

Université de Montréal

**Gestion adaptative des ressources dans les réseaux
maillés sans fil à multiples-radios multiples-canaux**

par
Jihene Rezgui

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Thèse présentée à la Faculté des arts et des sciences

en vue de l'obtention du grade de philosophiae doctor (Ph.D.)
en informatique

Août, 2010

© Jihene Rezgui, 2010.

Université de Montréal
Faculté des arts et des sciences

Cette thèse intitulée :

**Gestion adaptative des ressources dans les réseaux
maillés sans fil à multiples-radios multiples-canaux**

présentée par :
Jihene Rezgui

a été évaluée par un jury composé des personnes suivantes :

Fabian Bastin
président-rapporteur

Abdelhakim Hafid
directeur de recherche

Michel Gendreau
codirecteur

Patrick Soriano
membre du jury

Ahmed Karmouch
examineur externe

Fabian Bastin
représentant du doyen

Résumé

Depuis quelques années, la recherche dans le domaine des réseaux maillés sans fil ("Wireless Mesh Network (WMN)" en anglais) suscite un grand intérêt auprès de la communauté des chercheurs en télécommunications. Ceci est dû aux nombreux avantages que la technologie WMN offre, telles que l'installation facile et peu coûteuse, la connectivité fiable et l'interopérabilité flexible avec d'autres réseaux existants (réseaux Wi-Fi, réseaux WiMax, réseaux cellulaires, réseaux de capteurs, etc.). Cependant, plusieurs problèmes restent encore à résoudre comme le passage à l'échelle, la sécurité, la qualité de service (QoS), la gestion des ressources, etc. Ces problèmes persistent pour les WMNs, d'autant plus que le nombre des utilisateurs va en se multipliant. Il faut donc penser à améliorer les protocoles existants ou à en concevoir de nouveaux.

L'objectif de notre recherche est de résoudre certaines des limitations rencontrées à l'heure actuelle dans les WMNs et d'améliorer la QoS des applications multimédia temps-réel (par exemple, la voix). Le travail de recherche de cette thèse sera divisé essentiellement en trois principaux volets: le contrôle d'admission du trafic, la différenciation du trafic et la réaffectation adaptative des canaux lors de la présence du trafic en relève ("handoff" en anglais).

Dans le premier volet, nous proposons un mécanisme distribué de contrôle d'admission se basant sur le concept des cliques (une clique correspond à un sous-ensemble de liens logiques qui interfèrent les uns avec les autres) dans un réseau à multiples-sauts, multiples-radios et multiples-canaux, appelé RCAC. Nous proposons en particulier un modèle analytique qui calcule le ratio approprié d'admission du trafic et qui garantit une probabilité de perte de paquets dans le réseau n'excédant pas un seuil prédéfini. Le mécanisme RCAC permet d'assurer la QoS requise pour les flux entrants, sans dégrader la QoS des flux existants. Il

permet aussi d'assurer la QoS en termes de longueur du délai de bout en bout pour les divers flux.

Le deuxième volet traite de la différenciation de services dans le protocole IEEE 802.11s afin de permettre une meilleure QoS, notamment pour les applications avec des contraintes temporelles (par exemple, voix, visioconférence). À cet égard, nous proposons un mécanisme d'ajustement de tranches de temps ("time-slots"), selon la classe de service, ED-MDA (Enhanced Differentiated-Mesh Deterministic Access), combiné à un algorithme efficace de contrôle d'admission EAC (Efficient Admission Control), afin de permettre une utilisation élevée et efficace des ressources. Le mécanisme EAC prend en compte le trafic en relèvements et lui attribue une priorité supérieure par rapport au nouveau trafic pour minimiser les interruptions de communications en cours.

Dans le troisième volet, nous nous intéressons à minimiser le surcoût et le délai de re-routage des utilisateurs mobiles et/ou des applications multimédia en réaffectant les canaux dans les WMNs à Multiples-Radios (MR-WMNs). En premier lieu, nous proposons un modèle d'optimisation qui maximise le débit, améliore l'équité entre utilisateurs et minimise le surcoût dû à la relèvements des appels. Ce modèle a été résolu par le logiciel CPLEX pour un nombre limité de nœuds. En second lieu, nous élaborons des heuristiques/méta-heuristiques centralisées pour permettre de résoudre ce modèle pour des réseaux de taille réelle. Finalement, nous proposons un algorithme pour réaffecter en temps-réel et de façon prudente les canaux aux interfaces. Cet algorithme a pour objectif de minimiser le surcoût et le délai du re-routage spécialement du trafic dynamique généré par les appels en relèvements. Ensuite, ce mécanisme est amélioré en prenant en compte l'équilibrage de la charge entre cliques.

Mots clés : Réseaux maillés sans fil, contrôle d'admission, différenciation de services, réaffectation des canaux, trafic en relève, gestion adaptative, qualité de service, méta-heuristiques.

Abstract

In the last few years, Wireless Mesh Networks (WMNs) area brought a new field of advanced research among network specialized scientists. This is due to the many advantages which WMN technology offers, such as: easy and inexpensive installation, reliable connectivity and flexible interoperability with other existing networks (Wi-Fi, WiMax, Cellular, Sensors, WPAN networks, etc.). However, several problems still remain to be solved such as the scalability, the security, the quality of service (QoS), the resources management, etc. These problems persist for WMNs, therefore the researchers propose to improve the existing protocols or to conceive new protocols for WMNs.

In order to solve some of the current limitations met in the wireless networks and to improve QoS of real time multimedia applications in such networks, our research will be divided primarily into three parts: traffic admission control, traffic differentiation and handoff-aware channel assignment schemes.

In the first part, we propose a distributed admission control scheme for WMNs, namely, Routing on Cliques (a clique is defined as a subset of logical links that interfere with each other) Admission Control (RCAC). Particularly, we propose an analytical model to compute the appropriate acceptance ratio and guarantee that the packet loss probability in the network does not exceed a threshold value. The model also allows computing end-to-end delay to process flow requests with delay constraints.

In the second part, we design an efficient scheduler for Mesh Deterministic Access (MDA) in IEEE 802.11s-based WMNs, called Enhanced Differentiated-MDA (ED-MDA) to support voice and video applications with strict requirements

on delay and on blocking/dropping probability. ED-MDA together with Enhanced Admission Control, namely EAC, reserves the minimum amount of necessary resources while maintaining an acceptable handoff call dropping and high resource utilization.

The final section addresses handoff-aware channel assignment (CA) problem in Multiple Radios WMNs (MR-WMNs). In this section, we first propose a multi-objective optimization model that, besides maximizing throughput, improves fairness and handoff experience of mesh clients. In this model, the Jain's index is used to maximize users' fairness and to allow same channel assignments to links involved in the same high handoff traffic, thus reducing handoff-triggered re-routing characterized by its high latency. Second, we solved this model to obtain exact solutions by the CPLEX software for a limited number of nodes. We therefore propose to use centralized heuristics/meta-heuristics algorithms as an offline CA process to obtain near-optimal solutions for larger instances (real size network). Moreover, in order to adapt to traffic dynamics caused especially by user handoffs, an online CA scheme is proposed that carefully re-assigns channels to interfaces with the purpose of continuously minimizing the re-routing overhead/latency during user handoffs. This online scheme is improved using load balancing.

Keywords: Wireless mesh network, admission control, service differentiation, channel assignment, handoff, network management, quality of service, meta-heuristics.

Table de Matières

Résumé.....	1
Abstract	4
Liste des Tableaux	10
Liste des Figures	11
Glossaire et Traductions	14
Liste des Sigles et Abréviations	15
Liste des Définitions	18
Remerciements.....	19
Dédicace.....	21
Chapitre 1: Introduction.....	22
1.1 Contexte de recherche	22
1.1.1 Architecture des WMNs	23
1.1.2 Applications réelles déployant WMNs	26
1.2 Motivations et objectifs	27
1.3 Contributions et organisation de la thèse.....	29
1.4 Articles publiés/soumis durant cette thèse.....	31
Chapitre 2: Revue de la Littérature.....	33
2.1 Introduction	33
2.2 Les mécanismes d'accès au média sans fil.....	34

2.2.1 Mécanisme d'accès distribué DCF pour la norme IEEE 802.11	34
2.2.2 Mécanisme d'accès EDCA pour la norme IEEE 802.11e	36
2.2.3 Mécanisme d'accès optionnel MDA pour la norme IEEE 802.11s	38
2.2.4 Récapitulatif et comparaison des mécanismes d'accès.....	40
2.3 Qualité de service dans les réseaux maillés sans fil	41
2.3.1 Modèles de qualité de service pour les réseaux ad hoc	42
2.3.2 Dimensionnement des WMNs: augmenter la capacité	45
2.3.3 L'importance de considérer et de minimiser les interférences	46
2.3.4 Contrôle d'admission.....	49
2.3.5 Gestion adaptative.....	54
2.3.6 Différentiation de services dans les réseaux sans fil.....	56
2.3.7 Routage dans les reseaux sans fil.....	58
2.4 Les approches existantes d'affectation des canaux	62
2.4.1 Affectation fixe des canaux (AFC).....	63
2.4.2 Affectation dynamique des canaux (ADC).....	63
2.4.3 Affectation hybride des canaux (AHC)	64
2.4.4 Optimisation des différents critères dans les WMNs.....	64
2.4.5 Récapitulatif et comparaison des approches d'affectation des canaux	65
2.5 Synthèse.....	67
Chapitre 3: Distributed Admission Control in Wireless Mesh Networks: Models, Algorithms, and Evaluation	69
Abstract.....	69
3.1 Introduction	70
3.2 Related work.....	72
3.3 Network Model.....	74
3.3.1 Connectivity Graph.....	74
3.3.2 Cliques	76
3.3.3 Assumptions/Notations	78
3.3.4 Queuing Model	80
3.4 Proposed Analytical Model: Computing the Clique-Acceptance Ratio and End-to-End Delay	84
3.4.1 Traffic Characterization.....	84
3.4.2 PLP.....	86

3.4.3 Computing the Acceptance Ratio	90
3.4.4 Computational Complexity of Acceptance Ratio Computing	91
3.4.5 Computing the Delay in Each A-Clique	91
3.5 Routing on Cliques Admission Control Algorithm.....	93
3.6 Simulation Results.....	97
3.6.1 Simulation Configurations	97
3.6.2 Results Analysis.....	99
3.7 Conclusions and Future Work	107
Abstract.....	109
4.1 Introduction	110
4.2 Related work.....	112
4.2.1 Problems with EDCA QoS Guarantees in IEEE 802.11s	113
4.2.2 Problems with MDA-based Schemes to Guarantee QoS in IEEE 802.11s	114
4.3 Overview of ED-MDA	115
4.3.1 Assumptions and Notations	115
4.3.2 MDA without Traffic Differentiation	118
4.3.3 MDA with Traffic Differentiation: ED-MDA	119
4.4 The Real Time Connections Analysis in the ED-MDA per-flow Scheme....	122
4.4.1 Impact of MDAOP Periodicity and CFP Period.....	122
4.4.2 Bounds Conditions for CFP Adaptive Adjustment.....	123
4.5 Admission Control Scheme EAC	127
4.5.1 Possible CFP Values' Update Based on EAC Operation	127
4.5.2 Time-Slots Reservation over ED-MDA by EAC Algorithm.....	129
4.5.3 Impact of the State History Length: S	131
4.6 Simulation Results.....	131
4.6.1 Simulation Configurations	132
4.6.2 Results Analysis.....	133
4.7 Conclusions and Future Work	142
4.8 Acknowledgment.....	142

Chapitre 5: Optimization Model for Handoff-Aware Channel Assignment Problem for Multi-Radio Wireless Mesh Networks.....	143
Abstract.....	143
5.1 Introduction	144
5.2 Related Work.....	146
5.3 Formulation	149
5.3.1 Problem Description	149
5.3.2 Network Model.....	150
5.3.3 Cliques	151
5.3.4 Problem Formulation	152
5.4 Solving The Model.....	157
5.4.1 Heuristic Based on Greedy Algorithm.....	158
5.4.2 A Variable Neighborhood Search Method	159
5.4.3 A Tabu Search Algorithm.....	162
5.5 Proposed On-Line Channel Assignment Scheme.....	164
5.5.1 Channel Assignment Trigger.....	164
5.5.2 Online Channel Assignment Algorithm	165
5.6 Simulation Results.....	167
5.6.1 Case I: Known Handoff traffic	168
5.6.2 Case II: Dynamic CA with and without HALBTH	174
5.6.3 Case III: Dynamic CA based heuristics MHALB+/MTABU+	178
5.7 Conclusions and Future Work.....	182
Chapitre 6: Conclusion et recherches futurs	183
6.1 Résumé des contributions.....	184
6.2 Directions pour les travaux futurs	186
Bibliographie.....	188

Liste des Tableaux

Table 1. Comparaison des méthodes d'accès proposées dans la littérature	40
Table 2. Récapitulatif des travaux sur le contrôle d'admission.....	53
Table 3. List of Symbols/Parameters.....	79
Table 4. Changes of local maximum occupancy.....	83
Table 5. Simulation parameters	98
Table 6. Notations parameters	117
Table 7. QoS requirement: 4 service classes	119
Table 8. QoS requirement: 4 new service classes.....	121
Table 9. Alpha adjustment according to the blocking probability constraint state	128
Table 10: Simulation parameters	133
Table 11. Notations parameters and variables.....	152
Table 12. CPU time (sec) with HALBTH, MHALB and CPLEX: 10-nodes network.....	170
Table 13. CPU time (sec) with MHALB+, MATBU+ and CPLEX: 19-nodes network.....	181

Liste des Figures

Fig. 1. Architecture client.....	24
Fig. 2. "Backbone" d'un réseau maillé [Ch06].....	25
Fig. 3. Architecture hybride [Ch06]	26
Fig. 4. "Broadband Home Network" [AW05].....	27
Fig. 5. Accès au média sans fil en mode DCF (IEEE 802.11)	35
Fig. 6. Accès au média sans fil en mode EDCA (IEEE 802.11e)	37
Fig. 7. MDAOPs durant l'intervalle DTIM (802.11s) [So07].....	39
Fig. 8. Le modèle FQMM.....	43
Fig. 9. Le modèle SWAN [GC02].....	44
Fig. 10. Diffusion radio à une distance r	47
Fig. 11. Deux zones: de transmission (r) et de CSR ($[r, R]$)	48
Fig. 12. Différentiation de services	57
Fig. 13. Une taxonomie des approches d'affectation des canaux dans les WMNs.....	63
Fig. 14. Modèle du support de la QoS de notre approche dans les WMNs.....	68
Fig. 15. A wireless mesh network	75
Fig. 16. Connectivity graph: G_A	76
Fig. 17. A-cliques and B-cliques	77
Fig. 18. Traffic scenario: an example.....	83
Fig. 19. Inter-CH interactions.....	94
Fig. 20. Delay variation with MDA versus CA.....	99
Fig. 21. Throughput variation with MDA versus CA	100
Fig. 22. Loss variation with MDA versus CA.....	102
Fig. 23. Outage variation with MDA versus CA.....	103
Fig. 24. Blocking variation with MDA versus CA.....	104
Fig. 25. Overhead variation with MDA versus CA.....	105

Fig. 26. Blocking, packet loss probability of RCAC versus W.O.RCAC for different network sizes.....	106
Fig. 27. A typical WMN where transmission cliques vs. interferences cliques are shown, topology of Mesh routers mutually interfering within cliques	116
Fig. 28. (a) MDA per-flow schedule; and (b) ED-MDA per-flow schedule	119
Fig. 29. Main fields of the call request packet in IEEE 802.11s (ED-MDA)	122
Fig. 30. The simulated topology	132
Fig. 31. Packet loss (%) for voice calls using different access Methods.....	133
Fig. 32. Packet loss (%) for video calls using different access methods	134
Fig. 33. Delay for voice calls using different access methods	135
Fig. 34. Delay for video calls using different access methods	136
Fig. 35. HO calls dropping probability with and without ED-MDA.....	137
Fig. 36. New calls blocking probability with and without ED-MDA	137
Fig. 37. Impact of the beta value on the HO and new calls delay for an offered load equal to 10 Mbps with ED-MDA (S=2)	138
Fig. 38. Impact of β on the HO and new calls delay for an offered load equal to 10 Mbps with ED-MDA (S=10).....	139
Fig. 39. Impact of β on the HO and new calls Blocking/Dropping probability for an offered load equal to 10 Mbps with ED-MDA (S=2)	139
Fig. 40. Impact of the history length on the HO and new video calls delay for an offered load equal to 10 Mbps and β equal to 1.9 with ED-MDA	140
Fig. 41. Impact of the history length on the HO and new voice calls delay for an offered load equal to 10 Mbps and β equal to 1.9 with ED-MDA	141
Fig. 42. Impact of the history length on the HO and new calls blocking/dropping probability for an offered load equal to 10 Mbps and β equal to 2.1 with ED-MDA	141
Fig. 43. A 2-radios MR-WMN with 3 orthogonal 802.11 frequencies assignment	150
Fig. 44. An example of MR-WMN CA schemes	151
Fig. 45. Node move	160
Fig. 46. Cliques exchanges	161

Fig. 47. Delay with HALBTH, MHALB and optimal solution.....	168
Fig. 48. Packet loss rate with HALBTH, MHALB and optimal solution	169
Fig. 49. Throughput with HALBTH, MHALB and optimal solution	169
Fig. 50. Delay with MHALB.....	171
Fig. 51. Throughput with MHALB	172
Fig. 52. Loss with MHALB.....	172
Fig. 53. Jain's fairness index	173
Fig. 54. Throughput achieved by MCs (0,0,0,1)	173
Fig. 55. Throughput achieved by MCs (1/3,1/3,1/3,0).....	174
Fig. 56. Delay (a); packet loss ratio (b); throughput (c).....	175
Fig. 57. Delay (a); packet loss ratio (b); throughput (c).....	176
Fig. 58. Delay (a) and packet loss rate (b) with and without HALBTH+	178
Fig. 59. Delay with MHALB+ and MTABU+	179
Fig. 60. Throughput with MHALB+ and MTABU+.....	179
Fig. 61. Loss rate with MHALB+ and MTABU+	180

Glossaire et Traductions

Assignment	Affectation
Backbone WMN	Réseau fédérateur maillé sans fil
Carrier Sensing Range	Portée de l'interférence
Gateway	Passerelle
Cluster	Groupe
Custer-head	Chef de groupe
Handoff	Trafic en relève
Medium Access Reservation	L'accès au medium est avec réservation
Medium	Média de transmission
Mesh Client	Client maillé
Mesh Router	Routeur maillé
Passive Clustering	Regroupement passif
Route Discovery	La découverte d'une route
Routing loop	Boucle de routage
Software QoS	QoS logiciel
Time-slots	Tranches de temps
Transmission range	Portée de la transmission
WMN	Réseau maillé sans fil

Liste des Sigles et Abréviations

AC	Access Category
AP	Access Point
AFC	Affectation Fixe des Canaux
ADC	Affectation Dynamique des Canaux
AHC	Affectation Hybride des Canaux
ATM	Asynchronous Transfer Mode
AODV	Ad hoc On Demand Distance Vector routing protocol
CA	Contention Access
CBR	Constant Bit Rate
CFP	Contention Free Period
CH	Clique Head
CP	Contention Period
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CSR	Carrier Sensing Range
CTS	Clear-To-Send
CW	Contention Window
DCF	Distributed Coordination Function
DiffServ	Differentiated Services
DSDV	Distance Source Distance Vector routing protocol
DTIM	Delivery Traffic Indication Message
EAC	Efficient Admission Control
EDCA	Enhanced Distribution Channel Access
ED-MDA	Enhanced Differentiated-Mesh Deterministic Access
FFMC	Fulkerson-based Matrix Computation

FTP	File Transfer Protocol
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IntServ	Integrated Services
IP	Internet Protocol
MAC	Medium Access Control
MANET	Mobile Ad hoc Network
MC	Mesh Client
MDA	Mesh Deterministic Access
MDAOPs	MDA opportunities
MR	Mesh Router
MR-WMN	Multiples Radios-Wireless Mesh Network
MPR	Multi-Points relays
NAV	Network Allocation vector
NLOS	Non-Line Of Sight
ns-2	Network Simulator-2
OBS	Optical Burst Switching
OLSR	Optimized Link State Routing
PHY	PHYsical layer
PCF	Point Coordination Function
P-CS	Physical sensing range
PLP	Packet loss probability
QoS	Quality of Service
QdS	Qualité de Service
RCAC	Routing on Cliques Admission Control
RREQ	Route REQuest
RSVP	Resource ReSerVation Protocol

RTS	Request-To-Send
TC	Topology Control
TCP	Transmission Control Protocol
TOXP	Transmission Opportunity
TS	Traffic Spot
UDP	User Datagram Protocol
V-CS	Virtual Carrier Sensing
VoIP	Voix sur IP
WMN	Wireless Mesh Network

Liste des Définitions

Affectation	Une affectation des canaux aux cartes radio de chaque nœud.
Canal de communication radio	Un média ("medium") de transmission d'information sans fil. Il relie la source au destinataire et permet l'acheminement du message. Par exemple, ce canal est partagé par les nœuds selon l'algorithme CSMA/CA dans IEEE 802.11.
Interface radio	Chaque nœud est supposé équipé par une ou plusieurs cartes radio.
Clique	Ensemble de nœuds adjacents deux à deux.
Clique maximale	Clique pour laquelle l'ajout de n'importe quel autre nœud du graphe fait perdre la propriété de clique (adjacence deux-à-deux des nœuds).
Clique A	Nœuds partageant un ou plusieurs canaux et dont la distance entre eux est inférieure à r .
Clique B	Nœuds partageant un ou plusieurs canaux et dont la distance entre eux inférieure à R et strictement supérieure à r .
Nœud	Peut être MC ou MR, équipé de multiples-radios, il peut communiquer sur un ou plusieurs canaux.
r	La portée de la transmission, souvent fixée à 250 mètres. Les nœuds qui sont à une distance inférieure à r peuvent transmettre et recevoir entre eux.
R	La portée de l'interférence, souvent fixée à 550 mètres. Les nœuds qui sont à une distance comprise entre R et r sont en mesure de détecter l'activité sur le nœud mais pas en mesure de tirer l'information.
Route	Une suite de canaux radio.
Requête	Une demande du trafic entre source et destination.
Routing	Attribuer une route à une requête.
Différentiation de services	Classer les trafics selon leurs priorités.

Remerciements

Au terme de ce travail de longue haleine, je tiens à remercier tous celles et ceux qui m'ont aidée de près ou de loin à mener à bien ma recherche, tous ceux ou celles qui étaient disponibles et ont participé à rendre plus humaine, plus riche et enrichissante cette longue expérience. Mes remerciements iront aussi bien à mes enseignants qu'au cadre administratif de notre département, mais aussi à ma famille, mon époux et à mes ami(e)s.

Je remercie tout d'abord mon directeur de recherche M. Abdelhakim Hafid, qui, tout au long de quatre ans, m'a appuyée et m'a encouragée à donner le meilleur de moi-même. Je lui exprime toute ma gratitude pour avoir réussi à stimuler continuellement mes capacités et mon potentiel. Il a su être réconfortant et généreux sans jamais laisser de côté la rigueur ni la précision qui incombent à la recherche.

Je tiens également à remercier M. Michel Gendreau pour son encadrement, ses conseils laborieux et son soutien qui m'ont permis de mener à bien ce projet de thèse.

Mes vifs remerciements sont aussi adressés à tous les membres du jury : M. Fabian Bastin, professeur au département d'informatique et recherche opérationnelle à l'université de Montréal, M. Patrick Soriano, professeur à l'HEC Montréal et M. Ahmed Karmouch, professeur à l'université d'Ottawa, pour m'avoir fait l'honneur d'accepter d'évaluer mon projet de thèse.

Je remercie également tous mes collègues de laboratoire pour leur collaboration active et leur présence.

Je ne manquerais pas de remercier mon cher époux Radhouane pour son irremplaçable et inconditionnel soutien. Il a été toujours à mes côtés surtout lors des moments les plus difficiles de cette thèse.

Enfin, je n'oublierais pas d'exprimer toute ma reconnaissance à certains membres de ma famille (ma chère et adorable mère, mon frère Aziz et mes sœurs) et à mes ami(e)s proches, qui m'ont soutenue, ne serait-ce que par le cœur, et qui ont su être présent(e)s quand j'ai eu besoin d'eux.

Dédicace

Je dédie ce modeste travail à la mémoire de mon père. J'espère que, du monde qui est sien maintenant, il apprécie ce geste comme preuve de ma profonde reconnaissance.

Chapitre 1: Introduction

1.1 Contexte de recherche

Les réseaux maillés sans fil (Wireless Mesh Networks-WMNs) sont une classe émergente des réseaux sans fil. Ils reprennent le principe d'un réseau sans fil à multiples-sauts, c'est-à-dire que les communications entre deux nœuds peuvent être supportées par plusieurs nœuds intermédiaires, (appelés Mesh Router-MR) dont le rôle est de relayer l'information d'un point à un autre [AW05]. Ces MRs s'organisent d'une manière autonome de façon à former un réseau fédérateur maillé sans fil ("Backbone WMN"). Ils jouent également le rôle d'une passerelle/pont d'accès vers différentes technologies sans fil (réseaux Wi-Fi, réseaux WiMax, réseaux cellulaires, réseaux de capteurs, réseaux privés sans fil, etc.). Les nœuds MRs ont une mobilité très réduite, voire nulle, définissant ainsi un réseau qui, par rapport aux nœuds des réseaux ad hoc mobiles (Mobile Ad hoc Networks-MANET), n'ont ni contraintes d'autonomie en énergie ni problèmes de maintien de la connectivité. Les WMNs permettent d'intégrer plusieurs technologies sans fil et ont donc pour vocation de jouer le rôle d'un réseau d'accès vers l'Internet ou vers différents types de réseaux sans fil.

La communication entre équipements dans les WMNs est NLOS (Non-Line Of Sight) [Ch06], c'est-à-dire que ces équipements communiquent même s'il y a des obstacles sur le chemin LOS (Line Of Sight), par exemple des immeubles, des arbres ou des collines. La caractéristique qui distingue la technologie WMN est sa topologie dynamique, qui lui permet d'être dynamiquement auto-configurable, auto-curative (préventive). Dans le cas où un chemin tomberait en panne, un nouveau chemin prend sa place pour assurer la connectivité dans le réseau. Par conséquent, les WMNs sont fortement fiables. Tous ces dispositifs

permettent aux réseaux WMNs d'être intéressants grâce à l'automatisation de la maintenance du réseau, aux bas coûts de la construction et aux services de couverture robustes et fiables. Les réseaux WMNs sont caractérisés par la capacité de chaque participant d'agir à la fois comme client et comme routeur du réseau. Si un émetteur n'est pas à portée directe de la machine destination, les informations devront être transmises de proche en proche, le long d'un chemin établi et maintenu par le réseau en cas de modifications de la topologie (ces réseaux sont d'ailleurs qualifiés de multiples-sauts).

Les WMNs sont donc amenés à être le support de communication d'une grande diversité d'applications ayant différentes exigences de qualité de service (QoS) en termes de délai, de débit, de fiabilité, de confidentialité, etc. Ces applications concernent de nombreux domaines, tels que la domotique¹, les systèmes embarqués, les systèmes de santé, etc. Par conséquent, de nombreuses applications peuvent être développées grâce à la technologie WMN. Plusieurs compagnies ont déjà commercialisé les produits WMNs dans le marché Télécom. De ce fait, plusieurs villes ont mis en place les WMNs, citons comme exemple l'architecture MetroMesh [Ch06] qui est déployée dans la ville de Chaska au Minnesota. Les réseaux WMNs sont considérés comme l'une des clés de la technologie réseaux sans fil pour la prochaine génération.

Dans cette section, nous avons commencé par présenter les avantages des WMNs. Trois types d'architectures de systèmes seront exposés suivis de quelques exemples d'applications réelles déployant les WMNs.

1.1.1 Architecture des WMNs

Les réseaux WMNs se composent de deux types de nœuds : les MRs qui possèdent, en plus des fonctionnalités communes d'un routeur câblé, des

¹ Ensemble de techniques adjoignant à celles du bâtiment les ressources de l'électronique, de l'informatique et des télécommunications pour améliorer l'habitat humain (ref. www.granddictionnaire.com).

fonctionnalités additionnelles pour supporter le routage dans les WMNs. Ils sont aussi dotés de multiples-interfaces sans fil pour se connecter aux différentes technologies sans fil. Grâce aux fonctionnalités passerelle/pont ajoutées aux MRs, les WMNs incorporent les réseaux existants. Le deuxième type est un client maillé sans fil (MC). Il a deux rôles : un utilisateur final ou/et un routeur avec le minimum de fonctionnalités. Ces MCs ont une seule interface sans fil [Ch06]. L'architecture de WMN est classifiée selon trois modèles [Ch06] : architecture client, architecture "backbone" d'un réseau maillé et architecture hybride.

- L'architecture client est constituée de nœuds clients seulement. Ceux-ci jouent un double rôle : utilisateur final et routeur (cf. Fig. 1). Cette architecture permet une communication "peer to peer" à travers tous les nœuds du réseau. Ce type de réseaux est plus similaire aux réseaux conventionnels ad hoc puisqu'une seule technologie radio est utilisée. Cependant, ces nœuds clients doivent être équipés par plus de logiciels et de matériels pour supporter le routage.

