considering the applications' requirements as well as the network characteristics in selecting the most suitable (cross-layer) routing protocol for each user session.

Chapter 8

CrossAid (XAid): towards a cross-layer architecture for MANETs

It is obvious that designing a cross-layer architecture for MANETs which could be used for more than one objective such as QoS provisioning, security, and multicast transfer. Our contributions described in this dissertation aim only to enhance the service differentiation in 802.11-based wireless networks. Hence, we proposed several schemes to enhance routing of packets between each pair of nodes. These mechanisms improve the route selection procedure by considering important metrics such as the stability of wireless links, the rate of consuming the energy, and the end-to-end average delay.

In this chapter, we first discuss the benefits and the disadvantages of the use of cross-layer paradigm when designing new communication protocols for wireless networks. We outline our recommendations derived from the analysis of the simulation results of the proposed mechanisms described in the previous chapters. Firstly, we identify the scenarios where sharing useful parameters between different layers is quite recommended to enhance the routing of packets. Secondly, we show the inefficiency of the cross-layer design paradigm in other network and traffic scenarios.

We summarize the chapter by outlining the challenges in solving the QoS provisioning with new cross-layer communication protocols, identifying various classes of solutions, and illustrating how the cross-layer paradigm-based optimization can be performed using several examples.

The remainder of this chapter is organized as follows. In Section 8.1, we provide a deeper analysis of the main obtained simulation results of the proposed cross-layer mechanisms. A ns2 simulation-based quantitative comparison of S-AODV and D-AODV proposals is given in Section 8.2. All our proposals are included in a new architecture called CrossAid (XAid) which will be described in Section 8.3. This section also gives a list of recommendations on using cross-layer based routing protocols. Section 8.4 summarizes this chapter and outline future work.

8.1 Analysis of the main obtained results

In this section we summarize the results that we have extensively described in the previous chapters. Thus, we recall how much the performance of the different proposals is enhanced. Moreover, we illustrate the scenarios where inter-layer interaction is useful for routing in

$+$	minor enhancement
$\simeq$	similar performance
$+ +$	good enhancement
$+ + +$	significant enhancement
-	minor performance degradation
- -	performance degradation

Table 8.1: Notations used for the comparaison of the proposed cross-layer routing mechanisms

MANETs. Furthermore, we identify the scenarios where each routing scheme among those we have proposed (E-AODV, F-AODV, S-AODV, D-AODV) is the best to use.

In this subsection, we summarize the performance evaluation of our proposals. Their comparison is given with regard to the basic AODV and EDCA combination approach. We define the notations that we use in our comparison in Table 8.1.

8.1.1 Summary of the obtained Packet Delivery Ratio results

We illustrate the comparison results in terms of the Packet Delivery Ratio (PDR) for low and high traffic load scenarios in Tables 8.2 and 8.3, respectively.

Cross-layer protocols	**Low mobility**	**Medium mobility**	**High mobility**
E-AODV	-	$\simeq$	$+$
F-AODV	$+ +$	$+ +$	$+$
S-AODV	$\simeq$	$+$	$+ +$
D-AODV	$+$	-	-

Table 8.2: Results of the PDR for low loaded networks and with different mobility levels

Cross-layer protocols	**Low mobility**	**Medium mobility**	**High mobility**
E-AODV	$+ +$	$+ +$	$+ +$
F-AODV	$+ + +$	$+ + +$	$+ +$
S-AODV	$\simeq$	$+ +$	$+ + +$
D-AODV	$+$	-	- -

Table 8.3: Results of the PDR for highly loaded networks and with different mobility levels

Table 8.2 summarizes the PDR results of our proposals with low traffic load. A good performance improvement is observed with F-AODV at both low and medium mobility. Moreover, there is an enhancement with S-AODV at high mobility. However, there is a degradation of the performance with E-AODV at low mobility. Furthermore, the same remark is observed with D-AODV at both medium and high mobility.

As we can see in Table 8.3 a significant performance enhancement in terms of PDR is achieved by F-AODV especially at low and medium mobility level. At high mobility scenarios, although D-AODV performs poorly, the S-AODV protocol provides a significant PDR improvement while it maintains a similar performance as the basic AODV when considering a stable network. Moreover, E-AODV provides a good performance enhancement at all different mobility levels.

8.1.2 Summary of the obtained average delay results

Cross-layer protocols	**Low mobility**	**Medium mobility**	**High mobility**
E-AODV	≃	-	-
F-AODV	-	-	+
S-AODV	-	-	+
D-AODV	+ +	+ +	+

Table 8.4: Results of the average delay for low loaded network and with different mobility levels

Cross-layer protocols	**Low mobility**	**Medium mobility**	**High mobility**
E-AODV	+ +	+ +	+ +
F-AODV	+	+ +	+ +
S-AODV	+	+	+ +
D-AODV	+ + +	+ +	+

Table 8.5: Results of the average delay for highly loaded networks and with different mobility levels

The average end-to-end delay is only enhanced by D-AODV mechanism as shown in Table 8.4. A minor enhancement is observed with S-AODV and F-AODV at high mobility.

At highly loaded conditions, all the proposed protocols improve the average end-to-end delay metric (see Table 8.5). A significant average end-to-end delay performance enhancement is observed with D-AODV at low mobility level. A minor enhancement is achieved with the S-AODV proposal.

8.1.3 Summary of the obtained routing overhead results

Cross-layer protocols	**Low mobility**	**Medium mobility**	**High mobility**
E-AODV	-	+	-
F-AODV	- -	-	+
S-AODV	≃	+	+ +
D-AODV	-	-	- -

Table 8.6: Results of routing overhead for low loaded networks and with different mobility levels

The routing overhead is increased with all the proposed mechanisms at low traffic load conditions except for S-AODV, which provides a minimum overhead especially at high mobility

Cross-layer protocols	Low mobility	Medium mobility	High mobility
E-AODV	+	+ +	+ + +
F-AODV	+	+	- -
S-AODV	$\simeq$	+ +	+ + +
D-AODV	-	- -	- -

Table 8.7: Results of the PDR for highly loaded networks and with different mobility levels

scenarios (Table 8.6). The reason behind this is that the estimation of the stability metric does not require the exchange of specific control messages. The only overhead in S-AODV is the field containing the stability information added in the AODV control messages.

The D-AODV scheme needs more routing control packets than other mechanisms at high traffic load conditions, as shown in Table 8.7. Whereas, the routing overhead is significantly reduced with both E-AODV and S-AODV when the nodes in the network are highly mobile.

8.1.4 Summary of the obtained energy consumption results

Cross-layer protocols	Low mobility	Medium mobility	High mobility
E-AODV	+ +	+ +	+ + +
F-AODV	+ +	+ +	+ + +
S-AODV	$\simeq$	+	+ +
D-AODV	-	-	- -

Table 8.8: Results of the energy consumption for low loaded networks and with different mobility levels

Cross-layer protocols	Low mobility	Medium mobility	High mobility
E-AODV	+ +	+ + +	+ + +
F-AODV	+ +	+ +	+ + +
S-AODV	$\simeq$	+ +	+ +
D-AODV	-	- -	- -

Table 8.9: Results of the energy consumption for highly loaded networks and with different mobility levels

Both E-AODV and F-AODV consume low energy than the basic protocol as shown in Table 8.8. However, the energy consumption is increased with D-AODV. This is explained by the increase of the routing overhead used by D-AODV.