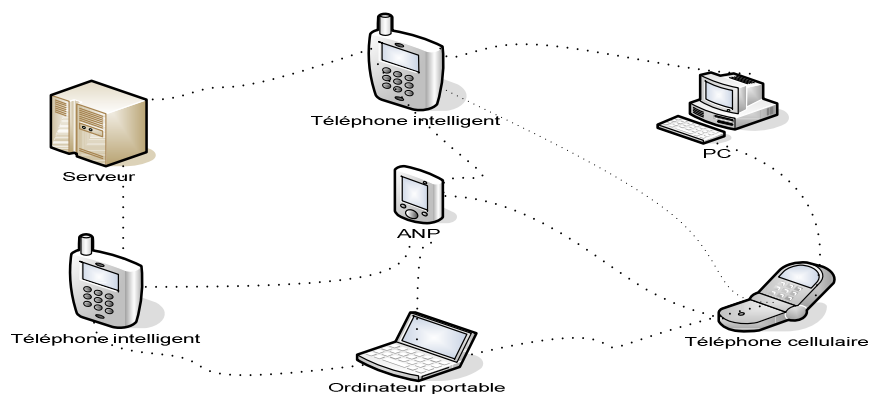


Fig. 1. Architecture client

- L'architecture "backbone" d'un réseau maillé est constituée des MCs et des MRs (cf. Fig. 2). Les MRs constituent le "backbone" de

l'infrastructure pour les MCs en leur permettant la connectivité. Ils réalisent les fonctions, tels que le routage, l'auto-configuration, l'auto-prévention. Plusieurs technologies radio peuvent être incorporées à ce type de réseaux maillés, par exemple 802.11, 802.16, etc. Grâce aux fonctionnalités pont/passerelle incorporées dans les MRs, cette infrastructure fournit une interface pour l'intégration des réseaux sans fil existants. Si les MCs utilisent la même interface radio, ils peuvent communiquer directement avec les MRs. Dans le cas où les clients n'auraient qu'une interface Ethernet, ils ne peuvent alors être connectés aux MRs qu'à travers cette interface.

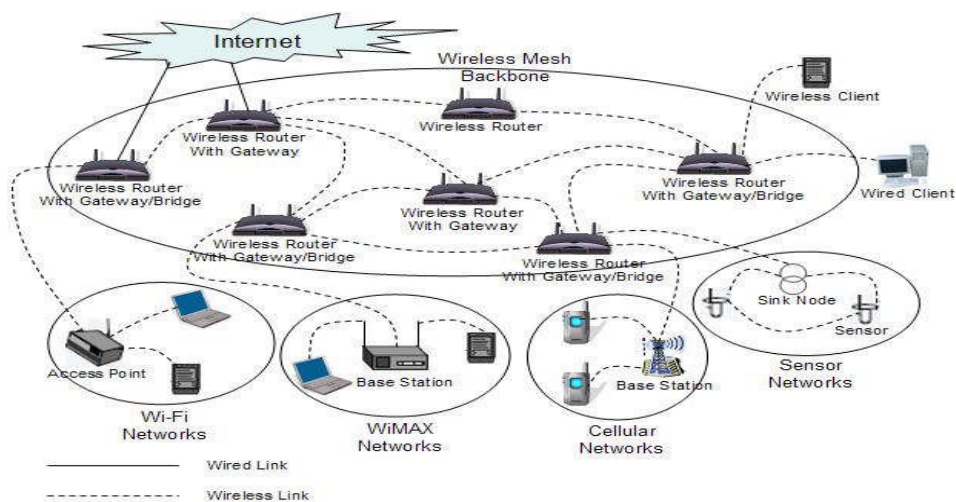


Fig. 2. "Backbone" d'un réseau maillé [Ch06]

- L'architecture hybride est composée des deux architectures précédentes (cf. Fig. 3). Les nœuds clients communiquent entre eux via les MRs ou peer to peer. Le "backbone" de l'infrastructure WMN permet la connectivité aux différents réseaux sans fil existants. Cette dernière architecture sera le modèle approprié pour la génération future.

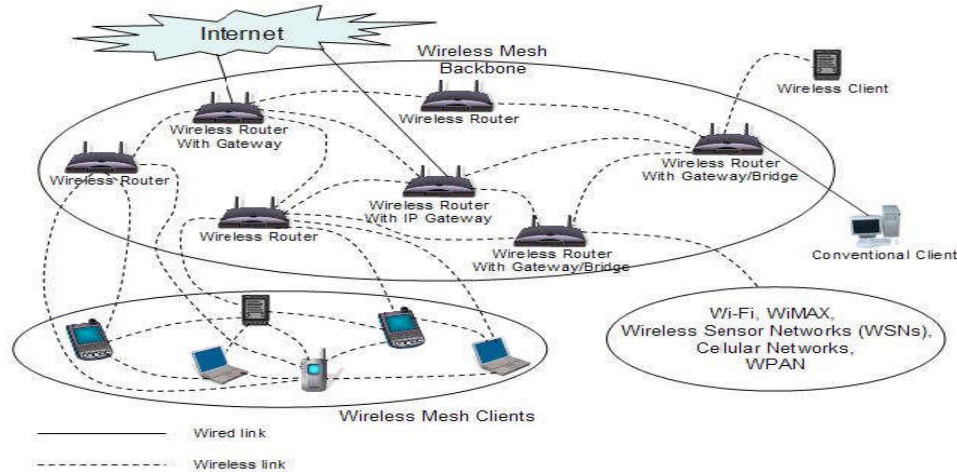


Fig. 3. Architecture hybride [Ch06]

1.1.2 Applications réelles déployant WMNs

L'application "Broadband Home Networking"² (cf. Fig. 4) est implémentée présentement à travers le protocole IEEE 802.11 [Ch06]. Les difficultés rencontrées en réalisant cette application sont les zones non couvertes et l'installation coûteuse des AP (Access point) multiples. En plus, deux nœuds de différents AP ne peuvent pas communiquer directement entre eux. Grâce aux WMNs, ces problèmes peuvent être résolus en remplaçant les AP par des MRs et en utilisant plus de ces MRs dans les zones non couvertes. Le déploiement de WMN dans une zone métropolitaine a plusieurs avantages. En premier lieu, le taux de transmission dans la couche physique est plus élevé que pour n'importe quel autre réseau cellulaire actuel. En second lieu, l'utilisation des WMNs est moins coûteuse pour les réseaux métropolitains sans fil à large bande d'accès si on la compare aux réseaux câblés ou optiques. En dernier lieu, à travers NLOS et les multiples-sauts parmi les nœuds, une zone de service plus large est offerte.

Nous constatons de même que les réseaux WMNs sont parfaits pour les applications de contrôle des immeubles, car celles-ci sont composées

² Application dans laquelle, un ensemble d'équipements communiquent à travers le réseau maillé tels que par exemple : TV, projecteur, caméra digitale, imprimante, lecteur MP3 etc.

d'équipements électroniques comme lumière, ascenseur, climatisation, etc. Il y a aussi d'autres applications appropriées à l'utilisation des WMNs, comme les systèmes de transport, le système médical, l'accès public à Internet, etc.

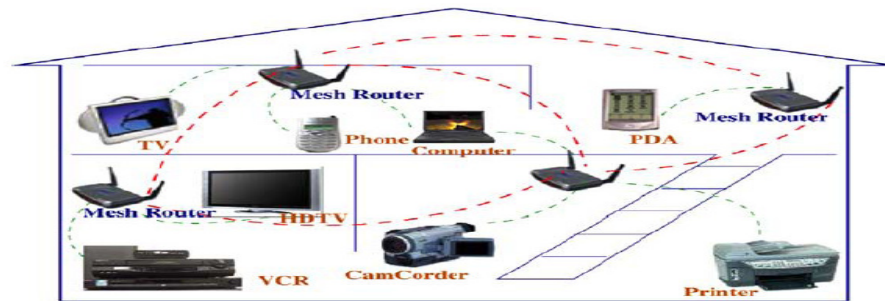


Fig. 4. "Broadband Home Network" [AW05]

1.2 Motivations et objectifs

Bien que la recherche dans le domaine des réseaux WMNs soit récente, elle est vaste et diversifiée. En effet, des travaux de recherche conséquents sont actuellement en train d'être effectués par des chercheurs dans le but de normaliser un ou plusieurs protocoles de routage pour les réseaux WMNs [AW05]. En outre, de nombreux autres travaux portant sur des protocoles de plus haut niveau (qualité de service, découverte de services, etc.) ont aussi été réalisés. Ces recherches s'appuient sur des simulations utilisant en particulier la norme 802.11 [EG08] de l'IEEE ou son extension directe 802.11e [NT04] [HM07], qui assure la différenciation du trafic selon des classes de priorité. De même, les chercheurs, en 2007-2008, ont adopté le protocole IEEE 802.11s [HM07] [HM08], qui est un protocole dédié aux WMNs et qui incorpore une méthode d'accès déterministe optionnelle appelée MDA, conçue pour les WMNs. Garantir la QoS pour les applications multimédia temps-réel demeure un défi dans les réseaux sans fil. Étant donné que les WMNs sont considérés parmi les solutions les plus attrayantes pour les problèmes des réseaux sans fil de nouvelle génération, ils

doivent être capables d'effectuer le routage de ce genre d'applications avec la QoS exigée.

Ainsi, l'objectif de ce travail de thèse est de contribuer à la définition d'une solution pour le support de la QoS dans le réseau fédérateur maillé sans fil, ainsi qu'à l'analyse des performances de cette solution. Le travail se situe plus précisément au niveau de la sous-couche MAC (Medium Access Control) et de la couche réseau. D'une part, il s'agit de spécifier, d'évaluer, puis d'implanter des mécanismes de gestion de ressources (ordonnancement, contrôle d'admission, etc.) afin d'offrir un service différencié avec certaines garanties de QoS. D'autre part, d'adapter le protocole de routage le plus approprié aux réseaux maillés sans fil afin de prendre en considération les paramètres de QoS, tel que le délai. En effet, les résultats de nombreux travaux de recherche sur le routage [PB03], [PB94], [CJ03] dans les réseaux MANET sont à reconsidérer [YJ08] puisqu'ils ont été conçus avec un souci d'optimiser la consommation d'énergie et sous l'hypothèse d'une forte mobilité des nœuds. Cependant, ces aspects ne sont plus valides dans le cadre des réseaux maillés sans fil. Par conséquent, beaucoup de travail reste à faire pour ce type de réseaux.

Cette thèse aborde le problème de garantir une QoS pour les applications avec des contraintes temporelles, telles que la voix et la visioconférence. Pour satisfaire ce type d'applications, nous nous intéressons principalement, dans le premier volet de la thèse, au contrôle d'admission des flux entrants en leur garantissant un court délai et un faible taux de perte de paquets. Sachant que les utilisateurs mobiles sont de plus en plus nombreux et fréquents dans les WMNs et que ces utilisateurs aiment pouvoir se promener sans avoir à se soucier de la possibilité d'interruption de leurs communications, nous traitons la différenciation entre plusieurs classes de service en présence du trafic en relève dans le deuxième volet de la thèse.

Le dernier volet de cette thèse s'attaque au problème de réaffectation des canaux en présence des appels en relève afin de minimiser le surcoût et le délai de re-routage et d'améliorer l'équité entre utilisateurs. Ce problème d'optimisation tient en compte maints aspects et critères, souvent contradictoires, bien qu'ils soient étroitement liés entre eux.

1.3 Contributions et organisation de la thèse

Les sujets de recherche évoqués ci-dessus ont abouti à plusieurs résultats, que nous avons soumis ou publiés sous forme d'articles dans des revues et des conférences internationales avec arbitrage. La thèse est constituée de trois articles de revues scientifiques que nous avons jugé les plus significatifs et complets parmi les onze que nous avons produits. Neuf de ces articles ont été acceptés (dont deux articles de revues et sept articles de conférences) et deux articles revues soumis. Les articles sélectionnés sont numérotés en gras à la fin de cette section. Trois chapitres, dans cette thèse, correspondent ainsi à trois articles de revues.

Dans le chapitre 2, après avoir introduit les WMNs et leurs caractéristiques, nous définissons en premier lieu des notions dont nous aurons besoin pour une meilleure compréhension des chapitres qui vont suivre. En second lieu, nous survolons les solutions existant déjà dans la littérature. En effet, une analyse critique de ces solutions nous permet d'identifier leurs limites et de mettre en place de nouvelles idées et propositions.

Le chapitre 3 propose un modèle analytique qui (1) calcule un ratio approprié d'admission du trafic afin d'assurer la QoS exigée par les flux entrants en termes d'un faible taux de perte de paquets; (2) calcule le délai de bout en bout pour les flux entrants avec une contrainte prédéfinie pour garantir un délai court et satisfaisant pour les applications avec des contraintes temporelles; (3) prend en compte les interférences et les ressources locales et celles dans un voisinage avant

de prendre des décisions relatives à l'acceptation et au rejet d'un trafic entrant. Ainsi, le contrôleur d'admission (RCAC) que nous avons proposé est capable de fournir la QoS exigée. Nous prouvons de même que la méthode d'accès CSMA/CA couplée avec RCAC donne de meilleurs résultats que ceux de la méthode d'accès MDA en termes de débit, de surcoût, de délai et de taux de blocage ou de perte de paquets.

Dans le chapitre 4, nous introduisons d'abord la différenciation entre classes de service multiples en présence du trafic en relève pour améliorer la QoS, notamment des applications multimédia temps-réel et/ou des utilisateurs mobiles. Nous démontrons ensuite que la méthode d'accès MDA renforcée par des mécanismes de contrôle d'admission et de gestion des tranches de temps, selon la classe de service en présence du trafic en relève, est meilleure que les méthodes d'accès DCF et EDCA en termes de longueur du délai, du taux de perte de paquets et surtout du taux de suppression du trafic en relève.

Dans le chapitre 5, nous proposons un modèle d'optimisation qui réaffecte les canaux de façon prudente en prenant en compte le trafic en relève. Ce modèle considère différentes fonctions objectif, tels que le débit, l'équité, le surcoût et la variation de l'équilibrage de la charge. Le solveur CPLEX [CP] est utilisé pour résoudre le modèle, et l'outil de simulation ns-2 [NS] pour la mesure de performance. En effet, ce modèle prend en considération la coexistence de ces nombreux critères pour retrouver une bonne solution sans pour autant favoriser l'un des critères au détriment des autres; une solution simultanément optimale pour les quatre critères étant impossible à garantir. En d'autres termes, une bonne solution est celle qui garantit la QoS exigée par les applications temporelles. Par exemple, dans une application voix, une solution qui donne un délai supérieur à 150 ms n'est pas considérée comme une bonne solution, même si elle satisfait les autres critères (débit, équité, l'équilibrage de la charge).

Le chapitre 6 résume les contributions majeures de cette thèse et expose les futures directions de recherche.

1.4 Articles publiés/soumis durant cette thèse

1. **J. Rezgui, A. Hafid, and M. Gendreau, "Distributed Admission Control in Wireless Mesh Networks: Models, Algorithms, and Evaluation", in *IEEE Journal Transactions on Vehicular Technology*, Vol.59, No.3, pp.1459-1473, 2010.**
2. **J. Rezgui, A. Belbekkouche and A. Hafid, "On Delay performance and Burst Assembly for Wireless Mesh and Optical Burst Switching Converged Metro Area Network", in *ACM/Springer Mobile Networks and Applications*, DOI: 10.1007/s11036-010-0249-3, 2010.**
3. **J. Rezgui, A. Hafid and M. Gendreau, "Admission Control and QoS Provisioning for IEEE 802.11s-based Mesh deterministic Access", Submitted to *Journal of Computer Networks*, 2010.**
4. **J. Rezgui, A. Hafid, R. Ben ali and M. Gendreau, "Optimization Model for Handoff-Aware Channel Assignment Problem for Multi-Radio Wireless Mesh Networks", Submitted to *IEEE Journal Transactions on Vehicular Technology*, 2010.**
5. **J. Rezgui, A. Hafid, R. Ben Ali and M. Gendreau, "Handoff-Aware Channel Assignment for Multi-Radio Wireless Mesh Networks", in *Proc. of IEEE ICC*, pp.1-6, 2010.**
6. **J. Rezgui, A. Hafid, R. Ben Ali and M. Gendreau, "Meta-heuristics for Channel (Re-) Assignment Problem for Multi-Radio Wireless Mesh Networks", in *Proc. of INFORMS Telecom*, 2010**

7. J. Rezgui, A. Hafid, R. Ben Ali and M. Gendreau, "Variable neighborhood search method for channel Assignment Problem in Multi-Radio Wireless Mesh Networks", to appear in *Proc of IEEE LCN 2010*, 2010.
8. J. Rezgui, A. Hafid and M. Gendreau, "Admission Control and QoS Provisioning in Multi-Service MDA in IEEE 802.11s-based Wireless Mesh Networks", to appear in *Proc of ISCT Qshine Conference*, 2010.
9. J. Rezgui, A. Hafid, M.Gendreau and B.Rong, "Congestion-aware Clique-based Handoff in Wireless Mesh Networks", in *Proc of IEEE MSN*, pp.388-395, 2009.
10. J. Rezgui, A. Belbekkouche and A. Hafid, "Novel Analytical Delay Model and Burst Assembly Scheme for Wireless Mesh and Optical Burst Switching Convergence", in *Proc of ICST QShine*, Las Palmas, Spain, pp.104-119, 2009.
11. J. Rezgui, A. Hafid and M. Gendreau, "A Distributed Admission Control Scheme for Wireless Mesh Networks", in *Proc of IEEE Broadnets*, London, UK, pp.594-601, 2008.

Chapitre 2: Revue de la Littérature

2.1 Introduction

Dans ce chapitre, nous présentons les principaux travaux de recherche traitant de la QoS dans le contexte des réseaux ad hoc et WMNs. Dans ce cadre, de nombreux travaux traitant de l'augmentation de la capacité ([KV05], [KV06], [KN05]) du réseau ont été développés. Néanmoins, cette augmentation ne signifie malheureusement pas, dans la plupart des cas, une garantie de la QoS. Ceci est expliqué par le fait que ces travaux n'ont pas introduit dans leurs modèles un mécanisme de contrôle d'admission pour bénéficier de cette augmentation. L'absence d'un mécanisme qui contrôle les flux entrants et qui s'assure d'avoir suffisamment de largeur de bande disponible avant d'accepter les flux, risque probablement de congestionner le réseau et de dégrader la QoS des trafics existants en termes de long délai et d'augmentation de perte de paquets.

Les différents mécanismes d'accès pour les réseaux sans fil seront présentés. Nous mettons l'accent sur leurs avantages et leurs inconvénients. D'abord, nous présentons la description de la méthode d'accès en mode distribué (protocole Distributed Coordination Function-DCF), suivie d'une introduction à la problématique liée à la QoS. Nous exposons ensuite la technique d'accès distribuée au média sans fil pour la norme IEEE 802.11e, à savoir EDCA (Enhanced Distribution Channel Access) tout en évoquant ses apports en termes de QoS. Puis, nous présentons le mécanisme d'accès déterministe (Mesh Deterministic Access-MDA) en détail; MDA étend DCF pour qu'il adapte une procédure de réservation déterministe en se basant sur des périodes de temps appelées MDA opportunités (MDAOPs).

Par la suite, un état de l'art concernant le mécanisme de contrôle d'admission sera détaillé pour les protocoles IEEE 802.11 et IEEE 802.11e respectivement. Les protocoles de routage les plus connus ainsi que le regroupement passif feront l'objet de la prochaine section. Même si le routage ne constitue pas le fondement de notre contribution, la connaissance des protocoles de routage du monde ad hoc est nécessaire

au travail que nous allons mener, en particulier lors de la proposition du protocole de routage distribué RCAC du chapitre 3.

Dans une étape suivante, nous étudions la nécessité des services différenciés pour maintenir une certaine QoS aux nouvelles applications multimédia avec des contraintes temporelles. Par la suite, sera présentée une revue de littérature mettant en perspective les solutions qui proposent de nouvelles techniques et heuristiques pour l'affectation des canaux dans les MR-WMNs. Finalement, une synthèse s'impose pour positionner nos contributions par rapport aux travaux de recherches proposés dans la littérature.

2.2 Les mécanismes d'accès au média sans fil

Nous présentons les trois mécanismes d'accès de base auxquels nous avons eu recours comme plateforme à nos algorithmes. Nous mettons l'accent sur leurs atouts mais aussi sur leurs limites pour faire comprendre l'importance des ajouts que nous avons proposés dans l'objectif d'adapter ces mécanismes aux contraintes de QoS requises par les applications multimédia.

2.2.1 Mécanisme d'accès distribué DCF pour la norme IEEE 802.11

Le standard IEEE 802.11 retient deux méthodes d'accès: la première, DCF, est une méthode distribuée, elle fait intervenir la totalité des stations du réseau; la deuxième, PCF (Point Coordination Function), est basée sur une approche centralisée dans laquelle l'accès au média se contrôle par le point d'accès. Dans ce qui suit, nous détaillons le fonctionnement du mécanisme DCF en mettant l'accent sur son inaptitude de supporter des trafics ayant des exigences strictes en matière de débit, de temps de transfert et surtout de récurrence temporelle (gigue).

2.2.1.1 Fonctionnement général de DCF

DCF est un algorithme distribué. Il est exécuté localement au niveau de chaque station afin de déterminer les périodes d'accès au média. La méthode DCF comprend quatre mécanismes, à savoir: 1) CSMA/CA (Carrier Sense Multiple Access with

Collision Avoidance); 2) algorithme "Backoff"; 3) principe de réservation du média; et 4) détection virtuelle des collisions.

L'accès au média est régi par le mécanisme CSMA/CA. La station émettrice exécute son algorithme de "Backoff" une fois que le support est libre. Après l'exécution de l'algorithme, la station débute son émission par un échange de paquets de réservation si le média est toujours libre. Au cours de son émission, la station gère un mécanisme de détection virtuelle des collisions.

La transmission des paquets par une station n'est possible que lorsque le média est libre pendant une certaine durée, DIFS (cf. Fig. 5). Dans le cas contraire, une procédure de "Backoff" est enclenchée.

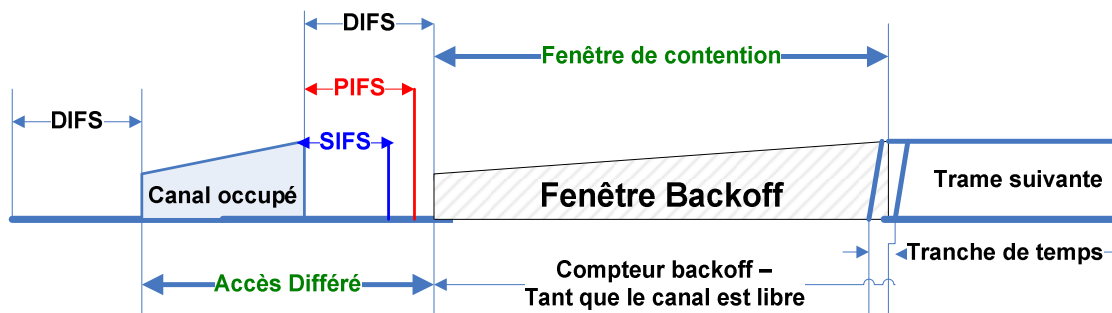


Fig. 5. Accès au média sans fil en mode DCF (IEEE 802.11)

Le mécanisme DCF utilise plusieurs temps d'attente pour la transmission de différents types de trames, appelés espaces d'inter-trames. Ces derniers sont déduits à partir d'IFS (Inter-Frame Space). IFS est la durée pendant laquelle une station doit attendre avant de transmettre sur le média. Nous comptons quatre sortes d'IFS: 1) SIFS (Short IFS): utilisé pour séparer les transmissions d'un même *dialogue*, par exemple entre une trame émise et son acquittement ou encore entre plusieurs fragments d'une même trame. Étant donné que le SIFS est le plus court des espaces inter trames, si une station commence à émettre on ne peut l'interrompre; 2) PIFS (Priority Inter Frame Space): utilisé par l'AP pour accéder prioritairement au média ($\text{PIFS} < \text{DIFS}$); 3) DIFS (Distributed IFS): utilisé par une station pour accéder au support de transmission, (DIFS est plus long que SIFS pour garantir une priorité absolue aux acquittements); 4)

EIFS (Extended IFS) est l'IFS le plus long: utilisé si une station reçoit un paquet erroné. Ainsi, EIFS permet de prévenir une collision avec le prochain paquet à venir.

La norme introduit la notion d'une tranche de temps ("time-slot") comme étant l'intervalle de temps qui permet à une station de savoir si une autre station a accédé au média au début du slot précédent (cf. Fig. 5). La valeur d'une tranche de temps dépend de la couche physique utilisée.

2.2.1.2 Limites de DCF en termes de QoS

Le mode DCF supporte les services de type "Best Effort" uniquement, sans aucune garantie de QoS [EG08]. En effet, les différents types de trafics dans DCF ont la même probabilité d'accès au média. Aucun mécanisme de différenciation n'est implémenté afin de garantir le débit, le délai de transfert ni la gigue pour les applications avec des contraintes temporelles et de hautes priorités. Certains travaux [XX06] se sont intéressés à l'équité de ce protocole d'accès au média en plus de s'intéresser à l'évaluation de ses performances. Pour une configuration simple, tel qu'un réseau à un saut, le protocole permet à chaque émetteur d'accéder au média avec la même probabilité. Il en résulte que ce protocole fournit une certaine équité en termes de nombre de trames émises. Cependant, il n'y a aucune garantie sur le volume de données effectivement transmis. En effet, la taille des trames n'est pas prise en compte dans la probabilité d'accès au média. Par ailleurs, des différences dans la topologie, telles que l'existence de zones plus denses que d'autres ou encore la présence d'interférences entre différents trafics peuvent avoir un impact important sur la probabilité de transmission avec succès des différents émetteurs dans une configuration à multiples-sauts tel que le cas des WMNs. Une différenciation de services s'avère impérative pour une meilleure performance réseau dans les WMNs surtout pour les applications multimédia temps-réel.

2.2.2 Mécanisme d'accès EDCA pour la norme IEEE 802.11e

La principale préoccupation du groupe de recherche, dans le cas de la spécification IEEE 802.11e, est de répondre aux exigences de QoS sans pour autant

sacrifier les intérêts des acteurs industriels concernés. Le mécanisme d'accès distribué, à savoir EDCA [NT04] permet une différenciation de services établie au niveau de la couche MAC. Dans le protocole IEEE 802.11 (DCF), tant que les requêtes sont courtes, chacune occupe le réseau peu de temps, et que les délais d'attente restent faibles, le problème ne se pose pas. En revanche, tout se complique énormément lors du transfert de gros fichiers comme de la vidéo ou de la voix. Pour remédier à ces lacunes, une nouvelle norme 802.11 intégrant la QoS, l'IEEE 802.11e (EDCA), a été proposée.

2.2.2.1 Le fonctionnement général de EDCA

La norme IEEE 802.11e vise à donner des possibilités en matière de QoS au niveau de la couche liaison de données. Elle définit aussi les besoins des différents paquets en termes de bande passante et de délai de manière à permettre une meilleure transmission de la voix et de la vidéo. Le standard IEEE 802.11e ajoutera ainsi des extensions afin d'améliorer la QoS pour les applications ayant des exigences qualitatives spécifiques [CC04] et [Ne04] tout en préservant la compatibilité ascendante avec les variantes des réseaux sans fil existantes.

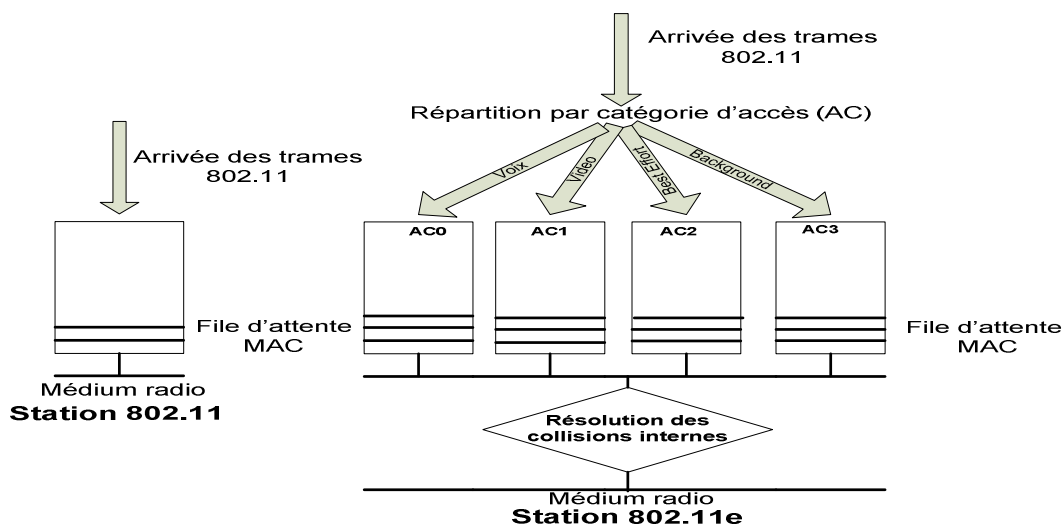


Fig. 6. Accès au média sans fil en mode EDCA (IEEE 802.11e)

L'EDCA est une amélioration du mode de communication traditionnel DCF de la norme IEEE 802.11. Ce protocole introduit un nouveau concept, celui de catégorie d'accès ou AC pour "Access Category" (cf., Fig. 6). Les catégories d'accès sont :

"Background", "Best Effort", la vidéo et la voix. EDCA fournit ainsi un accès différencié et distribué au média. Ce protocole attribue chaque trafic à une catégorie d'accès contenant des valeurs bien définies pour les paramètres d'accès DCF. L'accès au média pour une station dépend de la catégorie d'accès associée aux flux à transmettre.

2.2.2.2 Limites de EDCA en termes de QoS

Certes, le mécanisme de différenciation EDCA permet d'offrir des possibilités en matière de QoS. Les modifications introduites au niveau de la couche MAC assurent un traitement spécifique pour chaque type du trafic. Les recherches et les simulations [Ne04] réalisées prouvent que cette différenciation garantit une meilleure transmission de la voix et de la vidéo. Mais, certains problèmes persistent, telles que la dégradation des trafics à faible priorité et l'absence de différenciation entre un nouvel appel et un appel en relève ("handoff") [RH10e].

2.2.3 Mécanisme d'accès optionnel MDA pour la norme IEEE 802.11s

IEEE 802.11s est un protocole proposé pour modifier et améliorer IEEE 802.11. Ce protocole-là est conçu spécialement pour les WMNs. Il peut être utilisé pour des topologies statiques et aussi pour les réseaux ad hoc. Il est à noter que le standard de base dans IEEE 802.11s est EDCA. Il a les mêmes caractéristiques de base que celui proposé pour IEEE 802.11e. Une procédure optionnelle dédiée aux WMNs, appelée Mesh Deterministic Access (MDA) est incorporée dans le protocole IEEE 802.11s. Certains travaux de recherche récents [CL08], [GM09] ont amélioré la méthode d'accès MDA surtout en termes de surcoût et d'interférences sans porter attention au contrôle d'admission du trafic ni à la différenciation de services.

2.2.3.1 Le fonctionnement général de MDA

MDA a été proposé pour augmenter l'efficacité du protocole IEEE 802.11s [HM07] et réduire les risques de collisions. Il s'agit d'une méthode d'accès qui permet aux MRs d'accéder au média à des moments choisis à faible taux de collision possible dans un voisinage à deux-sauts. Comme certains nœuds qui supportent la fonctionnalité

MDA peuvent potentiellement interférer les uns avec les autres, MDA met en place des périodes de temps appelé MDAOPs dans le voisinage d'interférences. Ces MDAOPs sont d'abord négociés entre les MRs voisins via l'échange de messages de diffusion. Ensuite, les réservations MDAOPs sont effectuées dans les multiples unités de tranches de temps, durant l'intervalle périodique DTIM (cf., Fig. 7). Quand une réservation a été complétée avec succès, le propriétaire de MDAOP peut accéder aux tranches de temps appropriées en utilisant la méthode d'accès EDCA ou celle DCF. Toutefois, EDCA n'établit pas une politique de priorité stricte, mais seulement une priorité relative à certains nœuds par rapport aux autres. Il est à noter que pour utiliser la méthode d'accès MDA, un nœud doit être synchronisé.

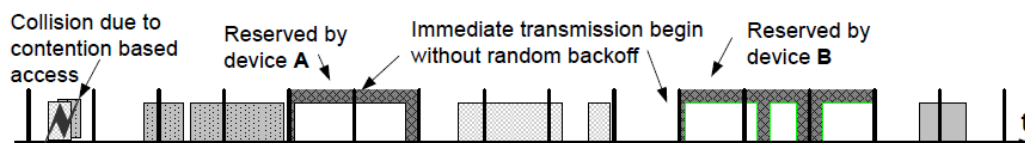


Fig. 7. MDAOPs durant l'intervalle DTIM (802.11s) [So07]

2.2.3.2 Limites des solutions basées sur MDA en termes de QoS

Les interférences en dehors du voisinage à deux-sauts peuvent encore se produire avec MDA, ce qui peut dégrader les performances des WMNs. Les contributions dans [CL08], [GM09] apportent une solution au problème d'interférences à l'extérieur de deux-sauts rencontré par MDA.

Dans [CL08], les auteurs s'attaquent à ce problème à l'aide d'une réallocation dynamique (appelée, MDA-DR) des tranches de temps MDA en conflit et à deux-sauts d'une demande d'un nœud. Dans la solution proposée, si une interférence est détectée à l'extérieur de deux-sauts, les flux sont déplacés de façon dynamique, en fonction de leurs durées écoulées dans le réseau. Toutefois, MDA-DR limite son amélioration de la performance au débit ce qui le rend approprié uniquement aux applications de données, tels que le transfert de fichiers et la navigation sur le Web, mais non aux applications avec des contraintes temporelles, par exemple la voix. En outre, il n'intègre pas un mécanisme de contrôle d'admission, qui lui permettra de limiter les pertes de paquets.

Dans [GM09], un accès au média avec réservation ("Medium Access Reservation"), dit MARE, est proposé comme une généralisation de la procédure de RTS/CTS dans un réseau sans fil à multiples-sauts. Plutôt que d'utiliser l'émission excessive des "beacons" de diffusion à chaque nœud MDA, MARE utilise une procédure RTS/CTS plus élaborée, qui est en mesure de notifier le groupe approprié des nœuds voisins, pour lequel un ensemble de tranches de temps a été réservé. MARE réduit sensiblement le surcoût par rapport à celui généré par MDA. Étant donné que l'approche MARE nécessite une taille constante de données, elle est plus appropriée pour le trafic voix. Toutefois, MARE n'intègre ni un mécanisme de contrôle d'admission ni celui d'un ordonnancement, ce qui limite son efficacité et ses garanties de QoS pour le trafic voix.

2.2.4 Récapitulatif et comparaison des mécanismes d'accès

Table 1. Comparaison des méthodes d'accès proposées dans la littérature

	QoS du trafic "Best Effort" : (1)	QoS du trafic voix : (2)	QoS du trafic en relève : (3)	QoS prenant en compte (1-3)
DCF [EG08]	X			
EDCA [Ne04]		X		
MDA [HM07]		X		
MDA-DR [CL08]		X		
MARE [GM09]		X		
Notre approche (chapitre 4)	X	X	X	X

Dans le chapitre 3, nous prouvons que pour garantir une QoS aux trafics avec des contraintes temporelles, nous devons doter la méthode d'accès DCF d'un

mécanisme efficace d'admission de contrôle des flux entrants [RH08]. Car, comme nous l'avons déjà mentionné, DCF est uniquement adéquat pour le trafic de type "Best Effort" [EG08]. Son amélioration EDCA permet la différenciation du trafic et la possibilité d'attribuer une priorité plus élevée aux trafics avec des contraintes temporelles, telle que la voix. Néanmoins, EDCA garde les inconvénients constatés avec DCF au niveau des interférences, que la méthode d'accès MDA contourne par un accès déterministe à des tranches de temps choisies. En outre, la méthode d'accès MDA ou son amélioration, telles qu'illustrées dans les travaux [CL08] et [GM09], ne prennent pas en charge la différenciation: (a) entre les différentes classes de service, et (b) entre les trafics en relève et les nouveaux trafics, ce qui va être traité dans le chapitre 4 [RH10e]. Une étude comparative de ces travaux de recherche est résumée dans la Table 1.

2.3 Qualité de service dans les réseaux maillés sans fil

Dans les réseaux de télécommunications, l'objectif de la QoS est d'atteindre un meilleur comportement de la communication, pour que le contenu de celle-ci soit correctement acheminé, et que les ressources du réseau soient utilisées d'une façon efficace. La QoS peut être définie comme le degré de satisfaction d'un utilisateur des services fournis par un système de communication. La QoS est définie dans [Qo99] comme la capacité d'un élément de réseau (ex : routeur, nœud ou une application) de fournir un niveau de QoS lors de l'acheminement des données.

Le RFC 2386 [CN98] caractérise la QoS comme un ensemble de besoins à assurer par le réseau pour le transport d'un trafic d'une source à une destination. Ces besoins peuvent être traduits en un ensemble d'attributs pré-spécifiés et mesurables en termes de :

- délai de bout en bout,
- variance de délai,
- bande passante,
- pertes de paquets.

Les besoins de la QoS sont différents selon le type de l'application. Par exemple, pour les applications temps-réel, comme la voix et la vidéo, le délai de bout en bout d'un paquet doit être limité, autrement le paquet est inutile. Les applications qui ne sont pas en temps-réel, comme le transfert de fichier ou la messagerie, se focalisent quant à elles sur la fiabilité des communications.

Le support de la QoS a été largement étudié dans les réseaux filaires. Le réseau ATM (Asynchronous Transfer Mode) supporte la QoS du trafic tout en le différenciant en plusieurs classes. Dans ce contexte, des solutions ont été proposées par l'IETF pour améliorer le réseau Internet afin de fournir la QoS aux communications multimédia. De même, des mécanismes ont été fournis pour gérer efficacement les ressources réseaux (bande passante, mémoires tampons) et répondre aux exigences des applications notamment celles avec des contraintes temporelles.

Cependant, il est très difficile de garantir une QoS prédéterminée à une application temps-réel dans un WMN, car il faut prendre en considération les spécificités de ces réseaux, à savoir: la bande passante limitée, le changement dynamique de la topologie dans le temps ainsi que le manque d'information complète sur l'état du réseau. En outre, la communication entre les stations mobiles, étant effectuée par voie radio, la qualité du lien sans fil reste peu fiable, et sujette à des variations selon la configuration et l'état du réseau.

2.3.1 Modèles de qualité de service pour les réseaux ad hoc

Dans cette sous section, nous décrivons deux modèles de QoS FQMM [XS00] et SWAN [GC02] proposés pour les réseaux ad hoc. Ces deux modèles intègrent le mécanisme de contrôle d'admission qui, selon nous, est à l'origine de la bonne performance de ces deux modèles.

2.3.1.1 FQMM (Flexible Quality of service Model for MANETs)

Le modèle FQMM [XS00] se base sur une architecture réseau non hiérarchique, composée d'une cinquantaine de nœuds mobiles, formant un domaine DiffServ (Differentiated Services) [PB96]. Il combine les propriétés des modèles filaires IntServ (Integrated Services) et DiffServ, en offrant une méthode d'approvisionnement hybride:

par flux, pour les trafics prioritaires, et par classe, pour les autres trafics. Dans un réseau, les nœuds peuvent avoir des rôles différents suivant les trafics existants. Le FQMM repose essentiellement sur la couche IP, où les fonctionnalités sont séparées en deux grands plans: le plan d'acheminement de données et le plan de contrôle et gestion (cf. Fig. 8). Dans ce modèle, le protocole de routage est supposé fournir des routes ayant suffisamment de ressources. Vu que les mécanismes de QoS offerts sont proches des protocoles filaires, il est possible d'interfacer le réseau avec l'Internet. Cependant, plusieurs mécanismes ainsi que l'interaction avec la couche MAC restent à définir pour s'adapter aux conditions variables du réseau WMN. Toutefois, nous préférons proposer un nouveau modèle (cf. Fig. 14) du support de la QoS qui se situera au niveau de la sous-couche MAC et de la couche réseau et qui est réellement adapté aux réseaux WMNs.

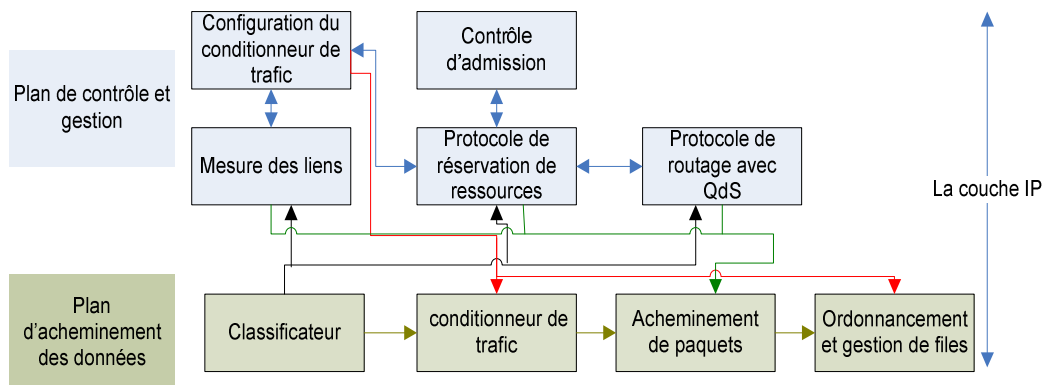


Fig. 8. Le modèle FQMM

2.3.1.2 SWAN (Service differentiation in wireless ad hoc networks)

SWAN [GC02] est un modèle basé sur un algorithme de contrôle distribué dans le but d'assurer une différenciation de services pour le protocole IEEE 802.11. Il offre la priorité (au niveau paquet) aux trafics temps-réel en contrôlant la quantité du trafic "Best Effort" admise par un nœud. Pour accepter un nouveau trafic temps-réel, le contrôle d'admission sonde la bande passante minimale disponible sur la route (validée et obtenue par un protocole de routage). Une décision à la source est alors prise suivant la bande passante obtenue. Pour maintenir la QoS des trafics déjà acceptés, le débit des trafics "Best Effort" est régulé en utilisant les mesures de délais au niveau MAC

comme paramètre. Un classificateur et un "shaper" permettent de différencier les deux types du trafic (cf. Fig. 9). En cas de congestion, les bits ECN (Explicit Congestion Notification) de l'entête des paquets IP sont positionnés pour permettre à la source de réinitialier le contrôle d'admission. Si la route ne dispose pas d'assez de bande passante, le trafic est supprimé. Ainsi, SWAN permet de fournir une QoS logiciel ("software QoS"). Toutefois, un flux prioritaire admis n'est pas sûr d'avoir des garanties pour la durée de la communication et peut à tout moment être violé par d'autres demandes de trafics. En plus, un mécanisme de contrôle de débit des flux "Best Effort" n'est pas à lui seul suffisant pour offrir des garanties aux applications temps-réel. De même, le protocole de routage ainsi que la couche d'accès au média sont de type "Best Effort". Enfin, ce modèle est conçu sans considération des nœuds au voisinage, i.e. les nœuds qui sont dans l'intervalle d'interférence.

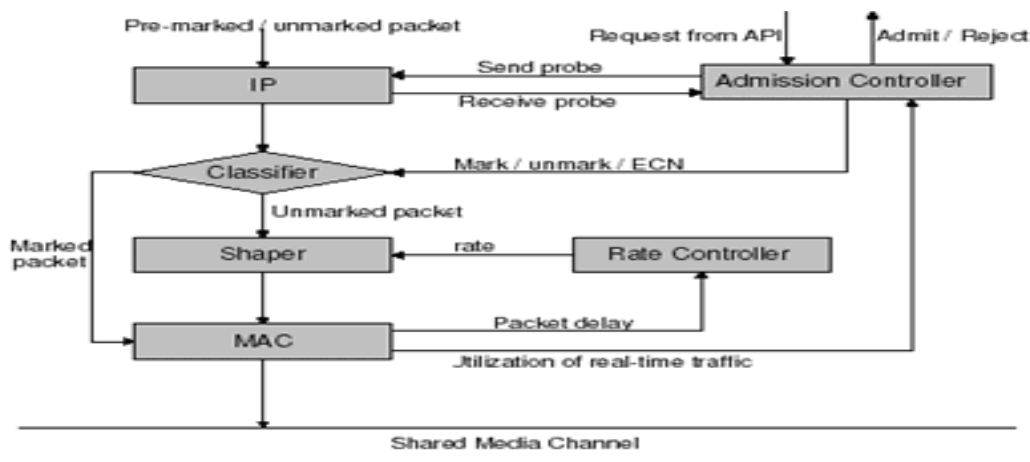


Fig. 9. Le modèle SWAN [GC02]

En fait, ces deux modèles sont de bons exemples à suivre pour supporter la QoS dans les réseaux ad hoc, mais ne sont pas adaptés aux réseaux WMNs et sont proposés pour des réseaux à un seul canal. La comparaison de notre approche (cf. Fig. 14) avec ces deux modèles et d'autres travaux est illustrée par la Table 2. Notons que dans notre approche, nous proposons plus de fonctionnalités, tels que: (a) un routage avec une QoS exigée (court délai, faible taux de perte de paquets) et un mécanisme de contrôle d'admission qui prend en compte les ressources locales et aux voisinages (voir chapitre

3), (b) la différenciation du trafic entre différentes classes de services et notamment entre un nouvel appel et un appel en relève (voir chapitre 4), et (c) l'affectation dynamique des canaux aux interfaces de façon prudente afin de minimiser le surcoût et le délai de re-routage et la gestion adaptative des ressources lors d'un trafic en relève ou de présence d'une congestion dans un réseau à MR-WMN (voir chapitre 5).

2.3.2 Dimensionnement des WMNs: augmenter la capacité

Augmenter la capacité est une exigence importante à satisfaire pour améliorer la QoS dans les réseaux sans fil. De nombreuses études [RC05] et [KM04] ont considéré le routage, l'affectation des canaux et l'ordonnancement. Certains chercheurs ont étudié le problème de sélection des routes efficaces pour maximiser le débit [KV05], [KV06], pendant que d'autres ont considéré l'affectation d'un seul canal et l'ordonnancement [KM04]. Tous ces travaux ont traité une partie du problème, ils se sont intéressés à la question de l'amélioration du débit en se contentant de se comparer à d'autres algorithmes, sans incorporer un mécanisme de contrôle d'admission des flux en cas de congestion. Par contre, pour certains [KN05], la question était de savoir comment caractériser la capacité par région dans le réseau en se basant sur des méthodes d'optimisation. Ils ont développé des algorithmes pour optimiser les protocoles de routage, l'affectation des canaux et l'ordonnancement dans le but d'obtenir la capacité minimale et maximale par région en considérant une fonction objectif. Dans [DA05], les auteurs ont essayé d'optimiser l'affectation des canaux afin d'obtenir le maximum de liens qui peuvent être actifs simultanément en tenant compte des contraintes d'interférence. Dans ce cadre, ils ont formulé le problème en programme linéaire pour résoudre l'affectation statique des canaux avec multiples-radios.

En équipant les MRs avec multiples-radios opérant sur des canaux non orthogonaux, la capacité du réseau augmente. Toutefois, l'affectation des canaux présente un défi à cause des interférences possibles qui peuvent affecter la performance du réseau. Dans [KV06], les auteurs ont proposé un algorithme d'affectation des canaux en considérant le problème d'interférence. Leur solution permet d'affecter les canaux aux radios en minimisant l'interférence. Il s'agit d'utiliser une nouvelle technique implémentée à chaque MR pour estimer l'interférence en utilisant un graphe

de conflit. Ils ont montré que la capacité est améliorée de 40% par rapport à l'affectation statique des canaux. L'avantage de cette approche est que l'affectation proposée a un impact moins considérable sur la topologie que d'autres méthodes proposées dans le contexte d'augmentation de la capacité.

Bien que toutes ces nombreuses études aient essayé d'augmenter la capacité des réseaux WMNs à multiples-radios et multiples-canaux, en optimisant le routage ou l'ordonnancement, leurs apports demeurent insuffisants pour garantir la QoS. Nous proposons un modèle du support de la QoS dans WMN (cf. Fig. 14), qui récapitule l'ensemble de contributions des trois prochains chapitres, dont la capacité a été optimisée par l'une des approches existantes [DA05], [KV06]. Cependant, nous considérons l'augmentation de la capacité comme une exigence nécessaire non suffisante pour garantir la QoS dans les réseaux sans fil. Pour remédier à ce problème, il vaut mieux gérer les ressources réseau par des mécanismes de contrôle d'admission afin d'améliorer réellement la QoS. C'est ce que nous allons détailler dans ce qui suit (pour les réseaux sans fil) avant d'évoquer l'importance de prendre en compte les interférences lors de contrôle d'admission des flux entrants [RH08] afin d'augmenter la capacité réseau [KV06].

2.3.3 L'importance de considérer et de minimiser les interférences

Pour améliorer les performances des WMNs, les interférences doivent être prises en compte. En effet, elles sont le premier facteur qui dégrade les performances d'un réseau sans fil. L'étude approfondie par Jain et al. [JP03] traite l'impact des interférences sur un réseau sans fil à multiples-sauts. Outre les collisions, les auteurs ont constaté que la perte des trames est un autre résultat des interférences entre canaux. La perte des trames peut se produire en raison des interférences cumulées résultant de nœuds n'appartenant pas au domaine de transmission silencieux de l'émetteur. En effet, l'ensemble silencieux d'un émetteur X est l'ensemble des nœuds qui est capable de détecter le média occupé si cet émetteur X transmet [YH07].

Dans la littérature, plusieurs modèles d'interférences ont été proposés tels que [HM02], [Le86], et [ZB07]. Cependant, Xu et al. [XY08] suggèrent l'utilisation d'une technique nommée "channel-bonding" pour réaliser une transmission de paquets à haut débit des données en utilisant un média plus élargi. Les résultats de simulation ont montré que, lorsque le trafic est faible, la technique "channel-bonding" réalise un délai minimal par rapport à la technique qui utilise de multiples-canaux. Par contre, dans le cas où le trafic est élevé, l'utilisation de multiples-canaux peut fortement atténuer l'influence des collisions de paquets et par la suite améliorer les performances du réseau. Aussi, est-il essentiel de mentionner l'importance de considérer non seulement les interférences dans la zone de transmission mais aussi dans la zone Carrier Sensing Range: CSR qui sont détaillés comme suit:

1) Transmission range (r): Les nœuds ne sont liés (ne communiquent) entre eux que si la distance entre deux nœuds est inférieure ou égale à r et que ces nœuds utilisent le même canal (le cas de a et b dans Fig. 10). En d'autres termes, le nœud a va être lié à tous les nœuds qui sont exposés à ses émissions (i) et aux nœuds dont l'émission interfère à la réception de b (ii), (cf. Fig. 10). Tous les nœuds dans la zone grisée seront liés à a et b s'ils utilisent le même canal. Pour détecter les nœuds dans la zone de transmission r , nous adaptons le concept de cliques de type A (voir chapitre 3).

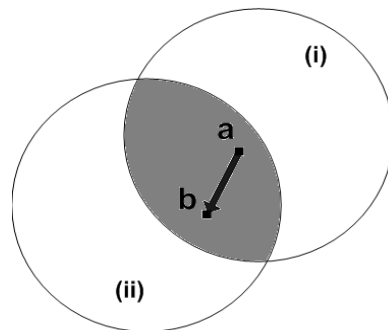


Fig. 10. Diffusion radio à une distance r

2) CSR [r, R]: Les nœuds ne sont liés (ne détectent pas du bruit) entre eux que si la distance entre deux nœuds est inférieure ou égale à R et supérieure strictement à r et qu'ils utilisent le même canal. Cette zone de transmission est utilisée pour bien estimer

la largeur de la bande passante au voisinage. Certains nœuds de la zone grisée, (cf. Fig. 11) n'ont pas de moyen de se rendre compte des situations de compétition avec un flux entrant sur le nœud a dans la zone blanche. Pour détecter les nœuds dans la zone CSR $]r, R]$, nous adaptons le concept de cliques de type B (voir chapitre 3).

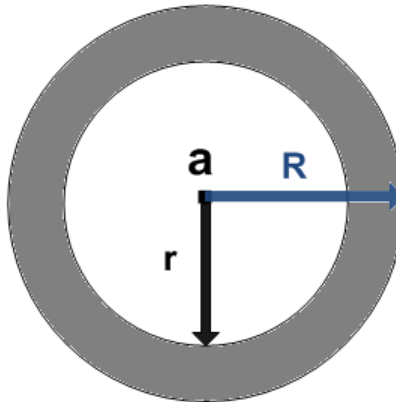


Fig. 11. Deux zones: de transmission (r) et de CSR ($]r, R]$)

L'efficacité d'un mécanisme de contrôle d'admission dépend de la façon dont la capacité du réseau est estimée. L'estimation est difficile à obtenir parce que, par rapport aux réseaux filaires, les liens dans WMNs sont intrinsèquement partagés, à cause des interférences. Ces liens sont difficiles à isoler, ce qui rend la performance des WMNs difficile à contrôler. Cependant, il est essentiel que le contrôle d'admission considère à la fois les ressources locales (zone de transmission 1) et les ressources au niveau des nœuds voisins (zone CSR 2) lors de l'analyse de la performance du réseau [YK05], [LG06]. La raison en est que les interférences 1) et 2) entre les liens sans fil causent une dégradation des performances, par exemple deux liens qui interfèrent et qui sont actifs en même temps donnent un débit plus faible que pour deux liens séparés. Dans le chapitre 3, nous adaptons le concept des cliques (rappelons, une clique est définie comme un sous-ensemble de liens logiques qui interfèrent les uns avec les autres) de la théorie des graphes pour résoudre ce problème 1) et 2), plus précisément pour identifier les nœuds qui sont en contention avec un émetteur par rapport à la zone de transmission et la zone CSR.

2.3.4 Contrôle d'admission

Peu nombreuses sont les études qui ont considéré dans leurs modèles de support de la QoS le mécanisme contrôle d'admission. Dans notre étude nous différencions entre les travaux de contrôle d'admission pour les différents protocoles 802.11 et 802.11e. Rappelons que le protocole 802.11 est un standard de l'IEEE qui définit une couche physique et une couche MAC pour les réseaux locaux sans fil. Il opère sur la bande de fréquence 2.4 GHz et permet des débits de 2 Mb/s. Le protocole 802.11e cherche à améliorer 802.11 de façon à pouvoir donner des garanties de QoS. Il supporte la différenciation de services pour les flux avec différentes priorités.

2.3.4.1 Le contrôle d'admission dans IEEE 802.11

Plusieurs travaux ont été faits pour garantir le support de la QoS dans les réseaux câblés ou sans fil. IntServ/RSVP et DiffServ [PB96] sont deux solutions pour procurer la QoS dans Internet. Cependant, elles ne sont pas adaptées pour les réseaux sans fil, car elles ne peuvent pas estimer les ressources disponibles ni calculer les ressources demandées par le flux entrant à cause du partage du canal dans les réseaux sans fil. Dong et al. [DY02] ont mentionné que le contrôle d'admission est une composante nécessaire et indispensable pour supporter la QoS, mais il n'y a pas un algorithme spécifique qui est désigné pour contrôler l'admission d'un flux entrant dans leurs travaux de recherches.

Dans SWAN [GC02], le contrôleur d'admission se met à l'écoute de toutes les transmissions des paquets avec leur portée de transmission pour réunir les informations sur la largeur de la bande passante et sur la congestion. Le contrôle d'admission est basé sur la largeur de la bande passante, mesurée tout au long du chemin de communication. La mesure de la bande passante s'effectue par l'envoi de probes messages. Cependant le sondage produit beaucoup de surcharge. De même, la perte d'un paquet probe engendre l'impossibilité de pouvoir déterminer la valeur courante de la largeur de bande passante. En plus, SWAN ne considère pas le cas où les nœuds peuvent s'interférer entre eux, même s'ils ne communiquent pas directement.

Dans AQOR [LC00] (An On-Demand QoS Routing Protocol), chaque nœud envoie périodiquement un paquet "Hello" qui annonce son existence et les informations sur le trafic concernant ses voisins. Basé sur cet échange, chaque nœud calcule la largeur de la bande passante disponible et consommée tout en tenant compte de son trafic et des trafics de ses voisins immédiats. Mais, AQOR comme SWAN ne considère que les ressources locales et non les ressources dans tous les nœuds en compétition (l'intervalle d'interférence CSR).

CACP [YK05] (Contention Aware Control Packet) est désormais le premier travail qui a invoqué la notion de la largeur de bande passante par rapport aux voisins, il estime les largeurs de la bande passante des nœuds occupées dans l'intervalle de l'interférence. Il faut avoir assez de largeur de la bande passante disponible localement et au voisinage pour une admission réussie d'un nouveau flux. Le CACP se base sur la source de routage pour construire l'ensemble des voisins en interférence et pour pouvoir calculer la largeur de bande passante consommée.

Lin et al. [LG06] présentent un mécanisme de contrôle d'admission dans les réseaux ad hoc. Leur proposition prend en considération les anomalies mentionnées dans CACP, telle que l'importance de considérer des nœuds voisins dans l'intervalle d'interférence. Leur mécanisme est intégré avec le protocole de routage ad hoc hop-by-hop, ce qui permet d'identifier des routes alternatives si le chemin le plus court est congestionné. Chaque nœud mesure la largeur de la bande passante disponible à travers le passif monitoring du canal sans fil. Leur proposition a permis d'améliorer l'estimation de la largeur de la bande passante en introduisant une formule qui considère une possibilité de réutilisation des transmissions parallèles. Lors des simulations, ils ont prouvé qu'en considérant les transmissions parallèles, nous pouvons augmenter l'admission du trafic en maintenant la QoS. Mais, ce travail considère un environnement avec un seul canal et une seule interface radio. Cependant, ce mécanisme, à son état actuel, ne permet pas une possibilité de l'étendre au contexte multiples-radios et multiples-canaux.

Les travaux [GC02], [LC00], [YK05] et [LG06] ont été proposés avec un souci de résoudre les problèmes de mobilité, d'autonomie en énergie et de connectivité. Cependant, les réseaux WMNs ne présentent pas ces problèmes. Aussi, tous ces travaux ont-ils été étudiés dans le cas des réseaux sans fil à un seul canal. De ce fait, nous aimerions développer ces idées en les adaptant aux MR-WMNs et améliorer ces propositions en proposant un mécanisme de contrôle d'admission efficace et un protocole de routage qui permettent le passage à l'échelle et qui supportent la QoS des applications multimédia temps-réel pour les MR-WMNs (voir chapitre 3).