The same remarks about the energy conservation results, observed at low traffic load with E-AODV and F-AODV, are observed at high network load (Table 8.9). However, D-AODV consumes higher energy than the basic protocol, whereas the improvement in terms of energy conservation increases with S-AODV at high mobility.

8.2 Performance comparison of S-AODV and D-AODV mechanisms

To quantify the importance of taking into account the network characteristics and the target metrics to optimize when selecting a routing protocol, we provide in this section a simulation-based analysis of the results obtained for S-AODV and D-AODV mechanisms. We evaluate the performances of these protocols under various network scenarios.

8.2.1 Simulation environment

The simulated scenarios consist of 50 nodes located in a uniform distribution within an area of 1500x300m forming a multi-hop network. These scenarios are generated by the enhanced random way-point mobility model [76]. The sources are CBR and generate UDP at 4 packets/second, each packet being 512 bytes. Note that the number of source nodes is 30 sources. The radio model is very similar to the first generation of WaveLAN radios with nominal radio range of 250m. The nominal bit rate is 2 Mbps. In our simulation wireless nodes move at an average speed of $15m/s$. We provide simulations for several pause time values.

SIFS	$16\mu s$
DIFS	$34\mu s$
ACK size	14 bytes
Data rate	36 Mbps
Slot_time	9 μs
CCA Time	$3\mu s$
MAC Header	28 bytes
Modulation	16-QAM
Preamble Length	$20\mu s$
RxTxTurnaround Time	$1\mu s$
PLCP header Length	$4\mu s$

Table 8.10: IEEE 802.11a PHY/MAC parameters used in ns-2 simulation

In the conducted simulations, we assume that each wireless station operates at IEEE 802.11a PHY mode-6, see network parameters shown in Table 8.10. We compare the performance of D-AODV and S-AODV protocols using the following metrics: packet delivery ratio, routing overhead, and average end-to-end delay.

8.2.2 Simulation results

In Figures 8.1, we plot the mean delay of the two mechanisms. It's obvious from the curves that the mean delay is improved well when using D-AODV for the case where there is no mobility (pause time=900). Indeed, the model enables packets routing over less congested nodes. However, this good performance decreases when the node mobility increases. Hence, the S-AODV mechanism performs better in such scenario with frequent changes. D-AODV allows re-routing and refresh routes including new nodes that have better quality than in the old routes which improves the end to end delay. Moreover, we remark that the improvement on delay increases with high network mobility. Furthermore, we can also observe this difference

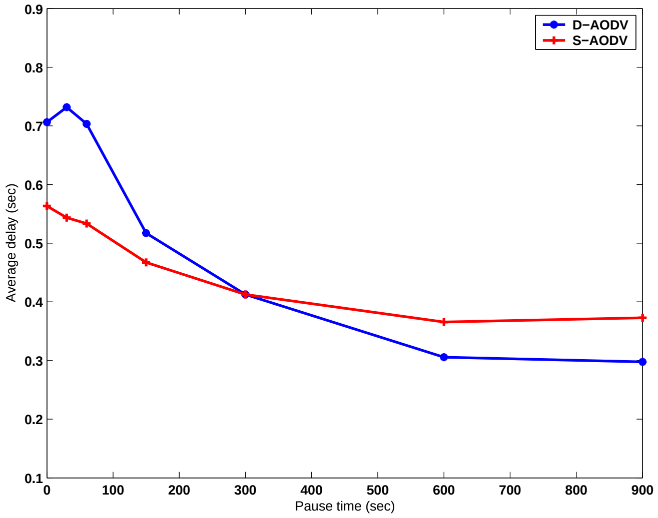

Figure 8.1: Results of the average end-to-end delay

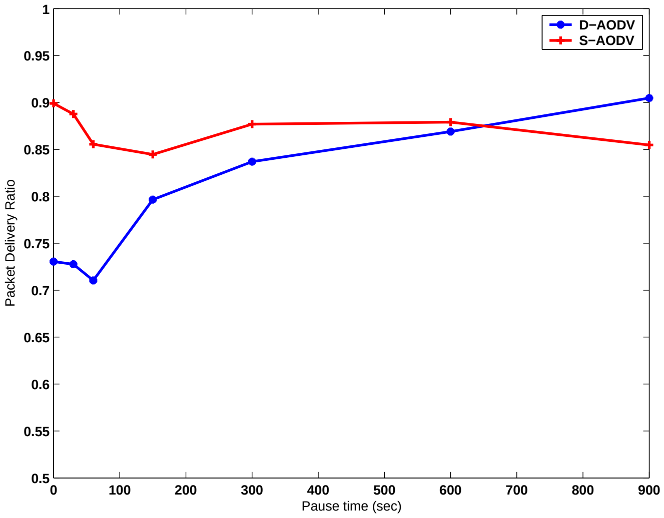

Figure 8.2: The results of the packet delivery ratio

on performance in Figure 8.2 and 8.3. Indeed, for the D-AODV scheme, the packet delivery ratio increases when the nodes are more stable. The routing overhead results obtained with this mechanism decreases when considering low mobility. These results demonstrate the importance of the adequate selection of cross-layer parameters regarding both network metrics and application requirements. On one hand, the efficiency of S-AODV is shown with high link changes. This mechanism is able to select stable routes even with mobile nodes but they follow the same movement direction. On the other hand, with D-AODV the performance improvement is obtained only when considering stable nodes. Overall, we can conclude from this analysis that the performance results depend strongly on the network characteristics.

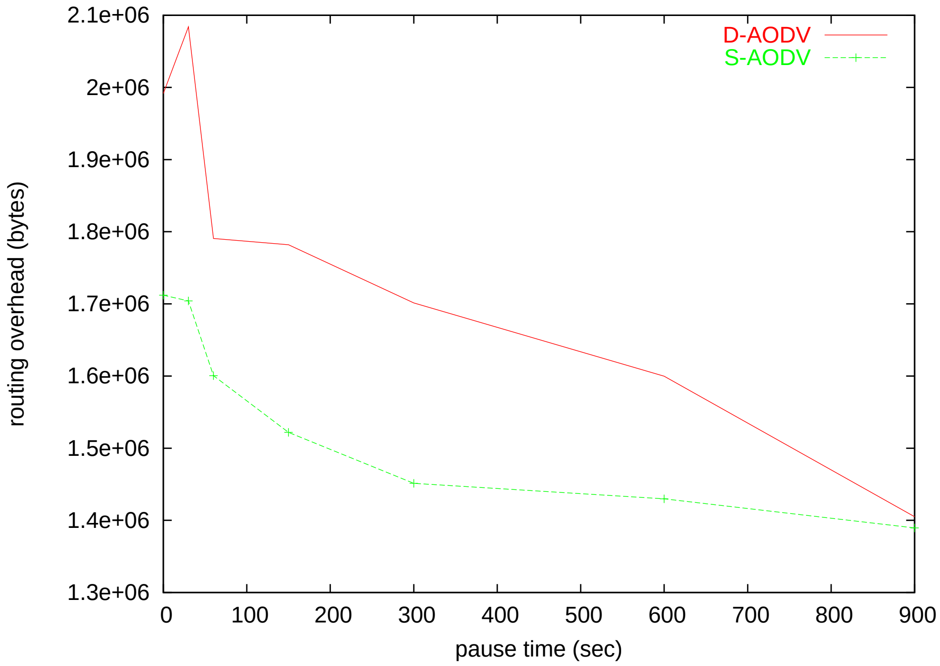

Figure 8.3: Results of the routing overhead

By this comparative study, we demonstrate that cross-layer routing mechanisms could not be efficient when network characteristics change frequently given that the estimation of the QoS metrics may not be accurate.