2.3.4.2 Le contrôle d'admission dans IEEE 802.11e

Peu d'algorithmes de contrôle d'admission se sont focalisés sur les réseaux sans fil IEEE 802.11, encore moins sur le protocole de standard IEEE 802.11e. Le premier travail [GZ03] considère le contrôle d'admission au niveau de chaque nœud. Par contre, les autres travaux [KL03], [XL04], [LW08] implémentent l'algorithme de contrôle d'admission au niveau du nœud AP.

Dans [GZ03], la méthode proposée requiert que chaque station mesure la charge du trafic du canal sans fil et la compare à un seuil donné. Si la charge est supérieure à ce seuil, il y a une surcharge exprimée par un long délai et une possibilité de dégradation de débit. La station concernée arrête la transmission des flux avec priorité moins élevée pour assurer la transmission des flux avec priorité élevée et continue à garantir la QoS réclamée. Quand la charge du trafic devient inférieure à un seuil donné, la station reprend la transmission des flux arrêtés pour augmenter l'efficacité du réseau. Dépendamment de comment les mesures et les calculs de la charge du canal sans fil sont effectués. Deux méthodes ont été élaborées: l'une pour le calcul de la largeur de la bande passante occupée et l'autre pour calculer le taux de collisions. La station prend la décision, dans les deux cas, grâce à une règle de seuil simple.

Dans [KL03], les auteurs ont proposé une stratégie de contrôle d'admission pour garantir la QoS exigée par chaque classe du trafic. De cette façon, le débit et le délai sont estimés selon un modèle analytique pour le mécanisme EDCF, le standard de

protocole IEEE 802.11e, avec des paramètres mesurés pour déterminer si un nouveau trafic est accepté dans le système ou non.

Dans [XL04], QAP (Quality Access Point) mesure l'utilisation du média. Il annonce la TOXP (Transmission Opportunity budget) à travers "beacon signals" pour chaque AC (sauf ACo). Si TOXP est épuisé pour un AC, le nouveau flux ne peut gagner un temps de transmission et le flux existant ne peut augmenter son temps de transmission. Le TOXP est un paramètre de différenciation défini dans IEEE 802.11e, c'est la période de temps pendant laquelle un flux appartenant à un AC a le droit d'initier une transmission dans le canal radio.

Dans [LW08], les auteurs étaient les premiers à proposer un algorithme de contrôle d'admission pour les réseaux sans fil multiples-sauts pour le mécanisme EDCA: l'amélioration directe de DCF dans le protocole de standard 802.11e. Dans ce cadre, ils ont présenté leur modèle par un graphe de contention où les nœuds sont les liens, suivi d'une estimation de capacité exigée par un flux entrant, appartenant à une catégorie prédéfinie de chaque clique maximale dans le graphe. Si le flux entrant a une capacité inférieure dans chaque clique à la capacité du réseau, il sera accepté. Cette approche traite malheureusement le cas d'un environnement multiples-sauts, mais à un seul canal ce qui n'est pas le cas actuel des réseaux maillés qui sont à multiples-sauts, multiples-radios et multiples-canaux. La modélisation proposée, où les auteurs transforment le graphe en un nouveau dans lequel ils remplacent les liens par des nœuds, peut s'avérer coûteuse. Par exemple un graphe avec n nœuds (donc n^2 liens) sera modélisé par un graphe de conflit, où il y a autant de nœuds que de liens (n^2) dans le graphe initial. Ce qui démontre que la complexité de tels modèles est de l'ordre de (n^2). La deuxième limitation de cette proposition concerne le passage à l'échelle (multiples-canaux, un seul radio). Le modèle, tel qu'il est, ne peut être étendu aux réseaux maillés à multiples-canaux à un seul radio que d'une façon unique, à savoir: de boucler sur le nombre de canaux (CA) et de construire le graphe de contention CA fois. Par rapport au passage à l'échelle (multiples-canaux, multiples-radios), si nous tenons à conserver leur modèle, nous aurons des graphes disjoints pour chaque canal, ce qui induit par conséquent un problème de commutation entre canaux. Dans le cas des

réseaux à multiples-canaux et multiples-radios, qui est un contexte assez fréquent dans les réseaux sans fil actuels, leur modèle ne fonctionne pas tel qu'il est.

Toutes ces études ont traité seulement le cas d'un seul canal, la plupart d'entre elles ont proposé une approche centralisée. Dans un contexte pratique, ce n'est pas réalisable avec un réseau de 3000 nœuds, comme c'est le cas de celui que Bell Canada veut implanter à Toronto. Prendre en compte les critères de multiples-radios et multiples-canaux, le trafic en relève et/ou les applications multimédia, un compromis/une optimisation entre/de plusieurs critères du réseau et la proposition d'une approche distribuée, n'est pas un travail évident. Cela nécessite beaucoup d'études et de travaux de recherches.

2.3.4.3 Récapitulatif des travaux sur le contrôle d'admission

Dans la table suivante, nous présentons les critères et les caractéristiques réseaux respectés par les différents papiers détaillés ci-dessus, ainsi que celles de notre approche proposée dans les trois volets de la thèse.

Table 2. Récapitulatif des travaux sur le contrôle d'admission

Réseau	[LC00]	[GC02]	[YK05]	[LG06]	[XS00]	[GZ03]	[KL03]	[XL04]	[LW08]	Notre approche
Un canal	X	X	X	X	X	X	X	X	X	
Une interface radio										
Multiples-canaux, Multiples-radios										X
Un saut					X	X	X	X		
Multiples-sauts		X	X	X					X	X
Différenciation du trafic (802.11)	X	X		X						

Différenciation de trafic (802.11e)			X		X		X	X		
Différenciation du trafic (802.11s)										X
Considérer le CFP										X
Centralisé	X			X			X	X	X	
Distribué		X	X		X	X				X
Considérer les nœuds au voisinage			X	X						X
Affectation statique des canaux										X
Affectation dynamique des canaux										X
Graphe de conflit									X	

2.3.5 Gestion adaptative

Durant les dernières années, beaucoup de travaux de recherche se sont focalisés sur le contrôle de la puissance [Kr04] et sur le routage [PB03], [CJ03] dans les réseaux ad hoc. Dans [Kr04], les auteurs ont étudié la gestion de puissance dans le but d'assurer le contrôle de la connectivité et de l'interférence. Peu nombreux sont les protocoles de gestion adaptative des ressources qui ont été proposés pour les réseaux ad hoc [WN99], [SS02], [SS03] et [OJ03]. Cependant, l'efficacité de ces modèles proposés a besoin d'être investiguée et bien étudiée dans le cas des réseaux maillés, surtout la détection du changement rapide de la topologie.

ANMP [WN99] (Ad hoc Network Management Protocol) a introduit le concept de groupe ("cluster") (voir section 2.3.7.2), où les nœuds dans le réseau participent dans

la construction du cluster et l'élection de "cluster-head" (un représentant/chef du groupe). La gestion dans ANMP est basée sur une approche hiérarchique où les "clusters-heads" regroupent les informations de gestion d'une manière centralisée de leurs membres dans leurs "clusters" correspondants.

Dans [SS02] et [SS03], les auteurs ont décrit une architecture appelée "Guerrilla", qui a pour but de faciliter l'adaptation, l'autonomie, l'économie, l'hétérogénéité et le passage à l'échelle dans les réseaux ad hoc. Par rapport à ANMP, Guerrilla permet aux représentants des "clusters" de prendre leurs propres décisions. C'est pour quoi cette architecture ne requière pas un ajustement périodique de sa configuration. Mais, Guerrilla et ANMP ne présentent pas un recouvrement suivi à une congestion du réseau. Ces deux modèles ont été élaborés pour un réseau à un seul canal. Cependant, l'étude du recouvrement, dans le cas d'une affectation dynamique des canaux, lors duquel des liens disparaissent et d'autres se créent, n'a pas été traité. Par contre, dans Guerrilla nous retrouvons une approche détaillée qui traite le cas des pannes des liens/nœuds efficacement.

Dans [OJ03], les auteurs décrivent un algorithme adaptatif qui ajuste dynamiquement les ressources des liens et la puissance de transmission pour maximiser la performance du système. Leur approche recherche les liens qui offrent la meilleure QoS et qui respectent les demandes des utilisateurs. Après, ces auteurs procèdent au calcul des puissances de transmission qui supportent ces exigences.

Une décomposition intéressante de la gestion des fonctions dans les WMNs a été présentée par Akyildiz [AW05]. Les composantes essentielles selon l'auteur sont (1) la gestion de la puissance (le but de la gestion de la puissance varie dans les WMNs; souvent les MRs n'ont pas la contrainte de puissance), (2) la surveillance réseau dans le but de veiller sur la performance statistique des nœuds. Ces données statistiques seront analysées par des algorithmes qui détermineront les anomalies potentielles contre lesquelles le serveur doit réagir. (3) La gestion de la mobilité, qui permet de fournir un accès sans fil aux utilisateurs mobiles. La dernière composante de ce modèle est (4) la gestion de la sécurité. Le dernier point ne constitue pas le

fondement de nos contributions. Nous constatons que cette décomposition n'inclut pas le mécanisme de contrôle d'admission. Cependant, un tel modèle risque de ne pas fonctionner correctement et de souffrir de la surcharge, la congestion et la dégradation de la QoS du réseau. L'approche que nous privilégions a ceci d'original qu'elle consiste en la combinaison d'outils complémentaires, à la fois pour l'étude, la modélisation et le développement de protocoles cross-layer [AW07], [AW08], comme pour la validation et l'évaluation des performances des WMNs.

2.3.6 Différentiation de services dans les réseaux sans fil

Les réseaux de télécommunications constituent aujourd'hui un formidable pont entre les hommes et la culture des autres, mais transporter des informations aussi différentes que la voix, les données et la vidéo nécessite des techniques de plus en plus élaborées. Initialement, la déferlante Internet n'a pas été conçue pour prendre en compte des informations de QoS, mais était juste fondée sur le principe du "Best Effort".

Le modèle "Best Effort" simplifie l'opération des équipements d'interconnexion tout en maximisant l'utilisation de ressources. Les paquets arrivants sont rejetés notamment quand la mémoire d'un routeur est saturée. C'est alors que les mécanismes de retransmission de TCP (Transport Control Protocol) recouvrent ces pertes, garantissant ainsi un transfert sans faille de bout en bout. Les algorithmes de "slow start" et de "congestion avoidance", implantés dans TCP, assurent de façon simultanée une distribution relativement équitable des ressources. Toutefois, même si le comportement de TCP répond aux besoins d'applications communément majoritaires dans le réseau comme Telnet ou FTP (File Transfer Protocol), son aptitude pour la transmission des flux avec des contraintes temporelles, dites des données isochrones³, laisse à désirer. La plupart des applications de visioconférence ou de voix sur IP (VoIP) se base sur le protocole UDP (User Datagram Protocol). L'utilisation de ce protocole n'implante pas de mécanisme de gestion de la congestion même s'il permet d'obtenir

³ Isochrone: se dit des flux de données dans les quelles l'écart de temps entre deux informations successives doit être constant. Au cas où le réseau de transmission introduirait un décalage, un mécanisme spécifique doit être mis en œuvre par le récepteur [CS03].

un débit relativement constant. Des mesures de contrôle sont impératives pour éviter la saturation du réseau en présence des flux temps-réel. En effet, ce contrôle permet d'assurer qu'un minimum de ressources soit disponible pour que les applications basées sur le protocole TCP puissent communiquer et que leurs performances ne soient pas affectées.

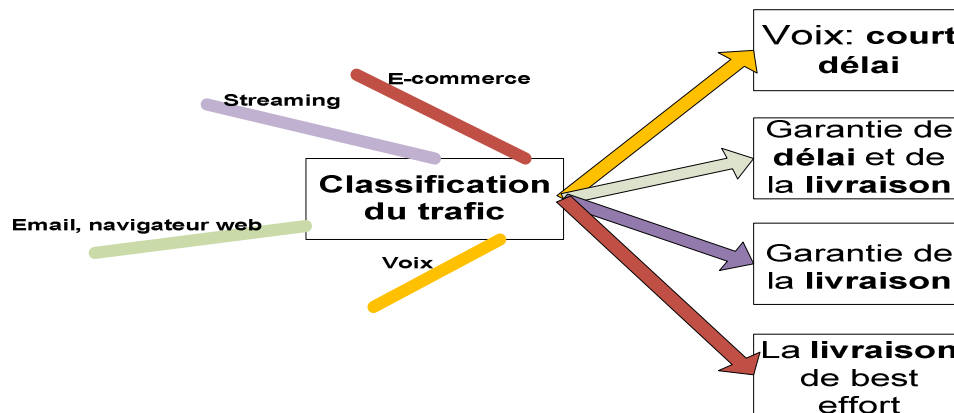


Fig. 12. Différentiation de services

La différenciation de services est une bonne solution pour satisfaire les applications avec des contraintes temporelles, c'est-à-dire, pour attribuer plus de priorité au trafic que nous jugeons important (cf. Fig. 12). La question réelle est de choisir quels sont les paquets qui doivent être supprimés en cas de congestion. La différenciation de services est la manière de préciser la priorité des différentes données, car certaines données sont plus importantes que d'autres. Dans la littérature, nous retrouvons d'une part l'approche Diffserv [PB96] que la plupart des constructeurs (ex, Cisco, Juniper, Procket) ont implémentée. Diffserv ne définit pas des classes de service; seules les fonctions sont définies, ce qui permet d'obtenir des classes flexibles. Dans Diffserv, deux types de routeurs sont considérés: les routeurs d'accès, qui procèdent à une gestion par flux et ont pour rôle de marquer les paquets et la possibilité de conditionner le trafic en entrée pour certaines classes, et les routeurs du cœur du réseau, qui procèdent à une gestion par classe, à la mise en tampon et à un ordonnancement selon le marquage. L'approche Diffserv ne considère pas un traitement par flux ce qui explique sa simplicité au cœur du réseau. Par ailleurs, pour différencier les services dans

IEEE 802.11, le mécanisme EDCA a été proposé (détaillé dans la section 2.2.2). Il est compatible avec le standard DCF. EDCA se base sur un algorithme probabiliste pondéré par la priorité pour l'accès au canal. Cette extension sera l'aboutissement de nombreuses recherches en termes de différenciation de services au niveau de l'accès au média radio. Dans le chapitre 4, nous exposons notre démarche pour garantir la différenciation de services en présence du trafic en relève pour répondre aux exigences des applications/utilisateurs mobiles avec contraintes temporelles.

2.3.7 Routage dans les reseaux sans fil

Dans le Chapitre 3, nous présenterons un protocole appelé RCAC qui garantit l'acheminement des flux de la source à la destination en leur garantissant un court délai et un faible taux de perte de paquets pour une approche distribuée. Pour cela, nous avons opté pour la présentation des contraintes communes à quelques protocoles de routage et pour la description des protocoles les plus connus et surtout ceux qui nous ont inspirés pour la proposition de notre protocole RCAC. Les liens sans fil subissent des interférences, cependant, les solutions de routage dans les WMNs doivent tenir compte des métriques radios les plus pertinentes. Dans les réseaux étendus, le protocole de routage est dominé par une surcharge et un encombrement du trafic. Alors, une surcharge constante de protocole de routage est inévitable, des chemins alternatifs sont continuellement calculés et actualisés, ce qui fait que le re-routage se fait avec au moins une perte minimale des paquets; du moins c'est ce que nous attendons d'un protocole de routage robuste et efficace.

2.3.7.1 Les protocoles de routage les plus connus

Les protocoles de routage sont conçus essentiellement pour l'établissement et l'entretien des routes, pour que les messages soient correctement acheminés dans le réseau. Les caractéristiques des réseaux WMNs rendent l'utilisation des protocoles filaires habituels inadaptée.

Plusieurs protocoles de routage proactifs, réactifs ou hybrides ont été proposés. Les chercheurs ont procédé, soit en améliorant les protocoles de routage de l'Internet, soit en mettant en œuvre de nouveaux algorithmes de routage. Dans le cas où les

algorithmes de routage proactifs (par exemple, DSDV [PB94], OLSR [CJ03]) sont utilisés, nous savons que chaque nœud dans le réseau doit maintenir des informations de routage et les stocker dans des tables. Le protocole entretient la consistance de ces tables, en envoyant périodiquement des mises à jour à travers le réseau. Par contre, dans le cas où un protocole de routage réactif est utilisé, les routes sont créées à la demande, respectant le principe d'une demande de réponse. Ainsi, un nœud qui veut commencer une communication avec un autre nœud lance un processus de découverte de route ("Route Discovery"). Si la route est trouvée, une procédure la maintient jusqu'à ce qu'elle ne soit plus utilisée. Parmi les protocoles de routage réactifs nous comptons par exemple: AODV [PB03] et DSR [JM04]. L'objectif de ces protocoles est de trouver le chemin le plus court entre une source et une destination. Ils prennent en compte l'état du nœud et la configuration du réseau quand une route est sollicitée. Il y a aussi une famille de protocoles de routage hybride qui mixe les protocoles de routage réactifs et proactifs (par exemple, ZRP [HP02]). Nous allons décrire dans ce qui suit deux exemples de protocole de routage proactif et un autre réactif.

- **Distance Source Distance Vector routing (DSDV)**

DSDV [PB94] est un protocole de routage proactif basé sur l'idée classique de l'algorithme distribué de "Bellman-Ford" en rajoutant quelques améliorations. Chaque station mobile maintient une table de routage qui contient: toutes les destinations possibles, le nombre de nœuds (ou de sauts) nécessaire pour atteindre la destination et le numéro de séquences (SN : Sequence Number) qui correspond à un nœud destination. Le SN est utilisé pour faire la distinction entre les anciennes et les nouvelles routes, ce qui évite la formation des boucles de routage. La mise à jour dépend donc de deux paramètres: Le temps, c'est-à-dire la période de transmission, et les événements. Un paquet de mise à jour contient: le nouveau numéro de séquence incrémenté du nœud émetteur et, pour chaque nouvelle route, l'adresse de la destination, le nombre de nœuds (ou de sauts) séparant le nœud de la destination et le numéro de séquence (des données reçues de la destination) tel qu'il a été estampillé par la destination. En effet, DSDV élimine les deux problèmes de boucle de routage "routing loop" et celui du "counting to infinity". Cependant, dans ce protocole, une

unité mobile doit attendre jusqu'à ce qu'elle reçoive la prochaine mise à jour initiée par la destination afin de mettre à jour l'entrée associée à cette destination dans la table de distance. Ce qui fait que le DSDV est un peu lent. DSDV utilise une mise à jour périodique et basée sur les événements, ce qui cause un contrôle excessif dans la communication. Mais, une étude précédente [YJ08] montre que l'algorithme de Bellman-Ford a un surcoût faible des messages de contrôle par rapport à la diffusion basée sur la découverte dans AODV [PB03].

Lors de la proposition de notre protocole de routage RCAC avec QoS, dans le chapitre 3, nous nous sommes basés sur DSDV en l'améliorant pour l'adapter dans un contexte des WMNs pour permettre un court délai et un faible taux de perte de paquets ainsi qu'un surcoût et une surcharge acceptable.

- **Optimized Link State Routing (OLSR)**

OLSR [CJ03] est un protocole de routage proactif. Il est une amélioration directe du protocole à état des liens filaires pour les réseaux mobiles ad hoc. Son optimisation se manifeste dans sa manière d'économiser lors des diffusions des ressources radios, en utilisant un système de relais multi-points (Multi-Points Relays : MPR). Dans ce système, chaque nœud choisit un sous-ensemble de ses voisins immédiats (MPR) pour retransmettre ses paquets en cas de diffusion et pour atteindre tous ses voisins à un saut ou deux-sauts. Dans le cas des protocoles proactifs, des messages de contrôle périodiques doivent être envoyés pour l'entretien des tables de routage. Deux types de messages sont utilisés pour OLSR: "Hello" et "TC" (Topology Control). Périodiquement, chaque nœud diffuse localement un message "Hello" contenant des informations sur son voisinage et sur l'état des liens. Chaque nœud prend connaissance de son voisinage au plus à deux-sauts grâce à ces deux types de messages. Ainsi, pour créer les tables de routage, chaque nœud génère périodiquement un paquet TC, qui est diffusé dans tout le réseau et qui contient la liste de ses voisins l'ayant choisi comme MPR. Seuls les voisins MPR rediffusent un paquet TC reçu pour éviter l'inondation. Cet algorithme réduit énormément la surcharge générée par le trafic de contrôle. La table de topologie peut être créée lors de la réception d'un message TC.

Basé sur cette table, chaque nœud peut calculer la table de routage qui permet de délivrer les paquets vers n'importe quelle destination dans le réseau.

- **Ad-hoc On Demand Distance Vector (AODV)**

AODV [PB03] est un algorithme de routage réactif à la demande. Il est basé sur le principe, vecteur à distance. Il ne construit de routes entre les nœuds que lorsqu'elles sont demandées par les nœuds sources afin de réduire le nombre de diffusions de messages. AODV introduit les numéros de séquence dans le but d'entretenir la consistance des informations du routage. Les numéros de séquences permettent d'utiliser les routes les plus récentes. AODV commence la recherche par une inondation d'une requête de route (RREQ : Route REQuest) dans le but de créer un chemin vers une destination. Toutefois, si la destination n'est pas connue au préalable, si le chemin existant vers la destination a expiré, ou s'il est devenu défaillant, la route peut ne pas exister. Néanmoins, AODV garde une table de routage, au niveau de chaque nœud de transit tout en maintenant les chemins d'une façon distribuée. Une transmission périodique du message "Hello" est effectuée dans le but de maintenir des routes cohérentes. Quoique, si on ne reçoit aucun message "Hello" après un laps de temps bien déterminé à partir d'un nœud voisin, le lien est considéré comme défaillant. Par ailleurs, quand la topologie du réseau ad hoc change, AODV ne présente pas de boucles de routage et offre une convergence rapide.

2.3.7.2 Regroupement passif (passive clustering)

L'intérêt que nous portons à la notion de regroupement passif revient au fait que nous nous en sommes inspirés pour adapter le concept de cliques de la théorie de graphe à notre contexte. Le regroupement passif est une approche de formation de groupes ("clusters") qui n'utilise pas de protocole de signalisation et ne nécessite aucune surcharge due aux échanges des informations du voisinage. En plus, le regroupement passif construit la structure de groupes avec les chefs de groupes et les nœuds de liaison sur demande; ce qui permet de minimiser les coûts de maintenance d'une structure de groupe. Kwon et al. sont les premiers à proposer un protocole de groupage passif, appelé PC [KG02]. Cependant, le groupage passif n'utilise aucun

message dédié à la construction des "clusters". Les "clusters" ne sont créés que lorsque nécessaire c'est-à-dire lorsqu'un nœud a une information à diffuser. Le protocole de regroupement passif utilise alors ces messages d'informations pour construire les "clusters" en ajoutant des champs aux paquets d'informations. Dans ce protocole, un nœud peut avoir différents statuts possibles tels que : "cluster-head", passerelle, ordinaire et non défini. Par défaut, le statut des nœuds est non défini. Seul un nœud ayant un statut non défini peut devenir "cluster-head". Si un tel nœud a un message à envoyer, il se déclare "cluster-head" et diffuse son statut en l'ajoutant à l'information qu'il doit envoyer. Les nœuds voisins d'un "cluster-head" deviennent des nœuds ordinaires, les nœuds voisins de plusieurs "cluster-heads" deviennent des passerelles. Les nœuds ordinaires ne relaient pas les messages de diffusion. Aucun message n'étant dédié à la maintenance de la structure. Les passerelles et les nœuds ordinaires activent des compteurs lorsqu'ils reçoivent des nouvelles de leur(s) "cluster-head(s)". S'ils restent sans nouvelles d'eux, le temps que leur compteur expire, les nœuds ordinaires reprennent un statut non défini et les passerelles prennent le statut de nœud ordinaire ou non défini suivant le nombre de "cluster-heads" qu'elles entendent encore.

2.4 Les approches existantes d'affectation des canaux

Il existe une vaste littérature sur les approches d'affectation des canaux dans les réseaux sans fil. Une étude approfondie leur est dédiée dans [SV04]. Ces approches peuvent être classées en trois grandes catégories (cf., Fig. 13): (1) affectation fixe des canaux (AFC): AFC attribue les canaux une seule fois et cette affectation des canaux ne change pas au fil du temps [RG04], [MD05] et [SG07], (2) affectation dynamique des canaux (ADC): ADC met continuellement à jour l'attribution des canaux pour améliorer les performances [RW07], [SG08], [SV04], [BC04], [RC04] et [RC05]; et (3) affectation hybride des canaux (AHC): AHC applique une approche AFC pour certaines interfaces et une approche ADC pour d'autres [RC04], [KV05], [KV06] et [RB06].

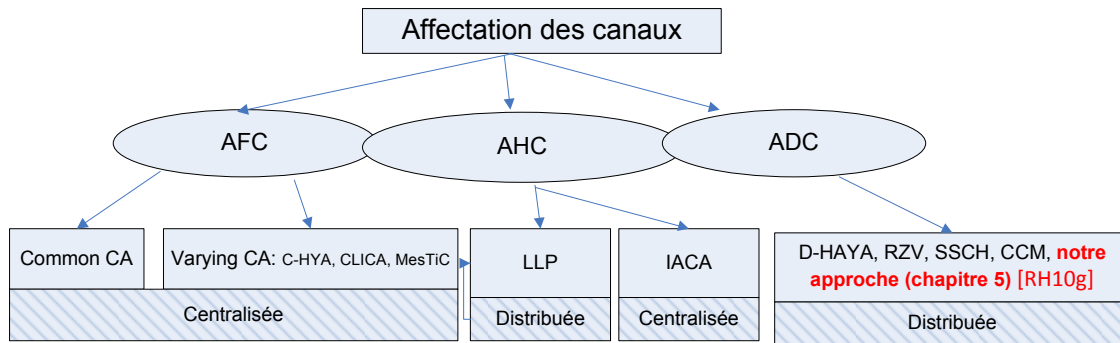


Fig. 13. Une taxonomie des approches d'affectation des canaux dans les WMNs

2.4.1 Affectation fixe des canaux (AFC)

Dans l'approche AFC, un ensemble de canaux est affecté de manière permanente aux interfaces en prenant en compte le temps de commutation entre interfaces. Ces canaux peuvent être réutilisés dans d'autres MRs, à une distance suffisante, de sorte que les interférences co-canal soient tolérables. L'avantage de cette approche réside dans sa simplicité; son inconvénient est qu'elle n'adapte pas sa réutilisation du canal pour tenir compte de changement de charge du trafic offerte. En fait, les approches basées sur AFC peuvent être subdivisées en (1) une affectation commune des canaux (Common Channel Assignment-CCA) [RG04] c'est-à-dire les interfaces radios de chaque nœud sont toutes affectées à la même série de canaux et (2) une modification de l'attribution des canaux (Varying Channel Assignment-VCA) [MD05], [RG07] c'est-à-dire les interfaces des différents nœuds peuvent être affectées à des ensembles différents des canaux, comme dans CLICA (Connected Low Interference Channel Assignment) [MD05], et dans MesTiC (Mesh-based Traffic and Interference aware CA) [RG07] (cf., Fig. 13).

2.4.2 Affectation dynamique des canaux (ADC)

Dans l'approche ADC, n'importe quelle interface peut être attribuée à n'importe quel canal, et les interfaces peuvent souvent changer d'un canal à l'autre. Par conséquent, lorsque les nœuds ont besoin de communiquer les uns avec les autres, un mécanisme de coordination [RG04], [MD05], [SV05], [SG07] et [BC04] doit veiller à ce qu'ils utilisent le même canal. L'avantage de l'approche ADC est la possibilité de

réaffecter une interface à n'importe quel canal offrant ainsi la possibilité d'utiliser plusieurs canaux avec un nombre minimal d'interfaces. Toutefois, les principaux défis sont les délais de commutation du canal (typiquement de l'ordre de quelques millisecondes dans les cartes sans fil 802.11) et la nécessité de mécanismes de coordination pour la commutation des canaux entre les nœuds. L'approche distribuée dite "Hyacinthe" (D-HYA) est une version distribuée de l'approche centralisée "Hyacinthe" (C-HYA) [RW07, RC05] qui réagit aux variations de la charge du trafic pour améliorer l'équilibrage de la charge et le débit. Toutefois, D-HYA pose un problème potentiel pour le routage à multiples-routes dans les réseaux maillés et ne considère pas la connectivité du réseau.

2.4.3 Affectation hybride des canaux (AHC)

Dans l'approche AHC, les concepts des stratégies fixes et dynamiques sont combinés par l'application d'AFC à certaines interfaces et ADC à d'autres [KV06]. Les approches basées sur AHC sont attrayantes parce que, comme avec AFC, elles utilisent de simples algorithmes de coordination tout en conservant la souplesse offerte par l'approche ADC. Link Layer Protocol (LLP) [KV06] classe les interfaces disponibles en interfaces fixes et commutables, c'est-à-dire elles peuvent être commutées sur des laps de temps entre des canaux non-fixes, dépendamment de la quantité du trafic. LLP attribue différents canaux à des interfaces fixes des différents nœuds. Ainsi, tous les canaux peuvent être utilisés. En outre, les cartes d'interfaces réseau sont utilisées uniquement pour la connectivité. Toutefois, LLP ne prend pas en compte la charge du trafic lors de l'attribution des canaux fixes. Interference-Aware Channel Assignment (IACA) [RB06] est une approche basée sur le graphe de conflits à multiples-radios. IACA offre des routes en alternative et évite les perturbations des flux de redirection du trafic sur un canal par défaut. Néanmoins, IACA ne considère pas la charge du trafic lors de l'attribution des canaux.

2.4.4 Optimisation des différents critères dans les WMNs

Optimiser simultanément un ensemble de critères (par exemple, débit, équité, interférences, etc.) dans les réseaux sans fil peut s'avérer impossible, mais devenu

incontournable dans la réalité. En effet, comme observé déjà dans la littérature concernant les réseaux 802.11 [JL05], maximiser l'équité réduit considérablement le débit global dans le réseau. Par contre, si l'équilibrage de la charge est maximal, le débit global est indirectement augmenté [DV06]. De même, l'attribution d'un même canal à des routeurs voisins minimise le surcoût de re-routage du trafic en relève (par rapport à un trafic en relève sur des canaux différents), mais elle peut causer des interférences importantes et donc dégrader les performances du réseau en termes de débit. Par conséquent, l'optimisation de l'un de ces critères affectera les autres. Ainsi, une approche à multiples-objectifs est définitivement recommandée pour ce genre de problème. Dans le chapitre 5, nous proposons un modèle d'optimisation à multiples-objectifs pour résoudre le problème de réaffectation des canaux dans un MR-WMN qui offre la possibilité d'optimiser la combinaison de plusieurs critères. Nous avons eu recours à un compromis entre les critères suivants: 1) le débit, 2) l'équité entre utilisateurs, 3) l'équilibrage de la charge du trafic et 4) le surcoût/le délai du re-routage du trafic en relève. Non seulement le problème d'affectation des canaux est un problème NP-difficile, mais encore le coût d'évaluation des fonctions à multiples-objectifs est très élevé par rapport aux simples fonctions objectif. Par conséquent, des méthodes heuristiques appropriées doivent être soigneusement conçues pour résoudre ce problème. Les travaux qui, dans la littérature, ont présenté le problème d'affectation des canaux comme un problème d'optimisation seront discutés ultérieurement; sachant que, dans le troisième volet de la thèse, nous proposons maintes méta-heuristiques pour résoudre ce problème pour un réseau de taille réelle.

2.4.5 Récapitulatif et comparaison des approches d'affectation des canaux

L'étude des solutions existantes, qui ont été proposées pour le problème d'affectation des canaux dans les WMNs, nous permettent de déduire que tous ces travaux de recherche n'ont pas considéré le trafic en relève lors de l'affectation des canaux, ce qui probablement provoque des communications interrompus et augmente le surcoût et le délai de re-routage dans le réseau sans fil. Réduire le surcoût de re-routage et permettre à la plupart des trafics re-routés de changer seulement le premier-saut tout

au long de la route du trafic en relève remédient certainement aux problèmes rencontrés par ces travaux [SX08].

En effet, dans plusieurs contributions, le problème d'affectation des canaux a été combiné à d'autres problèmes de performances et a été présenté comme étant un problème d'optimisation [MD05], [RW07], [SG08]. Dans [RW07], les auteurs ont choisit de maximiser une fonction qui reflète une équité (c'est-à-dire la capacité du réseau de permettre à un utilisateur de profiter d'une équité du débit total réalisé dans le réseau) proportionnelle entre les flux dans la couche MAC. Mais, ils ont négligé l'optimisation du débit total dans le réseau et n'ont pas montré s'il y a une corrélation entre les deux critères, à savoir l'équité entre utilisateurs et le débit dans le réseau. De même, ils n'ont pas considéré un trafic non-uniforme sur les liens logiques, qui sont généralement affectés par la demande du trafic et le protocole de routage. Subramanian et al. [SG08] traitent le problème d'affectation des canaux en considérant une seule fonction objective, à savoir la minimisation de l'interférence dans le réseau. Ils ont formulé le problème en utilisant un programme linéaire. Les auteurs ont étendu leur modèle pour adapter le trafic non-uniforme. Ils ont utilisé la méta-heuristique basée sur la recherche Tabu pour obtenir des solutions d'encadrement à ce problème NP-difficile, mais ils n'ont pas considéré l'équité entre MCs dans leur solution. En plus, la nature centralisée de l'approche d'affectation des canaux basée sur la recherche Tabu, fait que leur solution n'est adéquate que pour des réseaux "Managed" et il est à noter qu'elle ne prend pas en compte la variation dynamique du trafic.

Dans le chapitre 5, par opposition aux modèles/mécanismes existants, qui eux s'attaquent au problème considérant une ou deux fonctions objectif, et qui sont reliés principalement à l'interférence [MD05], [RW07], [SG08], nous proposons un nouveau modèle d'optimisation à multiples-objectifs pour le problème de réaffectation des canaux. Notre modèle considère et étudie l'optimisation de quatre fonctions objectif au mieux que possible pour avoir une bonne solution. Particulièrement, ce modèle propose en plus de maximiser le débit et l'équité entre utilisateurs, de minimiser le surcoût et le délai de re-routage des utilisateurs mobiles et/ou des applications multimédia et aussi

de minimiser la variation de la charge du trafic pour réaliser un équilibrage de la charge adéquat.

2.5 Synthèse

Dans ce chapitre, nous avons passé en revue les différentes solutions de dimensionnement (c'est-à-dire augmenter la capacité), de gestion des ressources (contrôle d'admission, etc.), de différenciation de services et d'optimisation d'affectation des canaux aux interfaces d'un MR-WMN. Malgré un certain nombre de solutions existantes dans la littérature, nous pouvons constater qu'il y a encore beaucoup de chemin à faire. En effet, nous ne pouvons pas trouver de schéma de gestion adaptative des ressources qui couvre tout le cycle de vie d'un réseau en ayant en vue de mieux contrôler le trafic entrant avec QoS exigée en termes de longueur du délai et du taux de perte de paquets.

Nous consacrerons la suite de cette thèse à la recherche de telles solutions. Ci-dessous, nous présentons notre approche (voir Fig. 14). Notre modèle proposé dans la Fig. 14 est composé de trois contributions majeures qui comportent des entités, parmi lesquelles on nomme principalement: contrôle d'admission, routage avec QoS, différenciation de services et réaffectation dynamique des canaux. Dans le chapitre 3, nous détaillons (1) l'entité contrôle d'admission ainsi que le routage avec QoS. Dans le chapitre 4, nous présentons (2) l'entité différenciation de services entre différent type de trafic. Dans le chapitre 5, nous détaillons (3) la réaffectation des canaux aux interfaces en présence du trafic en relève.

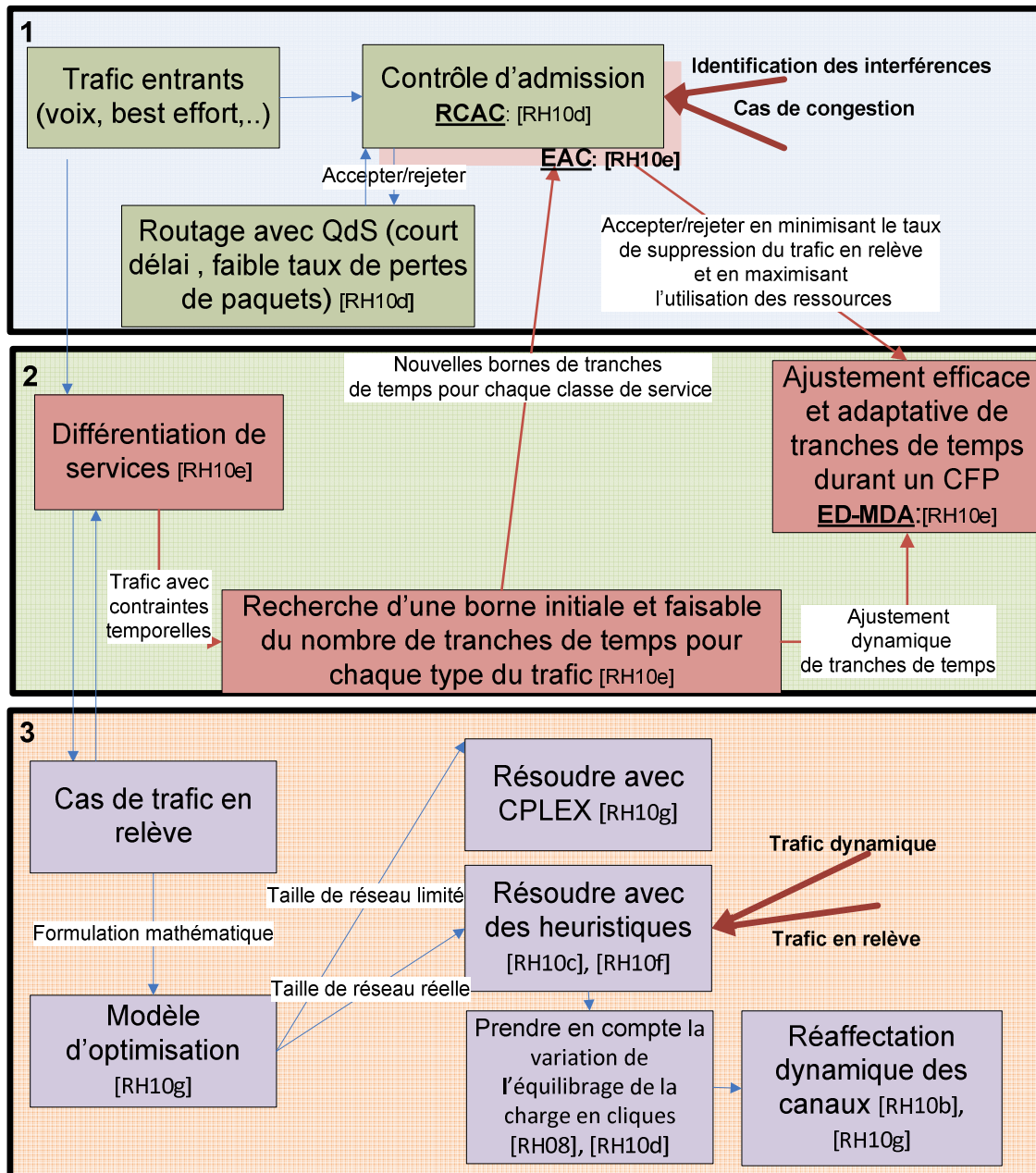


Fig. 14. Modèle du support de la QoS de notre approche dans les WMNs

Chapitre 3: Distributed Admission Control in Wireless Mesh Networks: Models, Algorithms, and Evaluation

Jihene Rezgui, Abdelhakim Hafid, Michel Gendreau

Abstract

Wireless mesh networks (WMNs) have attracted increasing attention from the research community as a high performance and low-cost solution to last-mile broadband Internet access. In WMNs, admission control is deployed to efficiently control different traffic loads and prevent the network from being overloaded. This paper introduces a distributed admission control scheme for WMNs, namely, routing on cliques admission control (RCAC). In particular, we propose an analytical model to compute the appropriate acceptance ratio and guarantee that the packet loss probability (PLP) in the network does not exceed a threshold value. The model also allows computing end-to-end delay to process flow requests with delay constraints. RCAC achieves scalability since it partitions the network into cliques, and only clique heads (CHs) are involved in the admission control procedure. Using extensive simulations, we demonstrate that our RCAC achieves high resource utilization by providing lower blocking probabilities in a dynamic traffic load environment while satisfying quality of service (QoS) constraints in terms of PLP and end-to-end delay. Moreover, we show that a contention access (CA) enforced with our RCAC outperforms the mesh deterministic access (MDA).

Keywords: Admission control, multichannel, quality of service (QoS), wireless mesh networks (WMNs).

Status: This paper is published in IEEE Journal Transactions on Vehicular Technology [RH10d]. A part of this article was presented on the following published paper:

- Distributed Admission Control in Wireless Mesh Networks, IEEE Broadnets, 2008, UK [RH08].

3.1 Introduction

Wireless mesh networks (WMNs) have recently emerged as a promising technology for next-generation wireless networks. A WMN consists of two types of nodes: 1) mesh clients (MCs) and 2) mesh routers (MRs). The MRs form a wireless mesh backbone infrastructure that forwards most of the traffic between MCs and Internet gateways. In general, MRs have very rare mobility and operate just like stationary routers, except that they are connected by wireless links using the very popular IEEE 802.11 wireless local area network (WLAN) standard or other wireless technologies such as WiMax. Using more than one radio interface in each MR allows for more channel diversity, resulting in less interference and, therefore, more throughput and capacity. However, this only improves the Best Effort traffic since supporting QoS for real-time traffic in WMNs remains an open challenge.

Admission control is one of the key traffic-management mechanisms that must be deployed to provide QoS support. A new traffic flow is accepted into the network only if there are sufficient resources. Admission control is the premise to the implementation of QoS routing, channel assignment, and multiple radios/channels scheduling, which have been proposed [KV05], [AB06], [KN05] to improve capacity and maximize throughput in WMNs. Indeed, when the network is overloaded, none of these schemes can prevent QoS degradation. The accuracy of admission control depends upon how well the network capacity is estimated. The estimation is difficult to obtain because, compared with wired networks, the links in WMNs are inherently shared, because of interferences, and difficult to isolate; this fact makes the performance of WMNs difficult to control. It is crucial that admission control considers both local resources and resources at neighboring nodes when analyzing the network performance [YK05], [LG06]; the reason is that interferences among links cause performance degradation, e.g., two interfering links that are simultaneously active often provide less throughput than two separated links.

In this paper, we propose a distributed admission control mechanism for WMNs, namely, routing on cliques admission control (RCAC); a preliminary version of RCAC has appeared in [RH08]. RCAC accepts a new flow request only when there are enough available resources to carry the flow while satisfying predefined thresholds of packet loss probability (PLP) and end-to-end delay. This will avoid situations in which uncontrolled resource usage leads to network breakdown (i.e., severe congestion). RCAC partitions the WMN into cliques, in which all vertices are adjacent to each other. A maximal clique is a clique that belongs to no other larger cliques. Only clique heads (CHs) are involved in the admission control procedure; this makes RCAC scalable for large-sized networks. Inside a clique, RCAC computes the available bandwidth while making use of local bandwidth information and neighboring bandwidth information; this is necessary to take care of interferences among one and two-hop-away nodes. To the best of our knowledge, RCAC is the first admission control mechanism for WMNs to consider two QoS parameters: 1) packet loss and 2) end-to-end delay in multichannel and multi-radio WMNs. RCAC attempts to answer the following question: For a given WMN, can new flows be accepted into the network while keeping the PLP under the PLP threshold and the end-to-end delay under the threshold value?

In our proposed approach, we take into consideration the knowledge of both local and neighboring resources in a distributed stochastic analytical model with two QoS parameters: 1) delay and 2) packet loss. We model interconnected CHs as a queuing network, and we approximate the PLP with the overflow probability in each clique; to this end, we estimate the total packet arrival at time t in each clique. The objective of our proposal is to compute an acceptance ratio for a given PLP. Indeed, for a threshold value of PLP, we are able to determine the number of flows that can be accepted into the network while satisfying the threshold. Our proposal also takes end-to-end delay into account when processing new flow requests.

The remainder of this paper is organized as follows. Section 3.2 presents related work. Section 3.3 presents our notations, assumptions, and network model followed by an analytical model in Section 3.4. Section 3.5 elaborates our admission control

algorithm and enhanced routing protocol in detail. Section 3.6 evaluates the proposed admission control via simulations. Finally, Section 3.7 concludes this paper.

3.2 Related work

The new challenges in WMNs require more research effort from different perspectives to provide QoS management. The authors in [DY02] report that it is necessary to have a mechanism for admission control. However, they do not present any specific solution. Studies specifically focusing on admission control over IEEE 802.11 [YK05], [LG06], [GC02], [KG07] have mainly considered the available bandwidth to decide whether the request is accepted or not. In the stateless wireless ad hoc network (SWAN) [GC02], the admission controller listens to all packet transmissions to collect information about bandwidth and congestion. It proceeds by sending probe messages. However, probing causes a lot of overhead and packet loss. In addition, SWAN does not consider the fact that two nodes could contend for a channel even without directly communicating with each other. The authors in [YK05] proposed the contention-aware admission control protocol (CACP) mechanism. The CACP provides admission control for flows in a single-channel ad hoc network based on the knowledge of both local resources at a node and the effect of admitting new flows on neighboring nodes. A scheme closely related to the CACP was proposed in [LG06], which integrates admission control with ad hoc routing and channel reuse due to parallel transmissions for a more accurate estimation of channel utilization. In [KG07], the authors developed a measurement based capacity utilization model for an IEEE 802.11-based mesh network; they also introduced routing metrics such as the maximum residual feasible path and new strategies like routing using call statistics. The proposed model provides the available capacity of each node for the admission of new voice-over-IP (VoIP) calls. The proposed approach [KG07] has been evaluated for a single-channel mesh network using only VoIP traffic. The applicability/adaptation of this approach in the case of multiple channels and traffic types has not been investigated.

Several schemes have been proposed to evaluate the expected bandwidth of a wireless network over IEEE 802.11e to provide a criterion for admission decision

[GZ03], [KL03], [XL04], [LN08]. In [GZ03], the authors proposed a method where each node measures both the occupied bandwidth and the average collision ratio; the measured value is compared with a given threshold; then, a decision to accept or not a new flow is made using a simple rule. The authors in [KL03] introduced an analytical model to evaluate the expected bandwidth and packet delay of each class of traffic based on the enhanced distribution coordination function (EDCF). This model provides a criterion for admission decision, as well as a theoretical analysis for EDCF. The proposed admission control strategy satisfies the required bandwidth and limits the packet delay of each traffic class to a predefined level. In [XL04], the quality access point measures the medium utilization and confirms the transmission opportunity budget (TOXP) through beacon signals for each Access Category (AC). If the TOXP is consumed for one AC, then the new flow could not gain transmission time, and current flows could not increase their transmission time. In [LN08], the authors proposed a centralized admission control mechanism; their model is based on the concept of conflict graph. The major problem with this model is that the utilization of the conflict graph is highly complex; even for a moderate-sized network, the number of interference constraints can be hundreds of thousands. The model works well in a multihop single channel for a small-sized network. However, the approach is centralized, which is not convenient for large networks.

The mesh deterministic access (MDA) studied in [HM07], [HM08], [CL08] extends the typical IEEE 802.11 medium instantaneous reservation procedure [which is known as the virtual carrier sensing (V-CS)] to a more advanced reservation procedure using scheduled MDA opportunities (MDAOPs) within a two-hop neighborhood. MDA aims to provide stringent medium access control (MAC) delay guarantees for real-time services such as VoIP. IEEE 802.11 V-CS [HM07] is performed by a four-way handshake procedure in which request-to-send, clear-to-send, data, and acknowledgment packets are exchanged between two communicating nodes, while a network-allocation vector is set by the other nodes in the physical sensing range. V-CS correctly works in a single-hop wireless network but can cause severe interference in wireless links that are multiple hops away and are sharing the same channel within an

overlapped transmission or interference range. Therefore, in multihop wireless networks, such as WMNs, a different access mechanism, namely, MDA, with less multihop interference, was adopted in the IEEE 802.11s draft amendment to provide stringent delay bounds. MDAOPs are first negotiated between neighboring mesh nodes by exchanging broadcast setup messages; then, MDAOP reservations in multiple time-slot units, during the delivery traffic indication message (DTIM) periodic interval, are performed. To limit the message broadcast signaling overhead, MDA-related messages are sent only within the two-hop neighborhood.

However, interference outside the two-hop neighborhood can still occur with MDA; this may degrade the WMN performance. The authors in [CL08] tackle this problem using dynamic relocation (DR) of conflicting MDA time slots that are two hops away from each requesting node. In their proposed solution, if enough interference is detected outside the two-hop neighborhood, then flows are dynamically relocated, depending on their elapsed duration in the network. In fact, "older" flows will be less likely to be relocated, i.e., its MDAOPs will be moved to another available time slot.

3.3 Network Model

In this section, we propose a network model for WMNs and illustrate how an equivalent queuing and stochastic network model can be constructed. In particular, we first define key concepts, namely, connectivity graph and cliques, and then present the assumptions/notations used in the rest of this paper. The network model will be used in Section 3.5 to develop the proposed distributed admission control mechanism for WMNs.

3.3.1 Connectivity Graph

We consider a multihop WMN, as illustrated in Fig. 15. In this network, MRs (e.g., MR_3 and MR_{17}) aggregate and forward the traffic from the MCs that are associated with them. They communicate with each other to form a multihop wireless

backbone network. This backbone network forwards the user traffic to the gateway access points (e.g., P_1 and P_2), which are connected to the Internet.

We formally model the backbone of a WMN as an undirected graph called connectivity graph $G = (V, E)$, where V represents the set of mesh nodes, and E represents the set of edges between these nodes. Among these nodes, $P \subset V$ are the gateway access points that connect to the Internet. In the rest of this paper, MRs and gateway access points are collectively called mesh nodes. $\forall (u, v) \in V$, an edge $e = (u, v) \in E$ if the distance between u and v , which is denoted as $d(u, v)$, is smaller than the minimum range, which is denoted $\min(r_u, r_v)$, of u and v (i.e., $d(u, v) \leq \min(r_u, r_v)$), where r_u and r_v represent the radio transmission ranges of nodes u and v , respectively. Since we consider a multiradio and multichannel WMN, channel assignment is needed.

The connectivity graph with channel assignment is denoted as $G_A = (V, E, A_G)$, where $A_G = \{A_G(u), \forall u \in V\}$, and $A_G(u)$ is the set of channels assigned to u . We denote NC as the number of channels per radio and NR as the number of radios per node; typically, we have $NR \leq NC$.

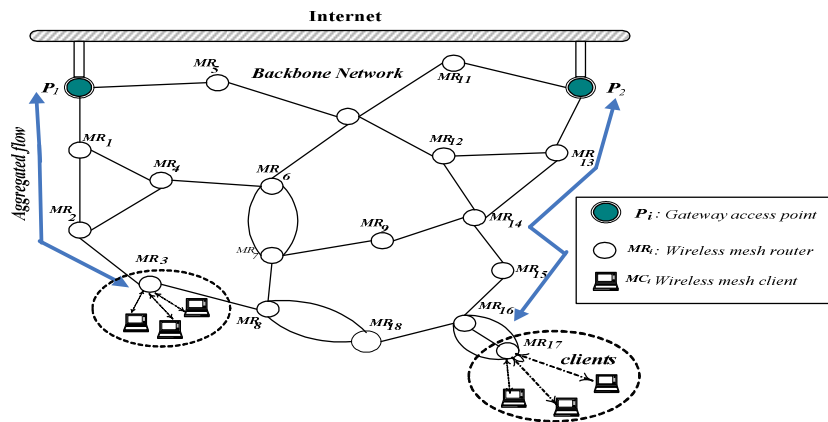


Fig. 15. A wireless mesh network

Fig. 16 shows an example of a connectivity graph $G_A = (V = \{A, B, C, D, E, F\}, E = \{(A, B) \dots (C, D)\}, A_G = \{A_G(A), \dots, A_G(F)\})$ in which we connect two nodes u and

v if they share the same channel and the distance between them is smaller than or equal to $\min(r_u, r_v)$. In this example, $NR = 2$ radios, and each node is labeled with its channel assignment; for example, D, with two radios, is assigned channels 1 and 2 ($A_G(D) = \{1, 2\}$). The radio transmission range of C and D is 200 and 250 m, respectively, the distance between C and D is smaller than $\min(200, 250)$, and they share channel 2; thus, the edge $(C, D) \in E$.

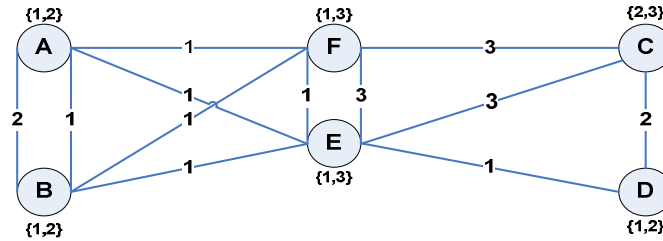


Fig. 16. Connectivity graph: G_A

3.3.2 Cliques

A clique is represented by an undirected graph where for each two vertices/nodes, in the graph, there exists an edge connecting them; all the edges in the graph use the same channel. A maximal clique is a clique to which no more vertices can be added.

The formal definitions of clique and maximal clique are given in the following.

Definition "Clique": Let $G = (V, E)$ be an undirected graph, where V is the set of vertices and $E \subset V \times V$ is the set of edges. A subset $S \subset V$ of vertices is called a clique if for every pair of vertices in S there is an edge in E , i.e., the sub graph introduced by S is complete.

Definition "Maximal Clique": A maximal clique S is a clique of which the proper extensions are not cliques i.e. for any S' if $S \subset S'$ and $S \neq S'$ then S' is not a clique.

We use maximal cliques to determine the nodes which compete to access the same channel. Therefore, two nodes i and j that belong to the same clique must not be active simultaneously. Our proposed admission control scheme takes into account both local and neighboring resources. Therefore, we define two types of maximal cliques (see Fig. 17): 1) A-clique is defined as a set of nodes, A , sharing the same channel and having a pairwise distance smaller than or equal to the minimum radio transmission range of the pair nodes (i.e., $d(u, v) \leq \min(r_u, r_v), \forall u, v \in A$); and 2) B-clique is defined as a set of nodes, B , that use the same channel and have a pairwise distance in the interval $]\min(r_u, r_v), R]$ $\forall u, v \in B$ where R is the interference range; B-cliques are used to identify nodes in the Carrier Sense Range (CSR).

Fig. 17 shows maximal cliques of type A and B computed using the connectivity graph shown in Fig. 16. For example $q1/w=1$ is an *A-clique* that consists of four nodes $\{A, B, F, E\}$. All of these nodes share the same channel 1. $\{A, D\}/w=1$ is a *B-clique*, where the distance between nodes A and D is in $]\min(r_A, r_D), R]$ and they share the same channel 1.

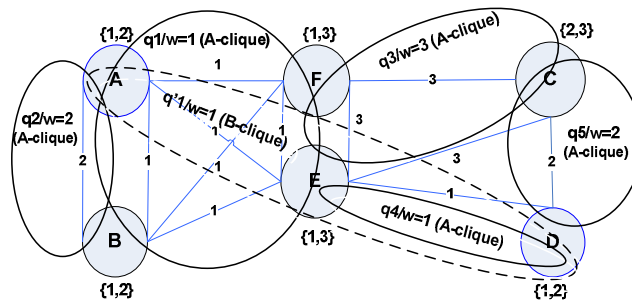


Fig. 17. A-cliques and B-cliques

Each *A-clique* is represented by its CH. In this paper, we use a simple algorithm that selects a node with the smallest degree as the CH; other algorithms can be used with no changes to our proposed admission control scheme. CHs are the only nodes involved in the admission control procedure; more specifically, they are responsible for computing the acceptance ratio, the available bandwidth, the maximum occupancy, and the average service time of *A-cliques* they represent (see details in the next sections).

Algorithm 1 details the election of CHs. In this algorithm, we have three inputs: 1) *A-cliques*; 2) node degree; and 3) MAC address. Here, the degree of node *X* is denoted as *X.degree*, which is equal to the number of *A-cliques* to which *X* belongs. The MAC address is used as a tie breaker when two or more nodes have the same smallest degree.

Algorithm 1. CH election algorithm

Input: *A-cliques*, Node degree, MAC address

Output: Head-of-*A-cliques*

```

1 X.state:= CH; /* Each node X declares itself as a CH*/
2 Every node X broadcasts a CH_state_notify (X.degree);
  /*All nodes in A-clique receive X-degree*/
3 Upon receipt of a notification from X, a node Y performs the following:
4 if Y.degree == X.degree {
5   Y broadcasts a CH_state_notify (Y.degree,Y.MAC_addr);
6   if (Y.MAC_addr > X.MAC_addr) Y.state=ordinary;
  /*so that not all nodes will be CH*/
7   else Y.state=CH;
8 }else if (Y.degree > X.degree) Y.state=ordinary;

```

In the rest of this paper, *A-clique* and CH will be used interchangeably.

3.3.3 Assumptions/Notations

The assumptions considered throughout this paper are given as follows: 1) Each *A-clique* may experience a different load and may have a different capacity than other *A-cliques*. 2) New flows arriving in each *A-clique* are uniform, independent, and Poisson distributed. 3) All the information exchanges and the acceptance ratio computation happen only once at the beginning of each control period of length *T*. 4)

The assignment of channels is static. 5) Network failures (link/node) are not considered. To improve the readability of this paper, we gather all the notations in Table 3.