8.3 The CrossAid (XAid) architecture

Cross-layer models are mainly introduced to enhance the performance of real time applications and achieve better QoS support. However, the proposed cooperative algorithms and parameters have to be rigorously selected, compared, and optimized. In the most cases, we have to take into account the benefits of each model that provides inter-layer cooperation comparing to its complexity. Indeed, there are some proposals that compute global or local metrics which are used to make decisions for route establishment, scheduling, tuning transmission rate, etc. However, using these metrics in a cross-layer model could be not efficient because they have sometimes inaccurate values which do not reflect the real situation around a given node. Moreover, since a node moves with an arbitrary speed and toward an arbitrary destination, the computed metrics (according to the participation of the node in communication and the traffic load level around it) could change during the time. Consequently, other nodes that consider the metrics of that node, to build routes for example, could have an inaccurate information since this later change according to mobility patterns, traffic load, and links capacity.

We believe that developing a cross-layer model for QoS support in MANETs is somehow a “danger”. On one hand, the modifications, which have to be added in the protocol stack and the complexity in introducing a new parameters and new algorithms to provide a ”good” inter-layer cooperation, are usually introduce a high complexity risk. On the other hand, this could be very interesting given that it captures the characteristics of the capacity, the expected behavior of node load to choose the ”best route” between sources and destinations in a way to achieve a global traffic load balancing. In other cases this may be useful to have knowledge about neighbor density and ”quality” to adapt transmission rate and to use scheduling strategies in an efficient manner.

If we recapitulate, we can say that the cross-layer design is a promised solution to address QoS support and service differentiation in mobile ad hoc networks, but it is affected by the unpredicted mobility of wireless devices and so the “lifetime“ of the availability of the accurate available information. We recommend the following requirements to efficiently design a QoS cross-layer model which leads to the architecture shown in Figure 8.4:

1. **Choosing the metrics:** choosing of a very useful and efficient metrics such as battery level, available bandwidth, and mobility rate.

2. **Computing the metrics:** the way of computing these metrics regarding one path (energy, lifetime of nodes, throughput, delay, etc.) have to be decided. The well-known approach is to minimize a cost function for a given link in the path between a source and a destination then consider the different costs computed for all links in the path. Depending on the nature of the metric, the cumulative value could be additive, concave, or multiplicative. Other techniques could be also used such are variance and max-min. Computation and complexity costs should always be taken into account.

3. **Adapting metrics’ values:** an adaptive method should be used to update the measured metrics: They could be updated even more when mobility increases and less in a stable network while taking into account traffic load variation and application requirements.

4. **Deciding to use or not the metrics:** As shown in Figure 8.4, considering the information useful for model selection, the more efficient model has to be chosen according to the two following parameters:

    (a) **Regarding to the network behavior:** in some cases, when the traffic load and its characteristics change rapidly (high mobility), it is very difficult to compute accurate values of the metrics that can be used to address QoS. Hence, the complexity of the cross-layer model becomes too high comparing to the expected performance enhancement and it is recommended in this case to use the legacy layered approach.

    (b) **Regarding to the user application:** each layer of the protocol stack responding to local variations and information from other layers. We have to evaluate the benefits and the disadvantages of the cross layer model for each specific user application.

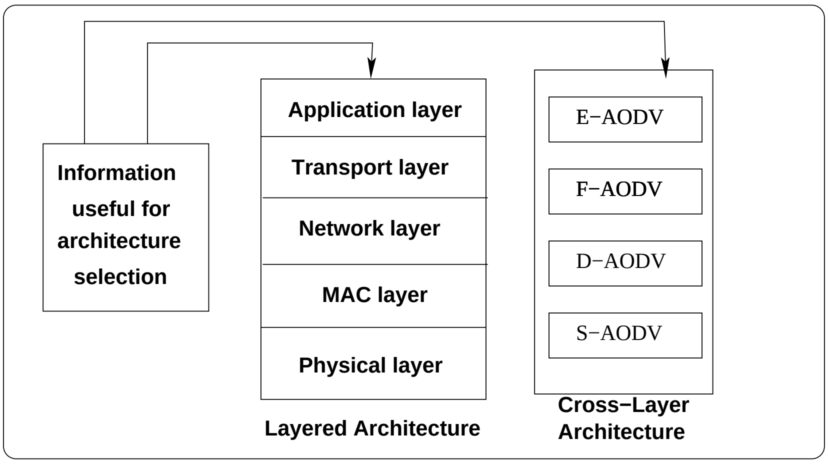

Figure 8.4: XAid: proposed new architecture design

The CrossAid architecture is the new cross-layer architecture shown in Figure 8.4. It uses the inter-layer routing mechanisms that we have proposed in this thesis. Each cross-layer routing scheme is used regarding the network characteristics and application requirements. The decision is made based after the analysis of the colleced measurement about the network stored in the bloc named "information useful for architecture selection" in Figure 8.4). This architecture is introduced not only to make a choice between layered or cross-layer architectures, but it contains implicitly the required information that we should consider to select the adequate cross-layer mechanism to use.

The design of the XAid architecture is derived from an extensive simulation-based (quantitative) analysis of our cross-layer routing schemes. We believe that we could extend this analysis by a qualitative study similar to what the authors of [3] have conducted for the GRACE (Global Resource Adaptation through CoopEration) cross-layer architecture. GRACE addresses the whole device and its resources using inter-layer cooperation. Every application that wants to make use of the GRACE framework must include some kind of cost model and should allow for multiple operation points. The cost model permits to predict resource consumption over time of the running application.

8.4 Chapter summary

This chapter starts by summarizing the performance results detailed in the previous chapters for all proposed cross-layer routing schemes. We then compared S-AODV and D-AODV routing mechanisms under several performance metrics. Finally, we describe the guidelines of designing a new cross-layer architecture called CrossAid, which implements our routing approaches described in this dissertation.

We believe that the decision to use which cross-layer routing mechanism is very coupled with the nature of the user application and the evolution of the network behavior. The very promising cross-layer design model consists in maintaining the layer isolation in the protocol stack while enabling a cross-layer interaction according to network and traffic characteristics. We have to establish whether cross-layer paradigm is suitable for all types of wireless networks and applications or not. Even if the answer is yes, it is necessary to maintain the layered approach, while enabling interactions between various protocols at different layers. Unless, the complexity of the new architecture could be expensive and inefficient regarding to minor performance enhancements as we have shown in the compared examples that we evoked in this chapter.

Chapter 9

Conclusion and outlook

Wireless ad hoc networks are collections of nodes communicating over a wireless channel. Since wireless signal power fades with distance and in the presence of obstructions, each node can communicate directly with only some of the other nodes that typically lie in its vicinity. Typically, the requirements of the traffic are assumed to be arbitrary and therefore is a necessity that nodes cooperate to forward packets to their final destinations. The lack of any wired infrastructure, the nature of the wireless channel, and the need for robustness and scalability create many challenging design problems in the link, network and higher layers of the OSI communication model.