Table 3. List of Symbols/Parameters

V	The set of nodes (routers) in the network.
<i>Clique</i>	Every vertex is adjacent to each other.
<i>Maximal clique</i> q_i	The clique which does not belong to any other larger clique: (<i>A-clique</i> / <i>B-clique</i>).
Q_A	The total number of <i>A-cliques</i> .
CH_i	Clique Head of <i>A-clique</i> q_i .
C_i	Available bandwidth of <i>A-clique</i> q_i .
$N_i(t)$	Number of active flows at time t in <i>A-clique</i> q_i .
N_{ct}	Set of <i>A-cliques</i> that can contend for channel k utilization, $k \in \{1..NC\}$.
θ_k	Bandwidth of k^{th} channel.
NC	Number of channels per radio; channels are identified/numbered from 1 to NC .
NR	Number of radios per node.
MO_i	Maximum occupancy, i.e., maximum number of flows that can be supported by <i>A-clique</i> q_i .
T	Length of the control period.
$PA_i(t)$	Packet arrival rate at time t in <i>A-clique</i> q_i .
$Loss_i(t)$	Packet loss probability (PLP) at time t in <i>A-clique</i> q_i .
a_i	Flow acceptance ratio in <i>A-clique</i> q_i .
R_i	Flow blocking ratio in <i>A-clique</i> q_i .
b_i	Average service time of CH_i
D	End-to-end delay.
m_i	New flow request arrival rate in <i>A-clique</i> q_i .
$E[X]$	The mean of variable X .
$A_i(t)$	Number of flows that are active in <i>A-clique</i> q_i at $x * T$ minus the number of flows that terminate during $]X * T, t]$, where $X * T \leq t < (X + 1) * T$ and X is an Integer.
$New_i(t)$	The sum of the number of flows that have been generated in <i>A-clique</i> q_i and the number of flows that started transiting <i>A-clique</i> q_i during $]X * T, t]$ where $X * T \leq t < (X + 1) * T$ and x is an Integer.

$Left_i(t)$	Number of flows in <i>A-clique</i> q_i that terminate during $[X * T, t]$ where $X * T \leq t < (X + 1) * T$ and X is an Integer.
P_{loss}	PLP threshold.
Δ_{delay}	End-to-end delay threshold.
r_v	Radio transmission range for node v . $\forall u \in V$, if u is in the r_v of node v , u and v are neighbors.
R	Interference range.
CSR	Carrier Sensing Range $[\min(r_u, r_v), R]$, $\forall u, v \in V$ if u is in the CSR of node v , u and v may interfere.
$C_neighbors$	$\forall u, v \in V$, if u is in the CSR of node v , u and v are $C_neighbors$.
$D_{degree}^{(u)}$	Node degree (number of <i>A-cliques</i> to which node u belongs).
S_k	The packet generating process of an individual flow k .
$E[S]$	Average flow generation (packets/s)
P_{q_i, q_j}	The proportion of the traffic generated in <i>A-clique</i> q_i that is routed through <i>A-clique</i> q_j
$\beta_{q'q}$	The fraction of traffic generated in <i>A-clique</i> q' and going to <i>A-clique</i> q .
$INTER_i$	The waiting time incurred by interferences between <i>A-clique</i> q_i and its $C_neighbors$ <i>A-cliques</i> .
L	The packet size.
B_{req}	Bandwidth requirement.
Γ_S	The moment generating function of S_k .
Γ_P	The moment generating function of $PA_i(t)$.

3.3.4 Queuing Model

The network consists of V routers, Q_A *A-cliques*, and Q_B *B-cliques*. All two-hop neighbors transmitting on the same channel are interfering neighbors. Each node may be either a source or a destination. Assume that the packet size is L (see Section 3.4.1 for the details of our traffic model) and that a client may transfer a packet to its MR as soon as it is generated. Then, the delay between the generation of a packet and its transfer to the MR is negligible. In practice, each MR has a finite physical buffer. In our approach, we aggregate all the physical buffers in the *A-clique* into a logical buffer

at the CH. The packets are served by the CHs in a first-come–first serve manner. We propose to model WMNs as a queuing network. The stations/nodes of the queuing network represent the CHs.

Contention Matrix: The contention matrix is equivalent to the connectivity graph (see Section 3.3-1). It shows nodes that are grouped in the same A-clique, and only one of these nodes can be active at any given time. Let us define the contention matrix C for all channels as

$$C_q^u = \begin{cases} w, & \text{if node } u \in A\text{-clique } q \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where the dimension of matrix C is $Q_A \times V$.

Let us now define the contention matrix for a number of channels NC as

$$C_q^u = \sum_{w=1}^{NC} w C^w, \quad w \in \{1..NC\} \quad (2)$$

where C^w defines the unit matrix related to a given channel w.

Node Degree: The degree of node u is defined as the number of A-cliques that the node u belongs to using the same or different channels. To compute the degree, we have to sum the lines in the matrix represented in (2) for each channel w as

$$\text{For } u \in V, \quad D_{\text{degree}}^{(u)} = \sum_{w=1}^{NC} \sum_{i=1}^{Q_A} C_{i,u}^w \quad (3)$$

Fig. 17 shows that the degree of node E is 3 because E belongs to three A-cliques: $q_1/w=1$, $q_3/w=3$, and $q_4/w=1$.

Maximum Occupancy: We assume that the bandwidth requirements of the incoming flows are multiples of F . For example, if F is equal to 100 kb, and the bandwidth requirement B_{req} of flow j is 1 Mb, then B_{req} is $10F$.

The maximum occupancy MO_i of A -clique q_i is equal to the number of (unit) flows that can be accepted/supported by q_i , i.e.,

$$MO_i = \frac{C_i}{F} = X \quad (4)$$

where C_i represents the minimum available bandwidth of all nodes belonging to A -clique q_i .

If more than MO_i flows exist in A -clique q_i , then q_i is overloaded. In this case, the probability of packet loss is very high. Our proposed scheme, i.e., RCAC, rejects flows when A -cliques are overloaded or the PLP is higher than P_{loss} .

Thus, the first step in RCAC is to evaluate the maximum occupancy in each A -clique so that the bandwidth requirement B_{req} does not exceed the available resources within the A -clique. Since each node has different channel views, the maximum occupancy is not simply a local concept. To demonstrate this relationship, we illustrate a scenario with six stations (Y, D, X, D₁, Z, and D₂), as shown in Fig. 18. The MAC layer protocol is IEEE 802.11 with radio transmission ranges of 150 m for Y, 250 m for X, 200 m for Z, and $R = 550$ m. The bandwidth of the wireless channel θ_1 is 2 Mb/s. X and Z are C_neighbors. Y is X's neighbor and is out of Z's CSR. Thus, we can conclude that A -cliques q_1 and q_2 are C_neighbors.

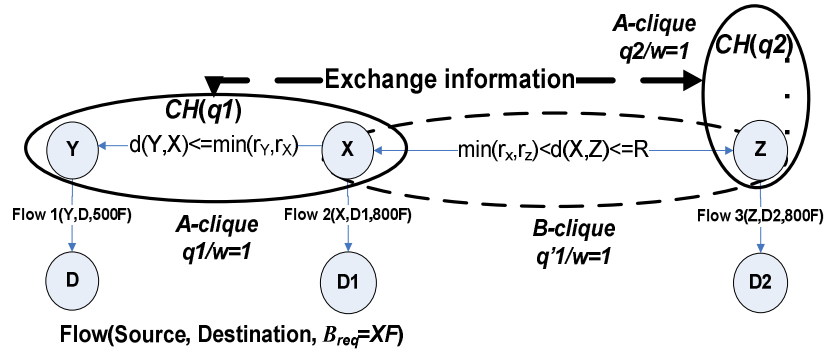


Fig. 18. Traffic scenario: an example

Table 4 shows the values of MO of the different cliques, as computed by Y, X, and Z. When flow 1 starts transmitting, only local information (the MO of q_1) is used; however, when flow 2 starts transmitting, not only the local MO is used but the neighboring MO of q_2 as well since 1) X belongs to a *B-clique* and 2) X and Z are *C_neighbors*; thus, the MOs of both *A-cliques* are reduced. If flow 3 starts transmitting, the congestion will occur (Table 4); to avoid this congestion, one has to check the availability of resources in both q_1 and q_2 before accepting flow 3. In our proposed scheme, to compute the maximum occupancy of node i , we use

$$MO_i = \min(MO_i(\text{local}), MO_j(\text{local}), \dots, MO_k(\text{local})) \quad (5)$$

where j, \dots, k are *C-neighbors* of i .

For the example shown in Fig. 18, the maximum occupancy of Z (after accepting flow 1 and flow 2 in the network) is $MO_i(Z) = \min(1200, 700) = 700$; since flow 3 requires 800, Z will simply reject it when using RCAC.

Table 4. Changes of local maximum occupancy

At time $t=t_0$, MO in q_1, q_2	Y ($MO_{q_1}=2000$)	X ($MO_{q_1}=2000$)	Z ($MO_{q_2}=2000$)
Flow 1 starts	($MO_{q_1}=1500$)	($MO_{q_1}=1500$)	($MO_{q_2}=2000$)
Flow 2 starts	($MO_{q_1}=700$)	($MO_{q_1}=700$)	($MO_{q_2}=1200$)

Flow 3 starts	$(MO_{q1}=700)$	<i>congested</i>	$(MO_{q2}=400)$
---------------	-----------------	------------------	-----------------

3.4 Proposed Analytical Model: Computing the Clique-Acceptance Ratio and End-to-End Delay

In this section, we first characterize the traffic model based on which the acceptance ratio could easily be computed. With the acceptance ratio, we derive the PLP. Finally, we determine the end-to-end delay for each path. The computed acceptance ratio and the end-to-end delay will be used in the admission control algorithm (RCAC) to accept or reject new flows into the network; a new flow will be accepted if its requirements in terms of loss rate and delay are satisfied.

3.4.1 Traffic Characterization

We denote the packet-generating process of an individual flow k as S_k , and we assume that individual packet-generating processes are independent identically distributed random variables with mean $E[S]$. Thus, the total packet-arrival rate $PA_i(t)$ in an A -clique q_i at time t is expressed as

$$PA_i(t) = \sum_{k=1}^{N_i(t)} S_k \quad (6)$$

where $N_i(t)$ denotes the number of active flows at time t in A -clique q_i .

To characterize $PA_i(t)$ as a Poisson process, we need to specify the mean of $PA_i(t)$. Using the moment generating functions of random processes $PA_i(t)$ and S_k , we obtain eq. (12) (see [GB07] for details). Let Γ_S denote the moment-generating function of S_k , i.e., $\Gamma_S(\theta) = E[e^{\theta S_k}]$, and Γ_P denote the moment-generating function of $PA_i(t)$. $\Gamma_P(\theta)$ can be computed as

$$\begin{aligned}
\Gamma_P(\boldsymbol{\theta})_{|N_i(t)=N} &= E[e^{\theta PA_i(t)} \mid N_i(t) = N] \\
&= E[e^{\theta \sum_{k=1}^{N_i(t)} S_k} \mid N_i(t) = N] \quad (7) \\
&= E[e^{\theta \sum_{k=1}^N S_k}] = \{\Gamma_S(\boldsymbol{\theta})\}^N
\end{aligned}$$

Therefore, we obtain

$$E[e^{\theta PA_i(t)} \mid N_i(t)] = \{\Gamma_S(\boldsymbol{\theta})\}^{N_i(t)} \quad (8)$$

Applying the mean to (8), we obtain

$$E[E[e^{\theta PA_i(t)} \mid N_i(t)]] = E[\{\Gamma_S(\boldsymbol{\theta})\}^{N_i(t)}] \quad (9)$$

Using

$$\begin{aligned}
E[X] &= \sum_x x p(x) = \sum_{xy} x p(x, y) \\
&= \sum_{xy} x p(x/y) p(y) = \sum_y \left(\sum_x x p(x/y) \right) p(y) \quad (10) \\
&= E_y[E_x[X/Y]]
\end{aligned}$$

We obtain

$$\begin{aligned}
E[e^{\theta PA_i(t)}] &= E[E[e^{\theta PA_i(t)} \mid N_i(t)]] \\
&= E[\{\Gamma_S(\boldsymbol{\theta})\}^{N_i(t)}] \quad (11)
\end{aligned}$$

Using (8) and (11), we have

$$E[PA_i(t)] = E[N_i(t)] E[S] \quad (12)$$

The number of active flows in *A-clique* q_i at time t can be expressed by the summation of the number of active flows $A_i(t)$ and new flows $new_i(t)$ (see Section 3.3.3):

$$N_i(t) = A_i(t) + New_i(t) \quad (13)$$

Therefore, in order to compute $E[PA_i(t)]$, we only need to compute $E[N_i(t)]$ since $E[S]$ is known *a priori* from the source traffic parameters (see next Section for details).

3.4.2 PLP

Assuming that the threshold loss probability is sufficiently small, we approximate the PLP by the overflow probability in each A-clique. We assume that the WMN is operating in moderate offered-load conditions so that we develop our analysis in the non saturated case. This assumption is reasonable since the RCAC objective is to prevent a high offered load from overloading the WMN and degrading the QoS of already-admitted flows. Moreover, most real traffic, either data or VoIP, exhibits ON/OFF behavior and therefore experience idle periods that help the WMN to operate in an unsaturated case but allow longer overflow periods. However, admission control such as our RCAC limits the overflow probability to very small values. On the other hand, it is known that when the overflow probability decays to zero, both measures, i.e., PLP and overflow probability, converge to the same value, and the difference becomes negligible. Therefore, we approximate the PLP, which is denoted by $Loss_i(t)$, by the overflow probability in each *A-clique* as

$$Loss_i(t) = Probability\{PA_i(t) > C_i\} \quad (14)$$

where $PA_i(t)$ denotes the total (new and cross) packet arrival rate into A-clique q_i at time t . In this paper, we formulate $PA_i(t)$ as a Poisson process. Such type of traffic behavior is expected when the network is accessed by a large number of users.

As mentioned earlier, we approximate the PLP by the overflow in each *A-clique* using (14). With the assumption of Poisson process $Loss_i(t)$ is expressed as follows:

$$Loss_i(t) = \sum_{k=C_i+1}^{\infty} e^{-\lambda t} \frac{\lambda t^k}{k!} = 1 - \sum_{k=0}^{C_i} e^{-\lambda t} \frac{\lambda t^k}{k!} \quad (15)$$

$$\lambda = E[PA_i(t)] = E[N_i(t)] E[S], \quad (16)$$

where the mean number of active flows in *A-clique* q_i at time t is given by $E[N_i(t)] = E[A_i(t)] + E[new_i(t)]$.

It is worth noting that both incoming (i.e., starting) and outgoing (i.e., terminating) flows are modeled as Poisson processes (see (17)).

In general, the packet service time in carrier sense multiple access/collision avoidance (CSMA/CA)-based IEEE 802.11 is not usually exponentially distributed due to the binary exponential backoff that is performed in random multiples of discrete time slots. However, in this paper, we assume that the packet service time is exponentially distributed for the sake of mathematical tractability. In fact, the authors in [HY04] demonstrate that the exponential distribution is a good approximation model for the IEEE 802.11 MAC layer service time.

Furthermore, recall that the packet arrival process of each admitted flow by RCAC to a given node is modeled as a Poisson, and RCAC rejects from the entire network all flows that, if accepted, cause ongoing admitted flows to experience significant packet drops. Therefore, since, in the worst case, only very few packets are dropped in a given node, the packet departure process of each admitted flow by RCAC from each node can be approximated as a Poisson process too, i.e.,

$$E[A_i(t)] = N_i(X * T) - E[left_i(t)] \quad (17)$$

where $X * T \leq t < (X + 1) * T$, and X is an integer.

In the following, we express traffic generated by local *A-clique* q_i and the transient traffic from adjacent *A-cliques*:

$$E[new_i(t)] = M_i \times \begin{bmatrix} a_1 m_1 \\ \vdots \\ a_i m_i \\ \vdots \\ a_{Q_A} m_{Q_A} \end{bmatrix} \quad (18)$$

where M_i is the i th row of the following matrix M .

$$M = \begin{bmatrix} 1 & P_{q_1 q_2} & \dots & P_{q_1 q_{Q_A}} \\ P_{q_2 q_1} & 1 & \dots & P_{q_2 q_{Q_A}} \\ \vdots & \vdots & \ddots & \vdots \\ P_{q_{Q_A} q_1} & \dots & \dots & 1 \end{bmatrix} \quad (19)$$

The dimension of the matrix M is $Q_A \times Q_A$, and it represents the proportions of the traffic generated in one *A-clique* that is routed through another, where Q_A is the total number of *A-cliques*. The value of a cell M_{ij} is equal to 1 if $i=j$, and equal to $P_{q_i q_j}$ if $i \neq j$, where $P_{q_i q_j}$ consists of the fraction of the traffic generated in *A-clique* q_i that is routed through *A-clique* q_j . To compute $P_{q_i q_j}$, we propose to use a heuristic, called Ford-Fulkerson-based Matrix Computation (FFMC).

FFMC takes as input the amount of traffic generated by each *A-clique*, the capacities of the links connecting the *A-cliques*, source *A-cliques* (i.e., *A-cliques* that generate traffic), and one destination; if there is only one gateway in the network, then the destination is that gateway; otherwise, the destination is a virtual node to which all gateways are connected. First, FFMC executes the (multisource) Ford–Fulkerson algorithm (which computes the maximum flow in a network [OL09]) on the input to compute the amount of traffic that passes through each of the links connecting the *A-cliques*. Second, it selects an *A-clique* from the set of *A-cliques*, which receives no

traffic from direct neighbors; we make the assumption that there is at least one *A-clique* that will be selected. Then, it computes the row for the selected *A-clique* in the matrix M using the output of the Ford–Fulkerson algorithm; FFMC fills the matrix using normalized values of the output of the algorithm. At the i^{th} step, we subtract the traffic originating from the *A-clique* corresponding to rows that have already been filled, and we apply the same procedure for the rest of the *A-cliques*.

Algorithm 2 presents the pseudo code of FFMC, which has two inputs $G=(S, E)$ and S_{q_i} and one output M, where 1) S represents the set of *A-cliques*, represented by their CHs, and E represents the set of edges between *A-cliques*; 1) the value/cost e_{q_j, q_k} associated with an edge connecting two cliques (q_j, q_k) is equal to the bandwidth between them; 2) S_{q_i} is the amount of traffic generated in *A-clique* q_i ; and 3) M is the traffic proportion matrix between *A-cliques*.

The variable β in Algorithm 2 is the fraction of traffic generated in *A-clique* q and going to *A-clique* q and is expressed as follows:

$$\beta_{q'q} := \frac{e'_{q'q}}{\sum \text{All outgoing traffic from } q'} \quad (20)$$

Algorithm 2. FFMC: pseudo-code

Input: $G=(S, E)$, S_{q_i}

Output: M

Variables : q_i : *A-clique* from S, C : is a set of *A-cliques*, $G':=(S, E')$, t : real, **Result:** *A-Clique*;

1 $C:=\emptyset$, $Result:=Null$, $M_{q_i, q_j} = \begin{cases} 0 & \text{if } i \neq j, i \text{ and } j \in \{1..Q_A\} \\ 1, & \text{otherwise} \end{cases}$;

2 $G':=Execute_Ford_Fulkerson_algorithm(G)$;

*/*the value/cost e'_{q_j, q_k} associated with an edge (q_j, q_k) (belonging to E') connecting 2 cliques q_j and q_k is equal to the fraction of traffic generated in q_j and transiting going to q_k .*/*

```

3 Repeat
4   Randomly select  $q_i$  from  $S$  such that  $q_i$  has no traffic coming from its neighbors;
5    $S := S - \{q_i\}$ ;
6    $C := C \cup \{q_i\}$ ;
7   Repeat
8     Result:=Choose an A-clique  $q$  in  $\bar{c}$  such that all of its incoming traffic comes
      from A-cliques in  $C$  and  $q' \in C$  and there is a flow from  $q'$  to  $q$ ;
      /* If an A-clique is not found, Result:=Null */
9     if (Result!=Null) {
10         $P_{q_i,q} := P_{q_i,q} + \beta_{q'q} \times P_{q_i,q}$ ;
11         $C := C \cup \{q\}$ ; }
12  until all A-cliques are in  $C$  or Result = Null
13  For every  $q_j$  and  $q_k$  in  $G'$  do
14     $e'_{q_j,q_k} := e'_{q_j,q_k} - S_{q_i} \times P_{q_i,q_j} t_{q_j,q_k}$ ;
      /*subtract all the traffic originating from  $q_i$  from the traffic of the graph
       $G'$ */
15   $C := \emptyset$ ;
16 until  $S = \emptyset$ ;

```

3.4.3 Computing the Acceptance Ratio

With our traffic model, the acceptance ratio a_i can be computed as follows.

1. Compute the number of active flows in the A-clique using (13).
2. Compute, $E[PA_i(t)]$ [see (12)] the mean of the packet arrival process in the A-clique using (12) with the mean $E[S]$ as a priori knowledge (see Section 3.4.3 to compute $E[N_i(t)]$)
3. Compute the PLP, $Loss_i(t)$, in each A-clique q_i .

4. Use the PLP and the predefined QoS constraint $Loss_i(t) \leq P_{Loss}$ to compute acceptance ratio a_i according to (21).

Thus, the acceptance ratio a_i for each *A-clique* q_i can be computed by solving the following equation, where the only unknown parameter is a_i , and P_{Loss} is the PLP threshold:

$$Loss_i(t) = 1 - \sum_{k=0}^{C_i} e^{-(Y E[S])t} \frac{(Y E[S])^k t^k}{k!} \leq P_{Loss} \quad (21)$$

where $Y = (N_i(X * T) - E[left_i(t)]) + (M_i \times [a_i m_i]_{1 \leq i \leq Q_A}) T$.

3.4.4 Computational Complexity of Acceptance Ratio Computing

The added performance and benefits realized by RCAC come at the cost of slightly higher complexity of the FFMC heuristic when computing the proportions of traffic routed between cliques. The complexity of the FFMC algorithm is on the order of $o(n^2)$, where $n = card(V)$ is the number of nodes. In fact, the complexity of computing the acceptance ratios mainly comes from filling the matrix M using Algorithm 2 (FFMC heuristic).

3.4.5 Computing the Delay in Each A-Clique

In this section, we present the delay analysis of the WMN model described in Section 3.3.4. Let N_{ct} denote the number of C_neighbors *A-cliques* of q_i identified by B-cliques. Before transmitting a packet, each node counts a random timer that is exponentially distributed with mean backoff duration $\frac{1}{\xi}$ [BA06].

The average service time of q_i is expressed as

$$b_i = \frac{1}{\xi} + \frac{L}{\theta_k} + INTER_i \quad (22)$$

where $INTER_i$ is the waiting time incurred by interferences between A -clique q_i and its C _neighbors A -cliques.

In the case of no interferences,

$$INTER_i = 0, \quad (23)$$

If interferences exist, we consider that all interfering A -cliques have the same probability to access the medium. In this case,

$$INTER_i = \frac{L}{\theta_k} \times \left(\frac{\sum_{j \in N_{ct}} PA_j C_j}{\sum_{j \in (N_{ct} \cup i)} PA_j C_j} \right), \quad (24)$$

where PA_j is the packet arrival rate in A -clique q_j .

The end-to-end delay for each path is determined by computing delay at each intermediate CH as

$$D = \sum_{i \in PATH} b_i \quad (25)$$

where $PATH$ is the set of the nodes in the end-to-end path from the user to the destination.

The delay parameter is the second QoS metric considered in our model. Before accepting a flow, RCAC checks whether the delay of the selected route is lower than the threshold value (Δ_{delay}), which is predefined according to the requirement of different traffic types. This constraint is expressed as follows:

$$D = \sum_{i \in PATH} b_i < \Delta_{delay} \quad (26)$$

3.5 Routing on Cliques Admission Control Algorithm

In this section, we present the details of the operation of RCAC; more specifically, we describe the operations and interactions among CHs in processing flow requests.

When a CH receives a new flow request from a node (called user) in the A-clique (in this case, it is called source CH), it determines 1) whether there is sufficient bandwidth to accommodate the flow (by computing the maximum occupancy [see (5)]) and to satisfy the loss probability threshold (by computing the acceptance ratio [see (21)]) and 2) whether the CH's delay [computed using (25)] is smaller than the delay threshold. If the response is yes, then the source CH sends the request, including the CH's delay, to the next-hop CH toward the destination CH; otherwise, a rejection is sent back to the user. Upon receipt of the request, a transit CH computes the sum of its delay and the delay included in the request; if the accumulated delay is smaller than the delay threshold, then the CH determines whether there is sufficient bandwidth to accommodate the flow and to satisfy the loss threshold. If the response is yes, then it propagates the request, including the accumulated delay to the next-hop CH. Each transit CH repeats this step. When the destination CH receives the request, it carries out a similar process; if the response is yes, then it sends "commit" toward the source CH [see Fig. 19(a)]. A timer is used by the source CH; if the timer expires before receiving the corresponding "commit", then the source CH assumes that the flow request cannot be accommodated and notifies the user [see Fig. 19(b)].

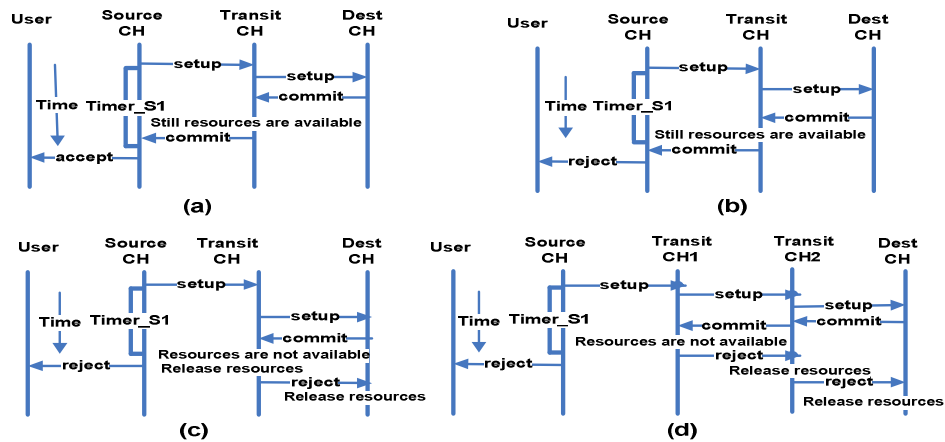


Fig. 19. Inter-CH interactions

The CHs involved in the flow request setup "conditionally" commit the network resources for the request only after receiving "commit". Indeed, this occurs in the second phase of the setup process, where each CH that receives "commit" forwards it to its previous-hop CH, toward the source CH, if resources are still available to accommodate the request; otherwise, it sends "reject" toward the destination [see Fig. 19(c)]. When the source CH receives "commit", it sends "accept" to the user who can start sending his/her traffic toward the destination. Upon receipt of "reject" a (transit or destination) CH releases the resources conditionally committed and forwards "reject" to the next hop toward the destination [see Fig. 19(d)]. RCAC, via CHs, supports concurrent processing of setup requests. It realizes concurrence control using an optimistic scheme. Indeed, it assumes that it is very unlikely that concurrent flow requests contend for the same resources at the same time, i.e., the success of one request causes failure of the other. This is the reason why, during the first phase, no resources are committed. The optimistic scheme allows CHs to accept more requests than other schemes (e.g., pessimistic scheme) and commit resources only when a request can be accommodated by all CHs involved in the flow request processing. Indeed, the optimistic scheme makes "optimal" the use of the network resources. Let us consider the following example: A transit CH receives a request req_1 , decides that it can accommodate the request, commits the corresponding resources, and forward the request to the next-hop CH. Then, it receives a request req_2 (with the same

characteristics as in req_1), decides that it cannot accommodate the request, and sends "reject" toward the source CH. Later on, it receives "reject" that corresponds to req_1 because one of the CHs toward the destination determined that it cannot accommodate req_1 . In this case, the two requests are rejected. If the optimistic approach has been used, then req_2 would have been accommodated (assuming that all CHs involved in processing req_2 were able to accommodate the request).

Algorithm 3 presents the pseudo code of the operation of RCAC in processing a flow request. More specifically, it presents the operation of the source CH, the transit CH, and the destination CH.

Algorithm 3. Admission control algorithm on A-clique head

Input: Flow request($user, D, B_{req}$) in q_j where $user$ is the source node, D is the destination node, B_{req} is the required bandwidth

Output: Admission decision: accept or reject

Variables : timer_S1;

At Source CH :

- 1 Receive the flow request from user with QoS requirement
 - 2 **if** (Local_Resource_Admission_Test ()==accept) {
 - 3 Forward setup request to the next hop CH towards D;
 - 4 Start up timer_S1;
 - 5 **while** (timer_S1 is not expired && commit message not received) **do** {Wait;}
 - 6 **if** (timer_S1 expired) send "reject" message to the user;
 - 7 **else if** (commit received) send "accept" message to the user;
 - 8 } **else** Send "reject" message to the user;
-

At Transit CH :

- 9 **if** (Local_Resource_Admission_Test ()==accept) {
- 10 Propagate the request with the pending admission control parameters to the next hop CH;
- 11 **if** (commit received && resources are still available) {
- 12 Commit "conditionally" resources;
- 13 Forward "commit" message towards the source CH;

```

14 } if (reject message received) {
15   Release resources;
16   Send “reject” message towards destination CH; }
17 }else Send “reject” message towards the source CH;

```

At Destination CH :

```

18 if ( Local_Resource_Admission_Test ()==accept) {
19   commit “conditionally” resources;
20   send “commit” message towards the source CH;
21   if (reject message received) Release resources;
22 }else Send “reject” message towards the source CH;

```

Function : Local_Resource_Admission_Test ()

```

23 Compute local  $MO_j$  (Equation (4));
24 Exchange information with C_neighboring cliques about  $MO$  to compute
   new  $MO_j$  (Equation (5));
25 Compute the acceptance ratio  $a_j$  by solving (Equation (21)) ;
26 if ( $N_j(t) < MO_j$  or  $\text{RANDOM\_UNIFORM}(0,1) > a_j$ ) return (reject);
27 Compute the average service time  $b_j$  (Equation (22));
28 if (the delay (D) from the source to this  $q_j$  is bigger than the threshold
   value  $\Delta_{delay}$  (Equation (26)) return (reject);
29 return (accept);

```

A CH needs a routing protocol to determine the next-hop CH toward the destination. In our simulations, we used a modified version of destination-sequenced distance-vector routing (DSDV) [PB94] called MyDSDV as the routing protocol to realize RCAC. MyDSDV is based on delay instead of the distance metric used in the original DSDV. It is used by a source or transit CH that receives a flow request to determine the next-hop CH toward the destination; a CH selects the next-hop CH with the shortest delay. The rationale behind our choice is that DSDV uses the Bellman–Ford algorithm, which is often preferred for mesh networks or sensor static wireless networks. A previous study in [YJ08] shows that the Bellman–Ford algorithm has a

lower control message overhead compared with the flooding based route discovery used in ad hoc on-demand distance vector routing (AODV) [PB03]. This being said, RCAC can be used with other routing protocols for WMNs.