Ad hoc networks have many characteristics that meet a lot of node heterogeneity. A fundamental issue in such multihop wireless environments is that network performance can degrade rapidly as the number of hops increases. Major problems to transmit data over available radio channels exist in every layer of the protocol stack. In one hand, issues like adaptive rate selection, adaptive antenna pattern, and adjust power control are addressed at the physical layer. In the other hand, link reliability, admission control, and access control to the shared channel are common issues of the network and the MAC layers. Moreover, there are several real-time application requirements that have to be guaranteed in order to provide QoS support and achieve service differentiation.

Intensive research work has been conducted by the Internet community to address the above mentioned issues separately. One new promising research direction to optimize data transfer in ad hoc networks is the cross-layer architecture. This new direction does not respect the original layered approach in which each layer operates independently. While the layered approach is simple, flexible, and scalable, it suffers from poor performance in ad hoc network even with the optimization applied to the evolved communication protocols. This is because, these protocols omit to consider the specific requirements of both the network and application constraints.

The aim of the conducted research activities was to propose and to validate new solutions based on the cross-layer design paradigm in order to enhance the QoS support in MANETs.

In this chapter, we first summarize briefly the contributions of this thesis work in Section 9.1. In Section 9.2, we discuss the future and long-term visions by enumerating several possible enhancements and extensions of our contributions, which we believe a fundamental

9.1 Summary of contributions and performance evaluation results

Both wireless and wireline networks need to support different types of network data service. The primary focus of this thesis is to study the support of these services over wireless networks and evaluate the gains that can be accrued by cross-layer techniques, which do away with the firm boundary that currently exists between the PHY and MAC layers, and the higher layers of the network protocol stack.

In this thesis we have sorted-out several open problems related to the design of communication protocols for QoS provisioning in MANETs. Compared to existing approaches, we have identified several new important metrics such as link stability and energy consumption rate that should be taken into account. By introducing these metrics we have proposed new routing mechanisms for MANETs and developed sophisticated algorithms to estimate their values in an accurate way.

Before presenting our new proposals by which data services can be supported in wireless networks using the cross-layer design paradigm, we described the limitations of each communication layer going from the application layer to the physical layer. We also discussed the most important previous research works based on exchanging useful information between adjacent layers. We made a first observation that despite of the performance improvement that this new design can achieve, there are some risks into changing the legacy layered architecture. Indeed, several issues, such as implementation, debugging, upgrading, and standardization, need to be addressed before the inter-layer interactions can be successfully exploited. Moreover, in this work, we justify whether the cross-layer paradigm is suitable for all types of wireless networks and applications or it may nor be useful for some of them under specific conditions. We strongly believe that it is necessary to maintain the layered approach, while enabling interactions between various protocols at different layers.

While working toward the proposition of new cross-layer mechanisms for the QoS support in MANETs, we have first focused on enhancing the EDCA 802.11e scheme. Hence, in the third chapter we proposed a new analytical model for the EDCA 802.11e mechanism and we detail several mechanisms to enhance this scheme for a better service differentiation without enforcing a cooperation with upper and lower layers. In the first part of the third chapter, we analyzed the performance of data burst transmissions, supported by the EDCA access mechanism. We developed an analytical framework to evaluate the impact of the $TXOP_{limit}$ on the overall performance. The analysis of the proposed framework allowed us to demonstrate that while the $TXOP$ parameter provides efficient service differentiation and preserves service to high priority traffic at high loads, it is especially prone to starving lower priority traffic. We suggested that the value of $TXOP_{limit}$ of the ACs, should be limited in order to guarantee a relatively satisfactory service level for both high priority and low priority classes. The model and the analysis provide an in-depth understanding and insights into the performance of the EDCA mechanism. They also provide helpful and powerful tools for further study, such as parameterization of the $TXOP_{limit}$ for some types of traffic characteristics for further QoS improvement for both WLANs and multihop networks. Indeed, EDCA parameters have to be properly set to provide prioritization of ACs and ensure minimum service guaranty to best effort traffics. Tuning them in order to meet specific QoS needs is a current research topic.

The second part of the third chapter described a new adaptive transmission opportunity scheme for QoS enhancement for IEEE 802.11 WLANs. To this end, we propose to adapt the value of $TXOP_{limit}$ as follow: at the beginning of each control period T, the TXOP duration is set according to the average packet size in each queue and its priority level. Then, we establish an analytic relationship between these values according to the medium utilization level while maintaining service differentiation. The goal is to enhance real-time applications and avoid starvation of low priority traffics.

We evaluated through simulations the performance of the ATXOP scheme and compared it with the basic EDCA. The results showed that our new approach outperforms the basic EDCA, especially at high traffic load conditions. Indeed, ATXOP increases efficiently the medium utilization ratio and it provides an overall goodput up to 25% higher than EDCA while achieving the expected delay differentiation.

We strongly believe that we could extend this scheme by considering the evaluation of its performances in multihop networks and providing interaction between MAC and routing protocols in order to tune the TXOP duration according to the network conditions that could be included in the routing packets. Therefore, the main key to QoS enhancement in wireless communications should be based on a coordination between all levels of the wireless protocol stack. Moreover, how to address an efficient cross layer QoS model based on energy conservation, stability, end-to-end delay, and other key metrics for real time applications. In other words, how to map between service differentiation and network QoS management.

In the second chapter, we investigated the QoS-based on EDCA scheme and we study analytically its performance in both WLAN and multihop network. Then, we introduced an adaptive MAC layer transmission opportunity based on EDCA approach to provide a good application performance and an efficient resource utilization in WLAN.

On-demand routing protocols are useful for mobile ad hoc network environment for their low routing overheads. However, they require to consider the reasons for link failure to improve its performance. Link failure caused by node mobility and lack of network resources. Therefore, it is essential to capture the aforesaid characteristics to identify the quality of links. Furthermore, the routing protocols that support QoS must be adaptive to cope with the time-varying topology and time-varying network resources. For instance, it is possible that a route that was earlier found to meet certain QoS requirements no longer does so, due to the dynamic nature of the topology. In such a case, it is important that the network intelligently adapts the session to its new and changed conditions. Indeed, it is not enough to find a shortest path but also with available resources as battery level. If battery energy is not taken into consideration in their design, it may lead to premature depletion of some nodes' battery leading to early network partitioning. Since computing complexity as well as communication establishment consumes significant power, we proposed in the fourth chapter a simple and efficient energy consumption rate based algorithm to establish routes between sources and destinations. Our performance evaluation using ns-2 simulator showed that the longevity of the network can be extended by a significant amount. Overall, we concluded that our mechanism demonstrates significant benefits at high traffic and high mobility scenarios. We expect that these scenarios will be common in ad hoc networking applications. Even though we implemented the algorithm on AODV, the technique is generic enough to be used with any on-demand protocol. Furthermore, we have also shown how our proposed mechanism can be added in the OLSR proactive MANET routing protocol.

Indeed, a large literature presents several works that have been done to optimize power consumption in ad hoc network. That considerable research has been devoted to low-power design of the entire network protocol stack of wireless networks in an effort to enhance energy efficiency. Therefore, the key to energy conservation in wireless communications should be based on a coordination between all levels of the wireless protocol stack. One of the main problems of wireless links is that all the nodes compete for the resources and channel access without taking into account knowledge about neighbor communications. Moreover, how to address an efficient QoS model based on MAC layer mode access and other key metrics for real time applications.