3.6 Simulation Results

In this section, we conduct a simulation study using ns-2 [NS] to evaluate and compare the performance of our proposed scheme, i.e., RCAC, with other existing schemes. We evaluate several performance metrics: 1) the end-to-end delay; 2) the throughput; 3) the packet loss and outage probabilities; 4) the flow blocking probability; and 5) the overhead. Note that the outage probability is defined as the ratio of the number of flows experiencing packet losses higher than the given threshold to the total number of accepted flows.

3.6.1 Simulation Configurations

In the first scenario (1), the WMN topology used in simulations is arranged as a regular grid of 5×5 IEEE 802.11 stations acting as MRs. More specifically, node spacing is 100 m, and the position of each node from the regular grid is perturbed by choosing a random angle uniformly in $[0, 2\pi]$ and a radius uniformly in $[0 \text{ m}, 25 \text{ m}]$. This perturbation is used in several existing WMN performance studies (e.g., [CL08]).

In the second scenario (2), the positions of MRs are uniformly distributed in a $1000 \text{ m} \times 1000 \text{ m}$ coverage area. The radio transmission range r takes the value of 150, 200, or 250 m, and the transmission interference R of each wireless station is 550 m.

Real-time traffic flows arrive at each wireless station according to a Poisson process; each flow generates on the average $E[S] = 20$ packets/s. We set a constraint on PLP not exceeding a given threshold P_{loss} and a constraint on delay not exceeding Δ_{delay} ; thus, RCAC accepts new flows only if these constraints are satisfied. The other parameters are presented in Table 5. It is worth noting that, in the simulation results, we are using WLAN IEEE 802.11 CSMA/CA for RCAC and without RCAC and WLAN IEEE 802.11 MDA for existing schemes in the related work to which we compare.

Table 5. Simulation parameters

P_{loss}	5%	Routing protocol for W.O.RCAC	DSDV
Δ_{delay} : scenario (1)/scenario (2)	50 ms/5ms	DTIM/ DTIM utilization threshold	32 ms/70%
r	250 meters	Channel assignment	Static channel assignment [BH08, BH07]
R	550 meters	$E[S]$	20 packets/sec
T	5 sec	Routing protocol for MDA/Dynamic relocation	AODV
L: Packet size	1000 bytes	Routing protocol for RCAC	MyDSDV
Traffic type	CBR	m_i	20 to 200 flows/ sec
Rate: scenario (1)/scenario (2)	1 Mbps/10Kbps	C_i	200 F, F=10 kb/sec
NR: scenario (1)/scenario (2)	1/2 radios	SIFS	10 μ s
NC: scenario (1)/scenario (2)	1/12 channels	Backoff Slot time	20 μ s
WithOut RCAC : W.O.RCAC	$a_i=1$, no delay and loss constraints	DIFS	50 μ s
Topology : scenario (1)/scenario (2)	Grid 5X5 / Uniform-distributed topology	Data rate	11 Mbps
Node spacing : scenario (1)	100 m	MAC Header	192 bits
Threshold (outage)	5%	Ack Frame Size	112 bits
Link Tx rate: scenario	18 Mbps	PLCP Preamble and Header	192 bits

(1)/scenario (2)	/1Mbps	Length	
PHY radio model	SINR-based [CL08]	Time slot	20 μ s

3.6.2 Results Analysis

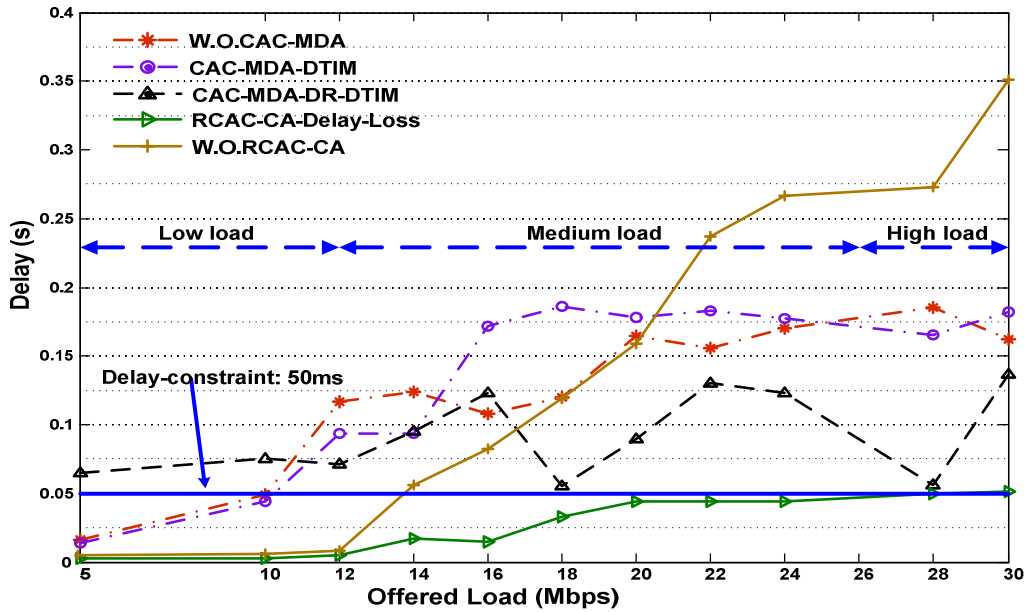


Fig. 20. Delay variation with MDA versus CA

MAC Access Method Analysis—CA versus MDA Using Scenario (1): As presented in Section 3.2, the IEEE 802.11 draft 2.0 allows an optional contention-free MDA method besides the well-known CSMA/CA contention-based access (CA) method. We studied the performance of both of these access methods when transmitting data on the channel of each clique. In Fig. 20, we present the average end-to-end delay experienced by network flows when using the various simulated schemes; in this section, we analyze the performance of the two MAC access methods without admission control, i.e., Without RCAC (W.O.RCAC)-CA and W.O.CAC-MDA schemes in Fig. 20. The average end-to-end delay is the sum of the access delays (both queuing delay, which is negligible in low load, and contention delay) experienced in hops along the path's flow from the MC to the gateway.

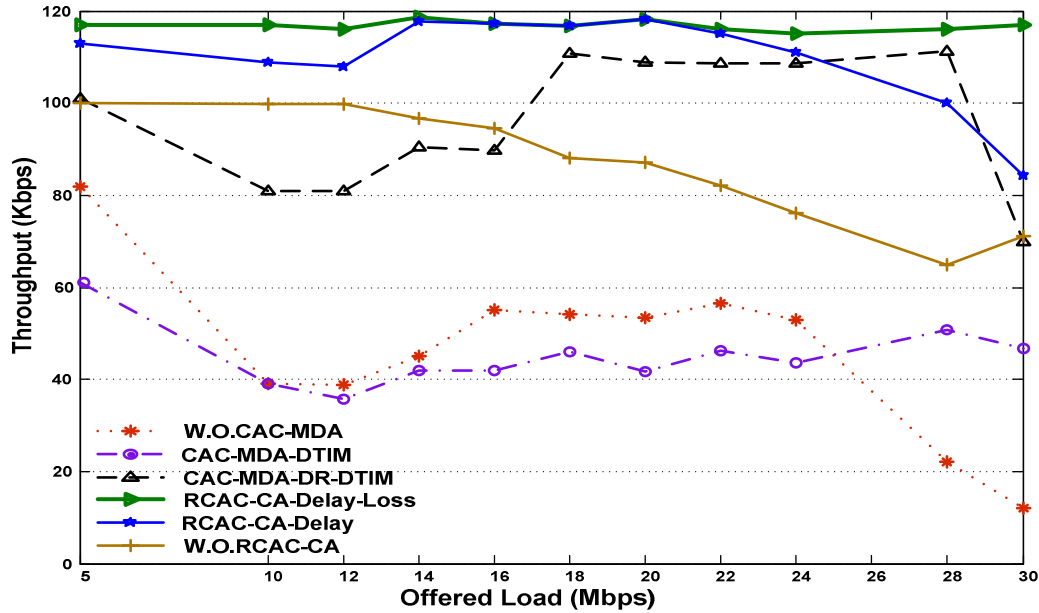


Fig. 21. Throughput variation with MDA versus CA

In a low-offered-load condition, where collisions are very rare, the CA method provides lower delays since it transmits almost instantaneously in a random time slot no later than 0.64 ms, which corresponds to a minimal contention window of $CW = 32$ time slots of 20 μ s. The CW size is initially set to 8, 16, or 32 but is doubled each time a node experiences a collision, until it reaches a maximum value of 1024. However, in this same light-offered-load condition, the MDA method waits longer periods before transmitting in specific reserved contiguous time slots, i.e., particularly MDAOPs, that can be scheduled to start as late as the end of a DTIM period of 32 ms. This scheduling is performed regardless of the absence of interferences and although earlier time slots are available, since it needs contiguous available time slots to transmit packets. Therefore, the average access delay is higher with MDA compared with CA when the offered load is low. On the other hand, in a high-offered-load condition, the end-to-end delay with MDA does not exceed 224 ms; it is bounded by the DTIM interval, which is equal to 32 ms, multiplied by the maximum number of hops in a path that is equal to 7 in our topology. Whereas the delay provided by CA increases without any bounds with the increase of the offered load, it results in many more nodes contending for the same channel, causing many more collisions and resulting in both longer binary exponential

backoffs and more frequent MAC retransmissions. Note that binary exponential backoffs and retransmissions are not used by the MDA; this results in much higher packet losses (see Fig. 22) for MDA compared with CA.

In Fig. 21, we compare the average throughput provided by the two access methods. We measure the throughput as the total number of bits that are correctly received for each flow in the WMN in the unit of time. Fig. 21 shows a much lower throughput provided by the MDA compared with CA. MDA throughput degrades much more in a high-offered-load condition. This is explained by the MDAOP reservation mechanism that automatically rejects the newly arrived MDAOP setup requests. This is caused by the contiguous time slots in the DTIM slotted interval that become less available when the offered load increases. This can be verified in Fig. 22 and Fig. 23 from the increase in the packet losses and the outage probabilities that are much higher for MDA compared with CA.

As opposed to MDA that behaves like an admission control since it bounds the access delay within a fixed DTIM interval, the CA method needs to be enforced with an admission control mechanism to limit its access delay. This admission control is done by efficiently restricting the number of flows contending for the same channel in a given clique, depending on several parameters. We present the performance of our proposed admission control scheme RCAC and compare it with other existing schemes in the next section.

Admission Control Analysis Using Scenario (1): We study the performance of two variants of our proposed RCAC: one based on a delay constraint only, namely, RCAC-CA-Delay, and the other based on both delay and loss constraints, namely, RCAC-CA-Delay-Loss. For the related work, since regular MDA, namely, W.O.CAC-MDA, does not provide a flexible constraint on access delay beyond the DTIM interval, we consider an MDA-based scheme called CAC-MDA-DTIM. This scheme acceptance parameter is based on a constraint on DTIM utilization that does not exceed a given threshold. The DTIM utilization threshold is the maximum fraction of the period of time that can be used by MDAOPs. We present the performance of CAC-MDA-DTIM

combined with another MDA improvement that considers the DR [CL08] of the reserved MDAOPs two hops away; we call this scheme CAC-MDA-DR-DTIM.

In Fig. 20, we notice that the average end-to-end delay experienced by the CA enforced by our proposed admission control RCAC-CA-Delay-Loss is bounded by a much lower delay, i.e., at least three times lower starting from an offered load of 16 Mb/s, compared with that of the MDA with admission control [So07]. This is explained by the lower access delay of CA when we limit the number of contending nodes with our RCAC-CA-Delay-Loss compared with that of MDA; the latter generally waits to transmit in specific reserved time slots, even though it is bounded with a DTIM utilization threshold.

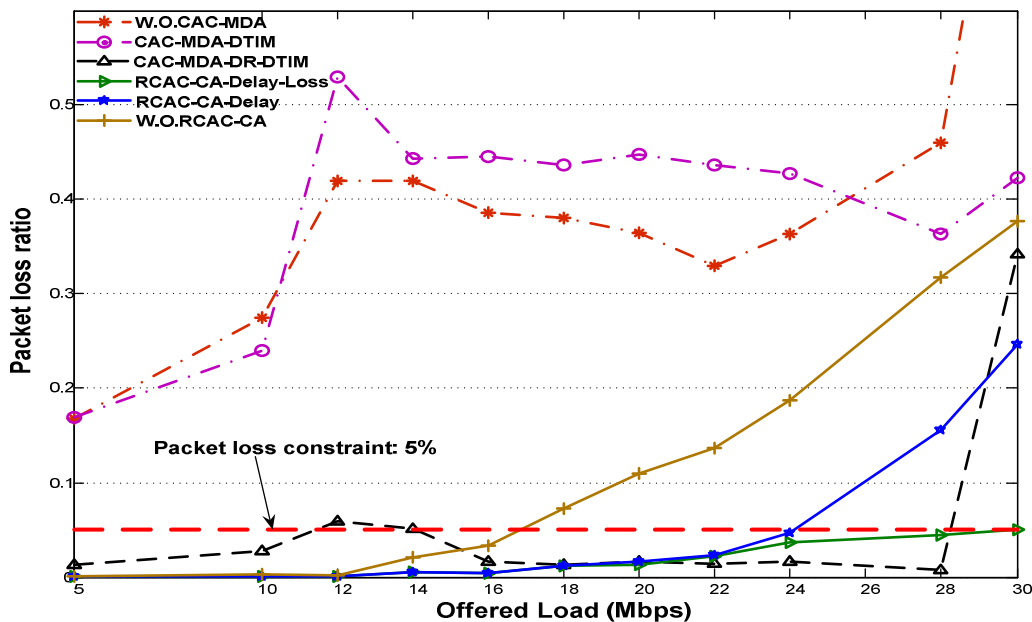


Fig. 22. Loss variation with MDA versus CA

In Fig. 22, we observe that MDA-based schemes, either with or without a DTIM utilization constraint, experience much higher packet losses compared with all the other schemes, even in low-offered-load conditions. This seems counterintuitive since MDA is supposed to, by means of MDAOP reservations, provide far fewer collisions/interferences and, therefore, fewer packet losses. However, as presented in [CL08], the MDAOP reservation process usually suffers from the DTIM fragmentation

problem, which can occur even in light-offered-load conditions. In fact, it is usually difficult to find contiguous time slots for reserving long MDAOPs for long packets. Therefore, MDA does not accept MDAOP reservations, and long packets are rejected. In fact, we tested the performance of MDA using small packet sizes of less than 160 B, and we found near zero packet losses comparable with that provided by the CA. Moreover, interferences outside the two-hop neighborhood are not prevented by MDAOP reservations. These issues are tackled using the DR of interfering MDAOPs; indeed, CAC-MDA-DR-DTIM keeps the packet loss to a near-zero value that becomes significant only after the offered load exceeds 28 Mb/s.

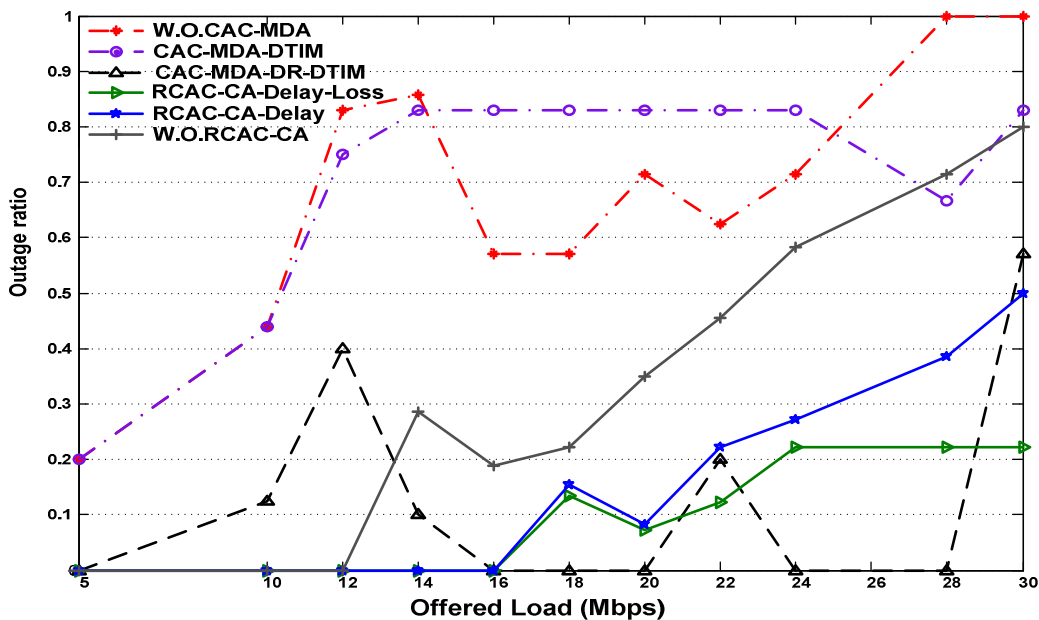


Fig. 23. Outage variation with MDA versus CA

In Fig. 22, we also notice that our proposed schemes RCAC-CA-Delay and RCAC-CA-Delay-Loss provide quite comparable packet losses under low- and medium-offered-load conditions. However, only RCAC-CA-Delay-Loss outperforms all the other schemes by providing a low and bounded packet loss for all offered-load conditions, even in high-offered-load conditions. In Fig. 21, the throughput variation with the offered load is tightly related to the packet loss variation in Fig. 22. Fig. 21 shows that our RCAC-CA-Delay-Loss provides the highest throughput compared with all the other schemes since it limits the number of contending flows by adjusting the

acceptance ratio a_i , depending on the offered load, so that the packet loss does not exceed the threshold P_{loss} .

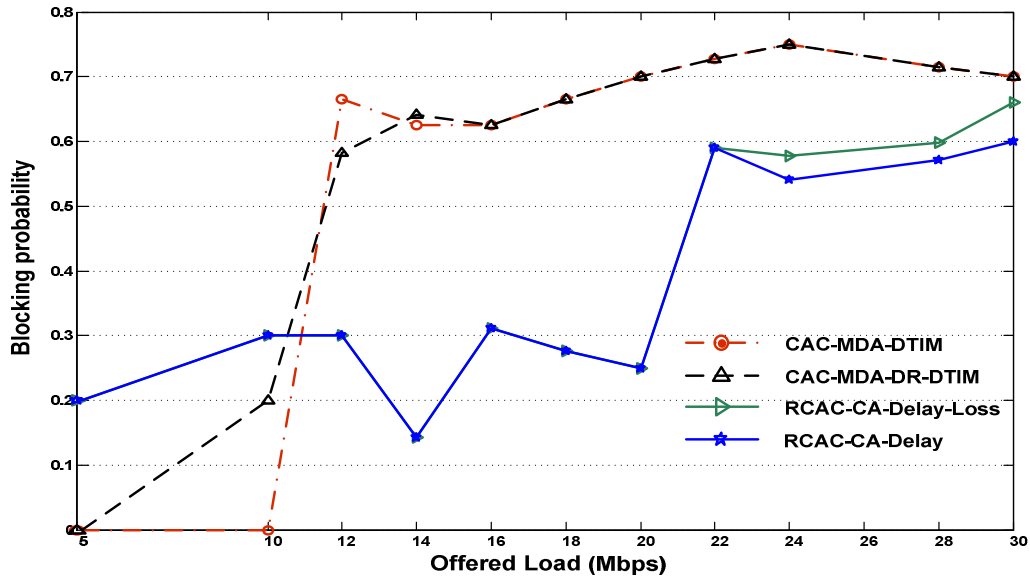


Fig. 24. Blocking variation with MDA versus CA

Fig. 24 shows the blocking probability when varying the offered load. Note that the blocking probability is expressed as

$$PR_i = \frac{\sum_{i=1}^{Q_A} m_i R_i}{\sum_{i=1}^{Q_A} m_i} \quad (27)$$

where $R_i = 1 - a_i$ and m_i is the new flow arrival rate in an A-clique q_i .

From Fig. 22, we notice that the packet loss threshold P_{loss} is reached starting from an offered load of 24 Mb/s since the blocking probability in Fig. 24 experienced by RCAC-CA-Delay-Loss becomes greater than that experienced by RCAC-CA-Delay. In fact, to guarantee a packet loss not exceeding the given threshold, our proposed RCAC-CA-Delay-Loss has to block only a little more flows compared with RCAC-CA-Delay.

One can argue that we do not need explicit admission control in WMNs if we use a deterministic access method such as MDA, since the latter guarantees delay

bounds required by real time services. However, our performance study shows that this access method, although it bounds delays, suffers a lot from packet losses and blocking. Some of the MDA improvements presented in the literature may significantly reduce the packet loss, but the blocking probability is still higher than that of our proposed RCAC. In fact, simulation results show that the regular CSMA/CA enforced with our delay-constrained and packet-loss-constrained admission control, i.e., RCAC-CA-Delay-Loss, is able to provide a more efficient utilization of the clique's radio channel since it provides a lower flow blocking probability compared with MDA-based schemes while guaranteeing both delay and packet losses, even in high-offered-load conditions.

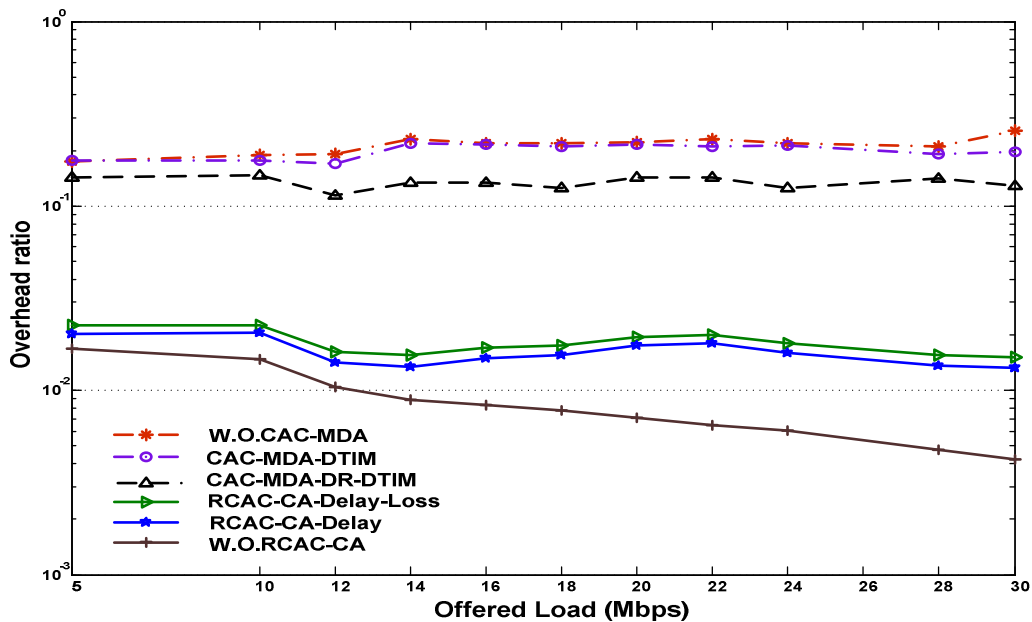


Fig. 25. Overhead variation with MDA versus CA

Fig. 25 shows the overhead generated by schemes based, respectively, on RCAC-CA and on MDA when varying the offered load. The overhead is computed as

$$Overhead = \frac{\text{Number of control packets}}{\text{Number of control packets} + \text{Number of Data}} \quad (28)$$

where the number of control packets for MDA-based schemes includes AODV overhead packets, MDA advertisements, MDA setup requests, etc. For the RCAC

schemes, the overhead of MyDSDV includes the DSDV overhead and inter-CH signaling (see Fig. 19). Without RCAC, only the DSDV overhead is considered.

In our case, we observe that RCAC-CA generates a bigger overhead compared with W.O.RCAC-CA. Indeed, the gains of RCAC-CA in terms of delay, throughput, PLP, and blocking come at the cost of a bigger overhead compared with the case without RCAC-CA. However, RCAC-CA still generates a smaller overhead compared with MDA-based schemes.

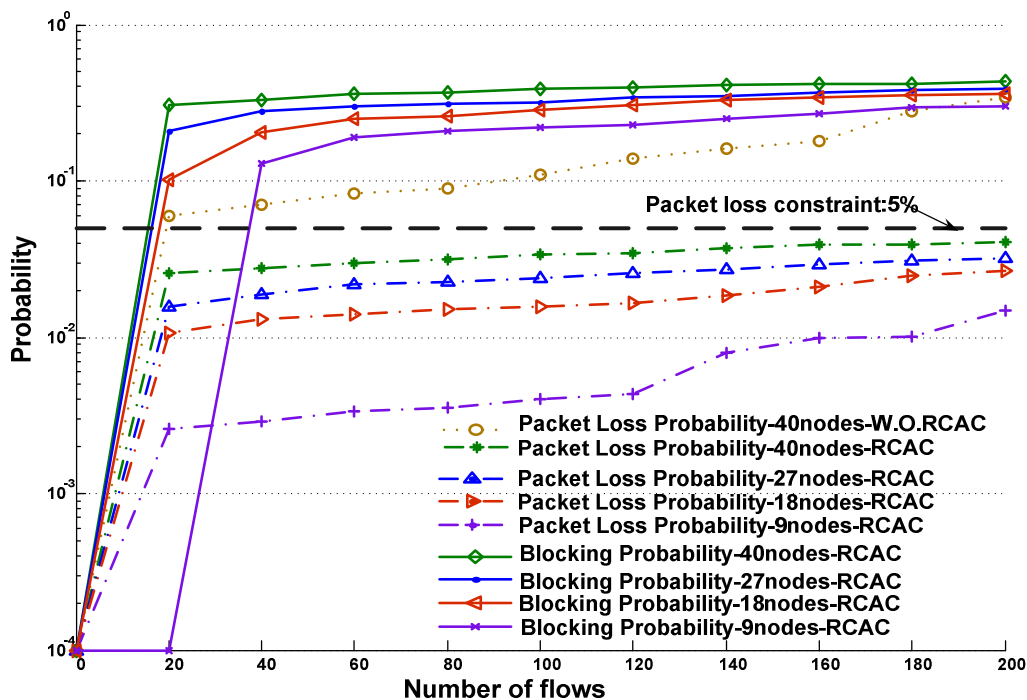


Fig. 26. Blocking, packet loss probability of RCAC versus W.O.RCAC for different network sizes

Admission Control Scalability Using Scenario (2): Specifically, Fig. 26 shows that, in the case of a nine-node network, RCAC rejects flows starting from the 40th flow, whereas in the case of 18-node, 27-node, and 40-node networks, RCAC rejects flows starting from the 20th flow. This can be explained by the fact that, in the case of 18, 27, and 40 nodes, interferences are more severe than in the case of nine nodes (we use the same geographic area size and random topologies); furthermore, we use manual

(ad hoc) channel assignment. Thus, different results may be achieved using different channel-assignment schemes (e.g., optimal schemes [BH08], [BH07]).

Fig. 26 shows the variation of PLP with RCAC and W.O.RCAC while varying the number of flows in a network of 40 nodes. W.O.RCAC corresponds to the basic scheme that always accepts and routes flows from sources to destinations using DSDV [PB94]. Fig. 26 illustrates that, even when the traffic increases, the PLP does not exceed 5% using RCAC (MyDSDV). However, without RCAC (W.O.RCAC), the PLP threshold is exceeded when the number of flows in the network approaches or exceeds 20. Up to this point, RCAC did not start blocking flows. When RCAC starts blocking flows, the PLP is more than three times higher without RCAC than with RCAC when the number of flows in the network approaches 60. The PLP keeps increasing with the number of flows if RCAC is not used. Furthermore, Fig. 26 illustrates the PLP variation versus the number of flows in WMN for different network sizes of nine, 18, 27, and 40 nodes. We observe that when both the traffic load and the network size increase, RCAC using MyDSDV maintains a PLP under the 5% threshold. These results confirm our analytical findings: RCAC controls the PLP in the network. In summary, the simulation results confirm that RCAC can efficiently satisfy the packet loss and the end-to-end delay requirements for real-time traffic in WMNs while providing better utilization of the network resources and not incurring extra overhead (smaller overhead compared with MDA-based schemes). Moreover, we compared the performance results of RCAC using CSMA/CA against an admission control using standard MDA and another admission control using an improved version of MDA [CL08]. We conclude that our scheme RCAC outperforms existing schemes based on MDA in terms of delay, throughput, packet loss, and blocking probabilities.