The proposed mechanism described in the fourth chapter can be classified as a *source-initiated and network-aided* approach as the source is responsible for selecting the path to be used and initiate the route determination/change procedure. Intermediate nodes along available paths toward the destination provide useful information to the source to help it selecting the best route. In the fifth chapter we have proposed another new scheme for routing protocols in MANETs, which is based on the energy consumption of nodes. Compared to the mechanism presented in the fourth chapter, this proposal allows intermediate nodes, and not only the source, to contribute on enhancement of the data forwarding process. This proposal, which we classified as a *source and network-initiated and assisted* approach, is based on a joint optimization of the routing and MAC protocols where the goal of the optimization is to achieve a good medium utilization, energy conservation, and provide a QoS enhancement of application performance.

The fifth chapter presented a cross-layer routing protocol called F-AODV, based on the co-operation between MAC and routing protocols. The objective of our proposal is to reduce the number of forwarding nodes (FN) in the network by considering their battery level and queue occupancy. On the other hand, it dynamically adapts MAC layer parameters of intermediate nodes, which allow them to achieve high probability of transmission success. The simulation results we obtained show that our model provides a total throughput significantly higher than the basic scheme. Besides, it provides a higher degree of fairness than the basic AODV protocol between competing user sessions. Furthermore, we compared the simulation results obtained with and without considering layer interconnections. To this end, we investigate the performance of AODV, E-AODV and F-AODV with different scenarios and network mobility patterns. The results that we got, show that the performance of the inter-layer cooperation paradigm depends on the network characteristics and the application constraints.

Globally, we show that E-AODV and F-AODV protocols perform better than the basic AODV protocol because they ensure a good routing management scheme for MANETs and prevent (or at least delay) data congestion in the network by traffic load balancing.

Obviously, the energy is one of the important metrics to consider when designing communication protocols for MANETs. However, there are other parameters, which have similar importance but they have been included only in limited number of new proposals in the open literature. One of them is the stability of wireless links, which measures how long each link between two each pair of nodes will be alive. In the seventh chapter, we proposed a new routing scheme, which takes into account the stability metric when establishing routes between wireless devices. We also develop the required algorithms to estimate in an accurate way this metric based on the measurements done at the MAC and network layers.

In the sixth chapter we proposed a cross-layer stability-based routing in mobile ad hoc networks. We described a new routing algorithm based on accurate stability parameters in dynamic network characteristics. In order to measure the stability of wireless links, each received frame at the MAC layer is processed even though the node is not the final destination. Information about the neighbor sending this frame is recorded and used by the routing protocol to increase the accuracy of the estimation of the stability vector, which contains the stability rate of each link with the neighbors. This stability vector is then used by the routing protocol to select the best next-hop to a given destination.

Our performance evaluation using ns-2 simulator show the importance of considering the stability information in route selection process. Overall, we conclude that our mechanism demonstrates significant benefits at high and unstable traffic scenarios. Even though we implemented the model in AODV, the technique used is very generic and can be used with any routing protocol such as DSR and OLSR. Furthermore, this proposal can be applied to single channel and multi-channel based medium access protocols, and there is no need for synchronization.

The stability metric is very important to select the best path between a source and a destination in MANETs. This can help also to reduce the packet losses due to link breakage. Another interesting metric, which is important mainly for delay-sensitive applications is the end-to-end delay. In the seventh chapter, we tackle the problem of the determination of routes by taking into account this metric and we develop a cross-layer routing scheme, which implements an intelligent algorithm operating between MAC and routing layer to select the path with the minimum end-to-end delay.

The seventh chapter presented a new MAC and routing cross-layer approach called D-AODV. The cooperation goal aims to provide a good delay and delay jitter for delay-sensitive applications but it could also be applied for other QoS parameters such as bandwidth, loss rate, and a route stability. The approach is an adaptive service differentiation based on buffer management and route establishment strategy. Our performance evaluation using ns-2 simulator showed the importance of considering the MAC-layer delays in route selection process. Overall, we conclude that our mechanism demonstrates significant benefits at high and unstable traffic scenarios. Even though we implemented the model in AODV, the proposed technique is generic enough to be easily included in any routing protocol such as OLSR by including the required information in Hello and TC (Traffic Control) control messages. Furthermore, this proposal can be applied to single channel as well as multi-channel based medium access protocols, and there is no need for hard synchronization between wireless devices. Additionally, the scheme could also be applied for the basic 802.11 that does support service differentiation at the link layer.

In the eighth chapter, we compared the performance obtained for the different routing schemes that we have proposed (E-AODV, F-AODV, S-AODV, and D-AODV) for low, meduim, and high mobility. The aim of this comparison is to demonstrate the importance of considering the applications' requirements as well as the network characteristics in selecting the most suitable (cross-layer) routing protocol for each user session. We also describe the guidelines for designing a new cross-layer framework, called **CrossAid (XAid)**, for QoS provisioning.

From the different contributions described in this dissertation, we conclude that the cross-layer paradigm is a good methodology to design new communication protocols for wireless networks in order to enhance the performance of real-time applications and achieve a better QoS support. However, the cooperative algorithms and the used metrics have to be rigorously selected, compared, and optimized regarding network characteristics and application requirements.

In summary, this thesis provides valuable insights and opens up a number of research directions for cross-layer deployment in fully distributed wireless environments. However, the scope of our work is intentionally limited to the cooperation between network and MAC layers for a better QoS provisioning in MANETs. We explore in the next section, the future directions of our work and we provide preliminary hints on how to address the issues which are not investigated in this dissertation and that we plan to tackle.

9.2 Future directions

There are several promising and important extensions to the research problems addressed in this thesis. Indeed, a number of issues remain open in the design of efficient schemes that evolve the cooperation of communication layers in order to improve the performance of the QoS support in MANETs and encourage the development of new applications for this kind of network. We will continue our approach of gaining insight using a combination of physical point of view.

First, we plan to investigate other optimization metrics such as the available bandwidth and the loss rate due to channel conditions.

Second, we aim to study the development of an unified metric which may consider more than one parameter. It is well known that finding the best path that optimizes two additive/multiplicative parameters is an NP hard problem. Hence, this issue remains a very crucial research area in the future.

Third, we will extend our first study of integrating the cross-layer routing proposals in OLSR routing protocol. We may also compare the performance of the new OLSR with the modified AODV when we incorporate the same routing optimization scheme.

Fourth, we will conduct a deep qualitative and quantitative study of our proposed architecture CrossAid as the GRACE architecture. Recall that the main target of GRACE refers to device as CPU consumption and not to the network.

Finally, we will explore the integration of other communications layers such as the physical and the transport layers in the routing decision.

Research Publications

- L. Romdhani and C. Bonnet, *A cross-layer feature for an efficient forwarding strategy in wireless ad hoc networks*, AINA 2006, 20th IEEE International Conference on Advanced Information Networking and Applications, April 18-20, 2006, Vienna, Austria.
- L. Romdhani and C. Bonnet, *A cross-layer on-demand routing protocol for delay-sensitive applications*, PIMRC 2005, 16th Annual IEEE International Symposium on Personal Indoor and Mobile Radio Communications, September 11 - 14, 2005, Berlin, Germany.
- L. Romdhani and C. Bonnet, *A cross-layer stability-based on-demand routing protocol for mobile ad-hoc networks*, IWWAN 2005, International Workshop on Wireless Ad-hoc Networks, May 23-26, 2005, London, U.K., 2004.
- L. Romdhani and C. Bonnet, *Achieving a Good Trade-off Between Complexity and Enhancement in Cross-layer Architectures,* 2nd IEEE International Conference on Information & Communication Technologies: from Theory to Applications - ICTTA2006 - April 24 - 28, 2006, Damascus, Syria.
- L. Romdhani and C. Bonnet, *Cross-layer's paradigm features in MANET: benefits and challenges*, PWC 2005, 10th IFIP International Conference on Personal Wireless Communications, August 25-27, 2005, Colmar, France.
- L. Romdhani and C. Bonnet, *Energy consumption speed-based routing for mobile ad hoc networks*, WWAN 2004, International Workshop on Wireless Ad Hoc Networking, March 23-26, 2004, Tokyo, Japan / Proceedings published in 24th International Conference on Distributed Computing Systems Workshops (ICDCS 2004 Workshops), IEEE Computer Society, 2004.
- L. Romdhani and C. Bonnet, *Adaptive transmission opportunity for QoS enhancement in (EDCA) IEEE 802.11e WLANs*, WMSCI 2005, 9th World Multi-Conference on Systemics, Cybernetics and Informatics, July 10-13, 2005 - Orlando, USA.
- L. Romdhani, C. Bonnet, *Energy consumption speed-based routing for mobile ad hoc networks*, Accepted for IJWMC, International Journal of Wireless and Mobile Computing special issue on Wireless Ad Hoc Networking, 2005.
- N. Qiang, L. Romdhani, and T. Turletti, *A Survey of QoS Enhancements for IEEE 802.11 Wireless LAN*, Journal of Wireless Communication and Mobile Computing (JWCMC), pp 1-20, 2004.
- L. Romdhani , Q. Ni, and T. Turletti*, Adaptive EDCF: Enhanced Service Differentiation for IEEE 802.11 Wireless Ad Hoc Networks.* IEEE WCNC'03 (Wireless Communications and Networking Conference), New Orleans, Louisiana, March 16-20, 2003.

Bibliography

[1] I. Aad and C. Castelluccia, Differentiation mechanisms for IEEE 802.11, In Proc. of IEEE Infocom 2001, April 2001.

[2] I. Aad and C. Castelluccia, Remarks on per-flow differentiation in IEEE 802.11, in Proc. of European Wireless 2002, February 2002.

[3] S. V. Adve et al., The Illinois GRACE Project: Global Resource Adaptation through CoopEration, workshop on Self-Healing, Adaptive and Self-MANaged Systems (SHAMAN), June 2002.

[4] S. Agarwal, A. Ahija, J. P. Singh, and R. Shorey, Route Lifetime Assessment Based Routing (RABR) Protocol for Mobile Ad-hoc Networks, In Proc. IEEE International Conference on Communications 2000 (ICC 00), V. 3, pp. 1697-1701.

[5] B. An and S. Papavassiliou, An Entropy-Based Model for Supporting and Evaluating Route Stability in Mobile Ad hoc Wireless Networks, IEEE Communications Letters, Vol. 6, No. 8, August 2002.

[6] G. Bianchi, Performance analysis of the IEEE 802.11 distributed coordination function, IEEE Journal on Selected Areas of Communication, vol. 18, no. 3, pp. 535-547, March 2000.

[7] A. Bruce McDonald and T. Znati, A Path Availability Model for Wireless Ad-Hoc Networks , In Proceedings of IEEE Wireless Communications and Networking Conference 1999 (WCNC 99).

[8] 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 Proceedings of the 4th International Conference on Mobile Computing and Networking (ACM MOBICOM'98), pages 85-97, October 1998.

[9] E. Carlson, and al., Performance Comparison of QoS Approaches for Ad Hoc Networks: 802.11e versus Distributed Resource Allocation, European Wireless, Nicosia, Cyprus, April 10-13, 2005.

[10] C. T. Calafate, P. Manzoni, M. P. Malumbres, On the Interaction Between IEEE 802.11e and Routing Protocols in Mobile Ad-Hoc Networks, 13th Euromicro Conference on Parallel, Distributed and Network-Based Processing (PDP'05), February 09 - 12, 2005 Lugano, Switzerland.

[11] H. S. Chahalaya and S. Gupa, Throughput and fairness properties of asynchronous data transfer methods in the IEEE 802.11 MAC protocol, PIMRC, pp 613-617, 1995.

[12] J. H. Chang and L. Tassiulas, Energy conserving routing in wireless ad-hoc networking, In Proceedings of the IEEE Infocom, pages 2231, Mar. 2000. ACM/IEEE.

[13] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks, ACM Wireless Networks, 8(5), September 2002.

[14] D. Chen, D. Gu and J. Zhang, Supporting Real-time Traffic with QoS in IEEE 802.11e Based Home Networks, Consumer Communications and Networking Conference (CCNC), pp. 205-209, January 2004.

[15] R. R. Choudhury and al., Using directional antennas for medium access control in ad hoc networks, ACM Mobicom atlanta, September 2002.

[16] T. H. Clausen and P. Jacket, Optimized Link State Routing (OLSR), IETF RFC 3626, October 2003.

[17] M. Conti, and al., Cross-Layering in Mobile Ad Hoc Network Design, Published by the IEEE Computer Society, February 2004.

[18] 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, pp. 3-12, Tel Aviv, Israel, March 2000.

[19] J. Deng and R. S. Chang, A priority Scheme for IEEE 802.11 DCF Access Method, IEICE Trans. in Comm., vol. 82-B, no. 1, January 1999.

[20] R. Dube et al., Signal Stability based adaptive routing for Ad Hoc Mobile networks, IEEE Personal Commumication, pp. 36-45, February 1997.

[21] R. Dube, C. D. Rais, K.-Y. Wang, and S. K. Tripathi, Signal stability based adaptive routing (SSA) for ad-hoc mobile networks, in IEEE Personal Communications, 1997.

[22] T. A. ElBatt, S. V. Krishnamurthy, D. Connors, and S. Dao, Power management for throughput enhancement in wireless ad-hoc networks, in IEEE International Conference on Communications, 2000, pp. 1506-1513.

[23] T. El-Batt and A. Ephremides, Joint scheduling and power control for wireless ad hoc networks, IEEE Transactions on Wireless Communications, Volume: 3, Issue: 1, pp. 74-85, January 2004.

[24] L. M. Feeney, An energy consumption model for performance analysis of routing protocols for mobile ad hoc networks, Mobile Networks and Applications Journal, V 6, N 3, pp. 239-249, 2001.

[25] M. Gerharz, C. de Waal, P. Martini and P. James, Strategies for Finding Stable Paths in Mobile Wireless Ad Hoc Networks, Proceedings of the 28th Annual IEEE Conference on Local Computer Networks (LCN), Konigswinter/Bonn, Germany, October 2003.

[26] M. Gerharz, C. de Waal, M. Frank, and P. Martini, *Link Stability in Mobile Wireless Ad Hoc Networks*, Proceedings of the 27th Annual IEEE Conference on Local Computer Networks (LCN), Tampa, Florida, November 2002.

[27] N. Gupta, S.R. Das, Energy-Aware On-Demand Routing for Mobile Ad Hoc Networks, IWDC 2002, Kolkata, India, December 2002.

[28] G. R. Hiertz and al. Throughput Analysis of IEEE 802.11e Wireless LANs and Efficient Block Ack Mechanism, 11th European Wireless Conference 2005.

[29] Y. Hu and D. B. Johnson, Exploiting Congestion Information in Network and Higher Layer Protocols in Multihop Wireless Ad Hoc Networks, ICDCS, Hachioji, Tokyo, Japan, March 24-26, 2004.

[30] D. He, J. Shengming, and R. Jianqiang, A Link Availability Prediction Model for Wireless Ad Hoc Networks, In Proc. of International Workshop on Wireless Networks and Mobile Computing, Taipei, Taiwan, April 2000.

[31] R. Hekmat and P.Van Mieghem, Degree distribution and hopcount in wireless ad-hoc networks, The 11th IEEE International Conference on Networks (ICON 2003), Sydney, Australia, Sept. 28-Oct. 1, pp. 603-609.

[32] IEEE 802.11.Standard 802.11: *Wireless LAN Medium Access Control (MAC) and Physical layer (PHY) Specifications*, 1999.

[33] IEEE 802.11e WG. Draft supplement to Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specications: MAC Enhancements for Quality of Service (QoS), IEEE Std 802.11e/D9.0, August 2004.

[34] IEEE 802.11 WG, Draft Supplement to Standard For Telecommunications and Information Exchange Between Systems-LAN/MAN Specific Requirements - Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), IEEE 802.11e/Draft 5.0, July 2003.

[35] V. Kawadia and P.R. Kumar, Power Control and Clustering in Ad Hoc Networks, in Proc. of Infocom 2003.

[36] E. S. Jung and N. H. Vaidya, A power control MAC protocol for ad-hoc networks, in ACM MOBICOM, 2002.

[37] C. E. Jones, K. M. Sivalingam, P. Agrawal, and J. C. Chen, A survey of energy efficient network protocol(I), Wireless Networks, vol. 7 (2001), pp. 343-358.

[38] P. Karn, MACA A New Channel Access Method for Packet Radio, ARRL/CRRL Amateur Radio 9th Computer Networking Conference 1990.

[39] Z. Kong and al., Performance analysis of IEEE 802.11e contention-based channel access, Selected Areas in Communications, IEEE Journal on Volume 22, Issue 10, Dec. 2004 Page(s):2095 - 2106.

[40] X. Li and Z. Bao-yu, Study on cross-layer design and power conservation in ad hoc network, Parallel and Distributed Computing, Applications and Technologies, PDCAT, pp. 324-328, 27-29 August 2003.

[41] Q. Li, J. Aslam, and D. Rus, Online power-aware routing in wireless ad-hoc networks, in Proc. of the Seventh Annual International Conference on Mobile Computing and Networking, pp. 97-107, July 2001.

[42] Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC), Enhancements for Quality of Service (QoS), 802.11e Draft 3.1, May, 2002.

[43] Perkins, Ad hoc on demand distance vector (AODV) routing, IETF, RFC 3561, Mai 2004.

[44] S. Mangold, S. Choi, P. May, O. Klein, G. Hiertz, and L. Stibor, IEEE 802.11e Wireless LAN for Quality of Service, In Proc. of European Wireless (EW2002), Florence, Italy, February 2002.

[45] J. P. Monks, V. Bhargavan, and W.-M. Hwu, A power controlled multiple access protocol for wireless packet networks, in Proc. of INFOCOM, 2001, pp. 219-228.

[46] S. Narayanaswamy, V. Kawadia, R. S. Sreenivas, and P. R. Kumar, Power control in ad-hoc networks: Theory, architecture, algorithm and implementation of the COMPOW protocol, in European Wireless Conference, 2002.

[47] Q. Ni, L. Romdhani, and T. Turletti, A Survey of QoS Enhancements for IEEE 802.11 Wireless LAN, Journal of Wireless Communication and Mobile Computing (JWCMC), pp. 1-20, 2004.

[48] Q. Ni, L. Romdhani, and T. Turletti, A Survey of QoS Enhancements for IEEE 802.11 Wireless LAN, Journal of Wireless Communication and Mobile Computing (JWCMC), 2004; 4: 1-20.

[49] Network Simulator-2. www.isi.edu/nsnam/ns/.

[50] R. Ramanathan and R. Rosales-Hain, Topology control of multihop wireless networks using transmit power adjustment, in Proc. of INFOCOM, 2000, pp. 404-413.

[51] RandomTrip: http://icawww1.epfl.ch/RandomTrip/

[52] L. Romdhani and C. Bonnet, Cross-layer's paradigm features in MANET: benefits and challenges, PWC 2005, 10th IFIP International Conference on Personal Wireless Communications, August 25-27, 2005, Colmar, France.

[53] L. Romdhani, Q. Ni, and T. Turletti, Adaptive EDCF: Enhanced Service Differentiation for IEEE 802.11 Wireless Ad Hoc Networks, IEEE WCNC 03 (Wireless Communications and Networking Conference), New Orleans, Louisiana, March 16-20, 2003.

[54] S. Selvakennedy, The Impact of Transmit Buffer on EDCA with Frame-Bursting Option for Wireless Networks, IEEE LCN'04, November 16 - 18, 2004.

[55] K. N. Sridhar, J. Lillykutty, and S. Rajeev, A Performance Evaluation and Enhancement of Link Stability Based Routing for MANETs, International workshop on mobile and wireless networking (MWN 2004), Montreal, Quebec, Canada, August 15, 2004.

[56] A. Shiozaki, Edge extraction using entropy operator, Comp. Vis., Graphics, Image Processing, vol. 36, pp. 1-9, 1986.

[57] A. Safwati, H. Hassanein, and H. Mouftah, Optimal cross-layer designs for energy-efficient wireless ad hoc and sensor networks, 2003. In the Proc. of the 2003 IEEE International Conference on Performance, Computing, and Communications, pp.123-128, 9-11 April 2003.

[58] V. Schaar, M. Krishnamachari, and al., Adaptive cross-layer protection strategies for robust scalable video transmission over 802.11 WLANs, IEEE Journal on Selected Areas in Communications, Vol. 21, N. 10, pp:1752 - 1763, December 2003.

[59] S. Singh and C. S. Raghavendra, Power efficient MAC protocol for multihop radio networks, in The Ninth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 1998, pp. 153-157.

[60] S. Singh and C.S. Raghavendra. PAMAS - power aware multi-access protocol with signalling for ad hoc networks. ACM Computer Communication Review (ACM CCR98), July 1998.

[61] S. Singh, M. Woo, and C. S. Raghavendra, Power aware routing in mobile ad hoc networks, in Proc. of ACM MOBICOM, 1998, pp. 181190.

[62] W. Su, S. Lee, and M. Gerla, Mobility Prediction and Routing in Ad Hoc Wireless Networks, International Journal of Network Management, Wiley & Sons, 11:3 30, 2001.

[63] M. W. Subbarao, Dynamic power-conscious routing for manets: An initial approach, in IEEE Vehicular Technology Conference, 1999, pp. 1232-1237.

[64] J. L. Sobrinho and A.S. Krishnakumar, Real-time traffic over the IEEE 802.11 medium access control layer, Bell Labs Technical Journal, 1996.

[65] Telecommunication Standardization Sector of ITU (1993) ITU-T Recommendation G.114. Technical report, International Telecommunication Union.

[66] I. Tinnirello and S.Choi, Efficiency Analysis of Burst Transmissions with Block ACK in Contention-Based 802.11e WLANs, ICC2005, Seoul Korea, 16-20 May 2005.

[67] C. K. Toh, Associativity-Based Routing for Ad Hoc Mobile Networks, International Journal of Wireless Personal Communications, Vol. 4, No. 2, pp. 103-139, March 1997 .

[68] C. K. Toh, Maximum Battery life Routing to support Ubiquitous Mobile Computing in Wireless Ad hoc Networks, IEEE Communications Magazine. June 2001.

[69] S. Toumpis, A. J. Goldsmith, Performance, optimization, and cross-layer design of media access protocols for wireless ad hoc networks, IEEE International Conference on Communications, vol. 3, pp.2234-2240 vol.3, 11-15 May 2003.

[70] A. Veres, et. al., Supporting service differentiation in wireless packet networks using distributed control, IEEE JSAC, Vol. 19, No. 10, pp. 2094-2104, October 2001.

[71] J. Wall and J.Y. Khan, An ARQ enhancement with QoS support for the 802.11 MAC protocol, Wireless Communications and Networking Conference, WCNC, March 2004.

[72] R. Wattenhofer, L. Li, P. Bahl, and Y.M. Wang, Distributed topology control for power efficient operation in multihop wireless ad hoc networks, in Proc. of INFOCOM, 1388-1397, 2001.

[73] Y. Xiao, Performance Analysis of Priority Schemes for IEEE 802.11 and IEEE 802.11e Wireless LANs, IEEE Transactions on Wireless Communications , Vol. 4, No. 4, July 2005.

[74] Y. Xu, J. Heidemann, and D. Estrin, Adaptive Energy-Conserving Routing for Multihop Ad Hoc Networks, Research Report527, USC/Information Sciences Institute, October, 2000.

[75] W. Yuen, H. Lee, and T.D. Andersen, A simple and effective cross layer networking system for mobile ad hoc networks, The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, vol. 4, pp. 1952-1956, 15-18 September 2002.

[76] J. Yoon and al., Sound mobility models, International Conference on Mobile Computing and Networking archive Proceedings of the 9th annual international conference on Mobile computing and networking, San Diego, CA, USA 2003.

[77] C. Zhao, W. B. Heinzelman, Adaptive Local Searching and Caching Strategies for On demand Routing Protocols in Ad Hoc Networks, international workshop on mobile and wireless networking (IMWN 2004), Montreal, Quebec, Canada, August 15, 2004.

List of Abbreviations

It's not what you say; it's what they hear.

AC	Access Category
AP	Access Point
ARQ	Automatic Repeat Request
AODV	Ad-hoc On-demand Distance Vector
CBR	Constant Bit Rate
CTS	Clear To Send
CFB	Collision Free Burst
CSMA	Carrier Sense Multiple Access
CSMA/CA	CSMA/Collision Avoidance
CSMA/CD	CSMA/Collision Detection
DCF	Distributed Coordination Function
E-AODV	Energy-AODV
ECN	Explicit Congestion Notification
EDCA	Enhanced Distributed Coordination Access
F-AODV	Forwarding-AODV
FDMA	Frequency Division Multiple Access
FIFO	First In First Out
GRACE	Global Resource Adaptation through CoopEration
HCF	Hybrid Coordination Function
IEEE	Institute of Engineering, Electrics, and Electronics

IETF	Internet Engineering Task Force
IP	Internet Protocol
MAC	Meduim Access Control
MACA	Multiple Access Collision Avoidance
MANET	Mobile Ad hoc Networks
OLSR	Optimized Link State Routing
PDR	Packet Delivery Rate
QoS	Quality of Service
QSTA	QOS STAtion
RFC	Request for Comments
RREQ	Route REQuest
RREP	Route REPLy
RTS	Ready To Send
RTT	Round Trip Time
S-AODV	Stability-AODV
TC	Traffic Category
TDMA	Time Division MultipleAccess
WiMAX	Worlwide Interoperability for Microwave Access
WLAN	Wireless Local Area Networks
WWAN	Wireless Wide Area Networks

Abstract

This dissertation focuses on the design, implementation, and evaluation of 802.11-based cross-layer mechanisms for the enhancement of the support of the QoS feature.

Before the cross-layer study, we have also explored the idea of enhancing separately a communications layer; namely the existing 802.11e MAC protocol which is designed for the QoS support. Although the improvements achieved, this study shows the limitations of the layered architecture that demonstrated its good performance in wired networks.

Due to the large number of cross-layer cooperation possibilities, we emphasize, in this work, on the cooperation between the MAC and the network layers. We believe that the cooperation between these two layers provides better performance improvement than the cooperation between other layers. In particular, we focus on the problem of routing data packets in a way that takes into account channel contention level, network characteristics, and higher-layer protocol requirements. We address the optimal routing with regard to links stability, average end-to-end delay, and energy conservation with and without assistance/initiation from the network. We design several cross-layer mechanisms that aim to overcome the issue of routing in MANETs while enhancing important QoS metrics (path stability, energy consumption, end-to-end delay, etc.). To this end, we extract the adequate parameters from both MAC and network layers and adapt them to provide QoS enhancement based on new inter-layer cooperation algorithms.

Furthermore, we identify the challenges that face the cross-layer architectures comparing to the traditional layered architecture for enhancing communication protocols.

Keywords: Quality of service in multihop mobile ad-hoc networks, cross-layer, cooperation, 802.11/e, QoS routing, energy conservation, stability issues, conception of architecture.

Résumé

Cette thèse se concentre sur la conception, le développement, et l'évaluation des mécanismes cross-layer basés sur le protocole 802.11 et qui visent à améliorer la qualité de service dans ce type de réseaux.

Avant d'étudier les pproches cross-layer, nous avons exploré l'idée d'améliorer, séparément, le fonctionnement d'une couche de communication; à savoir le protocole MAC 802.11e qui est conçu pour garantir la différentiation de service. Bien qu'on ait aboutit à des améliorations en utilisant un nouveau mécanisme adaptatif de la couche MAC, les résultats des études de performances montrent les limitations de l'architecture en couches qui a, par ailleurs, démontré ses bons résultats dans les réseaux filaires.

Il y a de nombreuses possibilités d'interaction inter-couche. Dans nos travaux, nous intéressons au partage de paramètres entre la couche MAC et la couche routage. Nous pensons que la coopération entre ces deux couches fournit une meilleure amélioration de performance que la coopération entre d'autres couches. En particulier, nous nous concentrons sur le problème de l'acheminement des paquets de données d'une manière qui prend en compte le niveau de congestion du canal, les caractéristiques de réseau, et les besoins de protocoles de couches hautes. Nous proposons un routage optimal en ce qui concerne la stabilité des liens, le délai de bout-en-bout, et la conservation d'énergie avec et sans assistance/initiation du réseau. Nous proposons plusieurs mécanismes cross-layer qui visent à résoudre le problème de routage dans MANETs tout en augmentant la performance des métriques de QoS importantes tel que la stabilité du chemin, la consommation d'énergie, le délai de bout-en-bout. À cet effet, nous identifions les paramètres adéquats à partir de la couche MAC et la couche réseau et nous les adaptons pour fournir de la qualité de service basée sur de nouveaux algorithmes de coopération inter-couches.

En outre, nous identifions les avantages et le inconvénients des architectures inter-couche par rapport à l'architecture en couche traditionnelle pour améliorer la performance de protocoles de communication dans un environnement sans fil.

Mots-clés: Qualité de service dans les réseaux ad-hoc multi-sauts, coopération inter-couche, 802.11/e, routage, conservation de l'énergie, stabilté, conception d'architecture.