3.7 Conclusions and Future Work

In this paper, we have proposed a new distributed admission control scheme based on cliques, namely, RCAC, to support real-time services in WMNs. Particularly, we have considered PLP and end-to-end delay as two major criteria in the design. Our admission control algorithm is carried out by each CH in parallel; thus, computation

efficiency and scalability are achieved. Simulations show that RCAC can effectively prevent the PLP and packet delay in WMNs from exceeding predefined thresholds. Furthermore, we have concluded from our performance study that the regular CSMA/CA enforced with our delay constrained and packet-loss-constrained admission control, i.e., RCAC-CA-Delay-Loss, is able to provide a more efficient utilization of the clique's radio channel since it provides a lower flow blocking probability compared with MDA-based schemes while guaranteeing both delay and packet losses, even in high offered-load conditions.

In a future work, we plan to add more sophisticated admission control for MDA to minimize the blocking probability while satisfying a constraint on packet loss not exceeding a given threshold. Recall that, in our study, we adopted the approach of static channel assignment, which is widely used in the research community. We will investigate the impact/adaptation of RCAC using dynamic channel assignment. Furthermore, we will extend RCAC to handle different types of traffic. Indeed, we will study the scenario of multiple traffic classes with different QoS requirements.

Chapitre 4: Admission Control and QoS Provisioning for IEEE 802.11s Mesh Deterministic Access

Jihene Rezgui, Abdelhakim Hafid, Michel Gendreau

Abstract

In this paper, we design an efficient scheduler for Mesh Deterministic Access (MDA) in IEEE 802.11s-based Wireless Mesh Networks (WMNs), called Enhanced Differentiated MDA (ED-MDA) to support voice and video applications with strict requirements on delay and on blocking/dropping probability. ED-MDA is combined with an Efficient Admission Control algorithm, called EAC, that guarantees QoS for different service classes and provides higher capacity in WMNs. EAC processes both new and HandOver (HO) calls to maintain a balance between two conflicting requirements: maximizing the resource utilization and minimizing the dropping rate. To establish priority between new and HO calls, especially to avoid the forced termination of an ongoing call, ED-MDA coupled with EAC reserves the minimum amount of necessary resources while maintaining an acceptable HO call dropping rate and high resource utilization. Particularly, our model provides an efficient adaptive adjustment of the Contention Free Period (CFP) duration to make efficient use of the scarce wireless resources while supporting different services with different QoS requirements, such as delay. Simulations show that ED-MDA together with EAC outperforms existing schemes in terms of delay and loss rate. Particularly, simulation results show the effectiveness of ED-MDA in decreasing the HO call dropping rate, especially when considering guard intervals.

Keywords: Admission control, handoff, network management, service differentiation, wireless mesh network.

Status: This paper is submitted to Journal of Computer Networks [RH10e]. The ideas presented in this paper are largely based on the following accepted paper:

- Admission Control and QoS Provisioning in Multi-Service MDA for IEEE 802.11s-based Wireless Mesh Networks, ICST Qshine, Texas, USA, 2010 [RH10h].

4.1 Introduction

Wireless Mesh Networks (WMNs) have recently emerged as a promising technology for the next-generation wireless networks [AW05]. A WMN consists of two types of nodes: Mesh Clients (MCs) and Mesh Routers (MRs). The MRs form a wireless mesh backbone infrastructure that forwards most of the traffic between MCs and Internet gateways. In general, MRs have very little mobility and operate just like stationary routers, except for being connected by wireless links using the very popular IEEE 802.11 WLAN standard or other wireless technologies, such as WiMax. Using more than one radio interface in each MR allows better channel diversity resulting in less interference and therefore more throughput and capacity. However, this only improves the performance for Best Effort traffic [EG08] since supporting Quality of Service (QoS) for real-time traffic in WMNs remains an open challenge. For example, real-time connections (RTC) require low delays and low packet losses, which are difficult to guarantee because of the random contention used by the traditional CSMA/CA MAC in IEEE 802.11-based WMNs. Nevertheless, some studies solve this problem by enforcing contention-based MAC with complex admission control schemes [RH10d].

In this paper, we are concerned with QoS provisioning as well as the mobility management in multi-service IEEE 802.11s WMNs. Thus, the allocation and efficient use of limited wireless resources are to be studied. The challenge is, to support multimedia applications in order to meet users's expectations in terms of delay and packet losses while maintaining a reasonable high use of radio resources. Generally, WMNs consist of MRs whose small transmission range may result in frequent handoffs of mobile MCs. This inherent property of the mobile MC may often lead to high packet delays and loss rates [RH09]. Thus, a suitable admission control protocol must be

designed to avoid the forced termination of an ongoing call so that a mobile MC could be able to freely move across the WMN while maintaining its ongoing communications uninterrupted. In multiple radios WMNs, it has been proved that there are still channels that may interfere even if the highest quality channel assignment heuristics [RG04] are adopted. The concept of clique [RH10d] is used in this paper to identify the conflicting links that need to be carefully managed to limit contentions. Recall that a clique is defined as a subset of logical links that interfere with each other.

Mesh Deterministic Access (MDA) aims to provide stringent medium control access (MAC) delay guarantees for real-time services such as voice over IP (VoIP) which is a condition that can hardly be satisfied in classical IEEE 802.11 standard. The MDA scheme studied in [HM07], [HM08], [CL08] and [GD09] extends the typical IEEE 802.11 medium instantaneous reservation procedure, also known as the virtual carrier sensing (V-CS), to a more advanced reservation procedure using scheduled MDA Opportunities (MDAOPs) within two-hop neighborhood. MDAOPs are first negotiated between neighboring mesh nodes by exchanging broadcast setup messages, then, MDAOPs reservations are performed in multiples of time-slot unit, during the Delivery Traffic Indication Message (DTIM) periodic interval. To limit the message broadcast signalling overhead, MDA-related messages are sent only within two-hop neighborhood. IEEE 802.11 V-CS [HM07] is performed by a four way handshake procedure in which request-to-send (RTS), clear-to-send (CTS), data and acknowledgment packets are exchanged between two communicating nodes, while a network-allocation vector (NAV) is set by the other nodes in the physical sensing range (P-CS). V-CS works flawlessly in single hop wireless networks but can cause severe interference in wireless links that are multiple hops away and share the same channel within an overlapped transmission or interference range. Therefore, in multi-hop wireless networks, such as WMNs, MDA, with less multi-hop interference, was adopted in the IEEE 802.11s draft amendment [HM07] to provide stringent delay bounds. Although, this adopted scheme of MDA has reduced, to a certain extent, the delay bounds it lacks the concept of differentiating frames with different priorities. Basically, MDA is supposed to provide a channel access with equal probabilities for all

stations contending for the channel deterministic access in a distributed manner. However, equal access probabilities are not desirable among stations with different priority frames.

Our contributions, in this paper, can be summarized as follows: (1) We propose a novel scheduler algorithm for per-flow MDA, called ED-MDA, which takes into account traffic differentiation; (2) We elaborate an adaptive adjustment of CFP (Contention Free Period) for each service class by prioritizing RTC in order to provide QoS guarantees in terms of delay and blocking probability that does not exceed pre-defined thresholds (i.e., for voice :150 ms delay and 6% blocking and for video 300 ms delay and 12% blocking resp.); (3) We propose a new admission control algorithm, called EAC, in conjunction with ED-MDA. EAC reserves the minimum amount of necessary resources (i.e., time-slots) to maintain an acceptable HO call dropping rate and provide high resource utilization; it also establishes priority between new and HO calls to avoid interrupted communication for highly mobile users; and (4) We present a centralized static algorithm (Algorithm 4) which searches for an initial feasible Upper Bound of the number of Time-Slots for each service class based on Erlang-b formula [KL75], called UBTS. Thus, starting with this good initial solution rather than a randomly generated one, ED-MDA coupled with EAC provides a better final adjustment of CFP/time-slots while improving the initial solution according to the state of the blocking probability constraint (Section 4.5).

The remainder of the paper is organized as follows. Section 4.2 provides a brief overview of related research. Section 4.3 presents our notations, assumptions, and network model. Section 4.4 proposes an analytical formulation of the differentiated MDA (ED-MDA). Section 4.5 describes our admission control algorithm EAC. Section 4.6 evaluates the proposed solution via simulations. Finally, Section 4.7 concludes this paper.

4.2 Related work

To the best of our knowledge, there exists only few contributions [HM07], [HM08], [CL08] and [GD09] focusing on the description and evaluation of MDA

proposed in IEEE 802.11s. Hiertz et al. [HM07] studied the MAC in IEEE 802.11s by first describing the mandatory Enhanced Distributed Channel Access (EDCA) and then the optional feature MDA. In opposition to EDCA, MDA is particularly designed to support multi-hop WMNs and is based on scheduled medium access. In addition, we note that QoS support in IEEE 802.11s is not sufficient to provide lower delays and losses required by multimedia applications. This being said, there exists a number of contributions [NG06], [WK06] and [BX10] that propose techniques, in the context of IEEE 802.11 WMN, to improve voice QoS and network capacity. However, these contributions neither take into account real and no real time traffic at the same time nor differentiate between new and HO calls.

4.2.1 Problems with EDCA QoS Guarantees in IEEE 802.11s

In contrast to Distributed Coordination Function (DCF), EDCA introduces features to support QoS; it describes eight traffic categories which are mapped to four access categories (ACs): voice, video, Best Effort and Background. Each AC has a specific probability that is the fraction of time allowed to access the wireless medium. Similar to DCF, EDCA does not support medium reservation, which makes it difficult to guarantee any QoS criteria. One can argue that IEEE 802.11e features, such as Hybrid Coordination Function (HCF) Controlled Channel Access (HCCA) and Point Coordination Function (PCF), can provide QoS guarantees. We believe this argument does not have a strong foundation for two reasons (1) HCCA and PCF are not implemented yet [HM07] by the current products; and (2) HCCA and PCF are not practical/realistic in WMNs since a WMN has a logical flat topology without a dedicated centrally coordinating device.

The shortcoming of EDCA can be summarized as follows: (1) EDCA has been designed for single-hop communications; (2) EDCA is considered as a prioritization mechanism which makes it hard to perform well in multi-hop mesh environment; (3) EDCA devices have no means of cooperation. Devices that are at the edge of the WMN have different views of the status of the wireless medium than devices in the core of the WMN. The edge devices detect the wireless medium when it is idle more often than

devices in the core of the WMN; and (4) EDCA is less efficient when confronted with high load because of unsuccessful transmissions observed by EDCA devices which involve increasing the size of the Contention Window (CW).

4.2.2 Problems with MDA-based Schemes to Guarantee QoS in IEEE 802.11s

MDA was proposed to increase efficiency of the IEEE 802.11s protocol by reducing the possibility of collisions. It is an access method that allows MRs to access the wireless medium at selected times with lower contention that would otherwise be possible within two-hop neighborhood. Broadcast setup messages are exchanged between neighboring MRs to first negotiate MDAOPs. These MDAOPs are performed in multiple of time-slot units, during DTIM periodic interval. When a reservation has been successfully completed, the owner of the MDAOP can access the appropriate times-slots using EDCA with parameters according to the highest priority or DCF and its corresponding MAC parameters set. However, EDCA does not establish a strict priority policy but only relative priority for some nodes with respect to others. This observation shows that the MDA protocol does not guarantee a collision free access to MDAOPs. In addition, we note that interferences outside the two-hop neighborhood can still occur with MDA; this may degrade WMN performance. The contributions in [CL08] and [GM09] address the problem of interferences outside the two-hop neighborhood in MDA. Cicconetti et al. [CL08] use dynamic relocation (called, MDA-DR) of conflicting MDA time-slots that are two hops away from each requesting node. In MDA-DR, the flows are dynamically relocated⁴ in the case where enough interferences are detected outside the two-hop neighborhood, and also depending on their elapsed duration in the network. In fact, “older” flows are less likely to be relocated, i.e., their MDAOPs will be moved to other available time-slots. However, MDA-DR limits its performance improvement to throughput which makes it suitable only for data applications, such as file transfer and Web browsing and not for delay

⁴ Reserved time slots are relocated into contiguous blocks to reduce the fragmentation of time slots while keeping a reasonable signaling overhead for the dynamic reservation changes.

sensitive applications, such as voice. Furthermore, it lacks an admission control scheme that limits packet losses.

In [GM09], Medium Access through REservation (MARE) is proposed as a generalization of the RTS/CTS V-CS procedure to multi-hop wireless networks. In this scheme, rather than using the excessive beaconing broadcast of each node in MDA, MARE uses a more elaborated RTS/CTS-like scheme that is able to consistently notify the appropriate group of neighboring nodes to which a given set of time-slots has been reserved. The proposed scheme was shown to substantially decrease the overhead compared to MDA; since it requires a constant-size data frames it seems suitable for voice traffic. However, no scheduling and admission control were developed for MARE which limits its efficiency as well as its QoS guarantees for voice traffic.

It is worth noting that several researchers focused on defining admission control mechanisms in IEEE 802.11 wireless networks to differentiate between HO and new calls [ZL01] or between different types of traffic with different QoS requirements [WL04], [BR10a]. However, current MDA or improved MDA schemes [CL08] and [GM09] do not support differentiation between different service classes and also between HO and new calls.

In this paper, we propose and analyze a novel Enhanced Differentiated MDA (ED-MDA) with an Efficient Admission Control (EAC) scheme that resolves the QoS of voice and video traffic and prioritizes handover calls over new calls which, to the best of our knowledge, has not been investigated in IEEE 802.11s WMNs.

4.3 Overview of ED-MDA

In this section, we present a network model for WMNs. In particular, we first define key concepts, namely, transmission and interferences cliques, and then present the assumptions/notations used in the rest of this paper. We also present an overview of the proposed scheme ED-MDA.

4.3.1 Assumptions and Notations

We consider a multi-hop WMN as illustrated in Fig. 27. MRs (e.g., MR₃ and MR₁₇) aggregate and forward traffic from/to MCs. These MRs communicate with each other to form a multi-hop wireless backbone network. This backbone network forwards traffic to/from the gateway access points (e.g., P₁ and P₂) so that a packet traverses multiple hops (MRs) to reach a destination. We model the backbone of a WMN as an undirected graph called the connectivity graph $G = (V, E)$, where V represents the set of mesh nodes, and E represents the set of edges between nodes (connections). Among these nodes, we notice the subset P such that, $P \subset V$ to be the gateway access points that connect clients to the Internet. In the rest of this paper, MRs and gateway access points are collectively referred to as mesh nodes. Finally, $\forall (u, v) \in V$, an edge $e = (u, v) \in E$ if the distance between u and v , which is denoted by $d(u, v)$, is smaller than the transmission range (i.e., $d(u, v) \leq r$), where r represents the radio transmission range of nodes u and v .

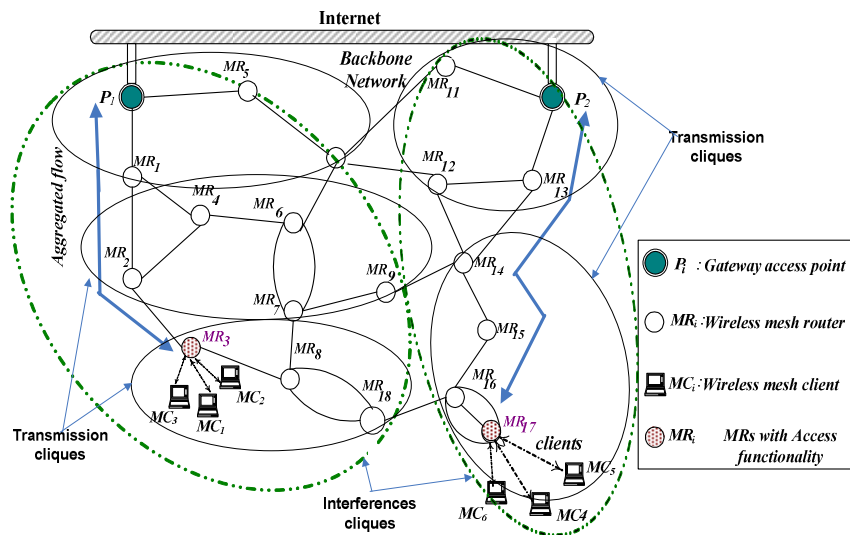


Fig. 27. A typical WMN where transmission cliques vs. interferences cliques are shown, topology of Mesh routers mutually interfering within cliques

We assume that the interference range is the same for all radio interfaces and is about twice the transmission range as it is a common assumption in IEEE 802.11 based networks [WW06]. Generally, we can define in any topological network two types of cliques (see Fig. 27): A transmission clique and an interference clique. Transmission

cliques are defined as sets of nodes, sharing the same channel and having a pair wise distance smaller than or equal to the radio transmission range of nodes in the pair (i.e., $d(u, v) \leq r$). Interference cliques are defined as sets of nodes that use the same channel and have a pair wise distance in the interval $]r, R]$, where R is the interference range. Interference cliques are used to identify nodes in the Carrier Sense Range (CSR). These interference cliques are briefly called cliques in the rest of the paper. Fig. 27 shows a sample MR-WMN that consists of two interference cliques and a number of transmission cliques inside each of the interference clique. Routers MR_{14} , MR_{15} , MR_{16} and MR_{17} compose the transmission clique where all routers links are in transmission range of each other. We used the algorithm presented in [DK88] to compute cliques shown in Fig. 27 and also for the 25-nodes MR-WMN shown in Fig. 30 (see Section 4.6). The logical topology is usually built using a Dijkstra shortest path based routing algorithm with the number of hops as the routing metric which is suitable for delay-sensitive RTC.

We note that the Erlang-b formula [KL75] describes the probability of call lost/blocked in a network (all time-slots are busy) until a time-slot becomes free. In Table 6, we assume that $i \in [1..4]$, $j \in [1..2]$ where i represents the service class and j is equal to 1 if it is an HO call, 2 otherwise (i.e., new call).

Table 6. Notations parameters

$AIFS_{AC[i,j]}$	Arbitrary Inter Frame Space of radio for service class $AC[i,j]$
τ	Time-slot duration
$L_{AC[i,j]}$	Packet size (including PHY and above) for service class $AC[i,j]$
O	MDAOP offset for a flow in a clique
π	MDAOP periodicity for a flow in a clique
δ	MDAOP duration for a flow in a clique
σ	DTIM utilization at time t

α	Fraction of the DTIM period reserved for MDA
T	DTIM period
C	IEEE 802.11 transmission rate in a clique
$T_{AC[i,j]}$	Upper bound of the number of time-slots for service class $AC[i,j]$
CFP	Contention Free Period to serve RTC
CP	Contention period to serve no RTC ($NRTC$)
E	Total offered load in Erlang-b
$B_{AC[i,1]}$	Dropping probability threshold of HO voice ($i=1$)/video ($i=2$) calls
$B_{AC[i,2]}$	Blocking probability threshold of <i>new</i> voice ($i=1$)/video ($i=2$) calls
$A_{AC[i,j]}$	Total number of HO ($j=1$) or <i>New</i> ($j=2$) [voice ($i=1$)/video ($i=2$)] calls admitted
$R_{AC[i,j]}$	Total number of HO ($j=1$) or <i>New</i> ($j=2$) [voice ($i=1$)/video ($i=2$)] calls rejected
$D_{AC[i,j]}$	The required delay tolerance in a clique for service class $AC[i,j]$
$D_{AC[i,j]}^{\max}$	The maximal delay for maximum number of hops m in a path for service class $AC[i,j]$

4.3.2 MDA without Traffic Differentiation

In MDA, the time between consecutive DTIM beacon frames is divided into time slots of length $32\mu s$. The periodic broadcast of beacon frames to all radios in a clique allows the synchronization of these DTIM intervals. Initially, MRs reserve the wireless medium for MDAOPs, which are reserved as multiples of time-slots during a given Contention Free Period (CFP) of a maximum access fraction ($MAF=\alpha T$) of the DTIM interval T (see Fig. 28-(a)). The remaining part of the interval as illustrated in Fig. 28-(a) is the contention period (CP) used for throughput-sensitive data

applications. Current MDA does not support traffic differentiation and has same behavior for all service classes in the network.

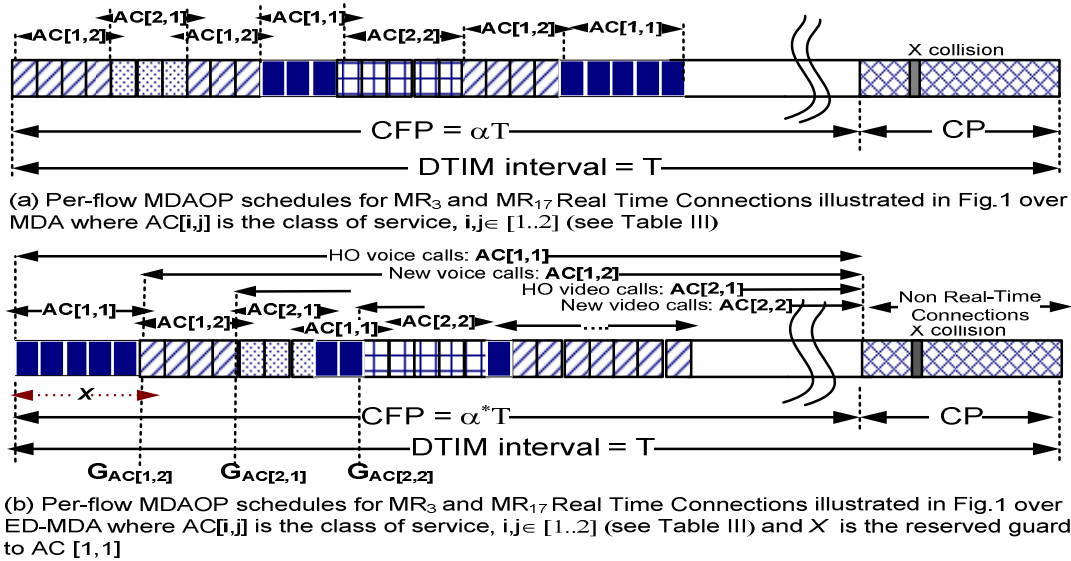


Fig. 28. (a) MDA per-flow schedule; and (b) ED-MDA per-flow schedule

4.3.3 MDA with Traffic Differentiation: ED-MDA

In this Section, we define the service classes supported by ED-MDA (see Table 7 and Table 8). Then, we describe the details of ED-MDA functionality in the presence of multi-service traffic and HO calls (see Fig. 28-(b)); particularly, we will show how HO calls are prioritized over new calls to maintain high resources utilization.

Table 7. QoS requirement: 4 service classes

Class-id	Delay sensitivity	Throughput sensitivity	Example
AC[1]	very high ($\leq 150\text{ms}$)	No	Voice
AC[2]	High ($\leq 300\text{ms}$)	High	Videoconference
AC[3]	Medium	Medium	Streaming
AC[4]	No	No	Best Effort

4.3.3.1 Service Classes

Generally, we can identify four types of classes related to the traffic nature either real time connections or not (see Table 7).

a) **VOICE TRAFFIC (REAL-TIME CONNECTIONS: RTC)**

Typically, voice traffic is delay-sensitive, but tolerates some frame losses. Acceptable upper bound limit of one way voice delay according to ITU-T G.114 must be smaller than 150ms [Si99]. Besides, voice packet loss (%) according to ITU-T P.862 (Perceptual Evaluation of Speech Quality, PESQ) should be kept below 10% [SX03]. Empirically, this value is found to reach 6.79% if the maximum allowed delay is 120ms [SX03].

b) **VIDEO TRAFFIC (REAL-TIME CONNECTIONS: RTC)**

Video traffic is delay-sensitive, but it tolerates some frame losses. In fact, the acceptable upper bound limit of one way video delay according to ITU-T H.261 must be smaller than 300ms [Si99]. Additionally, video packet loss (%) should be kept below 12%. Typically, a video traffic source generates frames at a constant rate over its active period. A popular example of RTC is a videoconference connection.

c) **VIDEO TRAFFIC (NO REAL-TIME CONNECTIONS: NRTC)**

In this kind of video traffic, the requirement is to reserve enough time-slots in each of the MDAOP periodicity sub-intervals in order to transmit 1 video packet (one downstream) of size $L_{\text{streaming}}$ over a clique having a transmission rate of C . This service class is considered as unidirectional calls when compared to voice/videoconference connections, the latter being bidirectional calls. An example of NRTC is video streaming.

d) **DATA TRAFFIC (NO REAL-TIME CONNECTIONS: NRTC)**

Data traffic is delay tolerable, but data packet loss (%) should be kept below 10%.

Recall that NRTCs traffic (AC [3] and AC [4]) in the proposed ED-MDA will be treated during the contention period (CP) and will not be studied in the scope of this paper. Indeed, we are only interested in supporting RTCs (AC [1] and AC [2]) that require low delays, low packet losses, and traffic differentiation support which are

difficult to guarantee with both IEEE.802.11 protocol [EG08] and current MDA [HM07].

4.3.3.2 Priority Setup in The Presence of Handover/New Calls

In the ED-MDA model we are studying, we consider a WMN consisting of several cliques. A clique serves heterogeneous users who require different service classes in the form of new calls generated in the clique or HO calls coming from adjacent cliques. According to the requirements of each service class in terms of blocking and delay constraints, it seems natural to consider in the DTIM interval this setup:

$$\Pr(AC[1]) > \Pr(AC[2]) > \Pr(AC[3]) > \Pr(AC[4]) \quad (29)$$

where $\Pr(x)$ denotes the priority of service class x .

Since we take HO calls into account and the no-delay-sensitive service classes are treated in CP duration in our proposed ED-MDA scheme (see Fig. 28-(b)), the problem under investigation is transformed and will be subject to different set of priority rules. The newly setup priority and QoS requirements as stated in (29) and Table 7 during the CFP period will become, in our model as follows:

$$\Pr(AC[1.1]) > \Pr(AC[1.2]) > \Pr(AC[2.1]) > \Pr(AC[2.2]) \quad (30)$$

where $\Pr(AC[i,j])$ denotes the priority of service class i , $i \in [1..2]$ and the sub-class-id j denotes whether the call is a HO or new call with $j \in [1..2]$.

Table 8. QoS requirement: 4 new service classes

Sub-class-id	Delay sensitivity	Throughput sensitivity	Service class
AC[1.1]	Very high ($\leq 150\text{ms}$)	No	Handoff voice calls
AC[1.2]	Very high ($\leq 150\text{ms}$)	No	New voice calls
AC[2.1]	High ($\leq 300\text{ms}$)	High	Handoff video calls
AC[2.2]	High ($\leq 300\text{ms}$)	High	New video calls

We assume that each call request (i.e., call) of service class $AC[i,j]$ includes the corresponding blocking probability and delay thresholds, $B_{AC[i,j]}$ and $D_{AC[i,j]}$ for $i \in [1..2]$ and $j \in [1..2]$, respectively. In addition, the call request contains a field called the relative blocking probability, $RB_{AC[i,j]}$ (see (31)), for accumulated blocking probability in WMN. $RB_{AC[i,j]}$ is updated at each intermediate WMN node and it is defined to be the total number of admitted calls per the total number of both admitted and rejected calls in WMN (computed according to the service class). Fig. 29 shows the main fields of the call request. $RB_{AC[i,j]}$ will be compared to the $B_{AC[i,j]}$ threshold according to each service class in our proposed admission control algorithm (see Section 4.5) and it is expressed as shown by (31).

$$RB_{AC[i,j]} = \frac{A_{AC[i,j]}}{A_{AC[i,j]} + R_{AC[i,j]}}, i \in [1..2], j \in [1..2] \quad (31)$$

Source	Destination	AC [i.j]	$B_{AC[i,j]}$	$D_{AC[i,j]}$	$RB_{AC[i,j]}$
--------	-------------	----------	---------------	---------------	----------------

Fig. 29. Main fields of the call request packet in IEEE 802.11s (ED-MDA)

4.4 The Real Time Connections Analysis in the ED-MDA per-flow Scheme

In this section, we first motivate the need to adjust MDAOP periodicity and CFP according to service classes in ED-MDA. Then, we analytically derive the MDAOP duration in number of time-slots $\delta_{AC[i,j]}^k$ and the MDAOP periodicity $\pi_{AC[i,j]}^k$ parameters in ED-MDA scheme in order to support QoS of RTC.

4.4.1 Impact of MDAOP Periodicity and CFP Period

It is known that one of the main objectives of MDA [HM07] is to provide stringent delay bounds for multimedia applications. However, current MDA does not provide traffic differentiation with different QoS requirements, which makes it hard to provide short delays in the presence of different types of traffic simultaneously in the network. According to the required delay for each service class and the maximum

number of hops in a path, we express, in sub-Section 4.4.5, the suitable MDAOP periodicity $\pi_{AC[i,j]}^k$.

We believe that an efficient adaptive adjustment of the CFP period (the adjustment of CFP is described and discussed in detail in sub-Section 4.5.1), must be considered in order to make efficient use of the time-slots, thus improving the QoS by exploiting the priority of each service class.

Consequently, taking periodicity and CFP duration adjustment into account allows the network to guarantee QoS to RTC in terms of smaller blocking/dropping probability, low delays and higher network capacity for each service class.

4.4.2 Bounds Conditions for CFP Adaptive Adjustment

In this sub-section, we introduce our model along with the underlying assumptions, so as to statistically describe the characterization of the traffic being studied and the guard computation of service classes in ED-MDA.

4.4.2.1 Traffic Characterization

We assume that our model follows the $M/M/1$ queuing model, the call arrival is described by a Poisson process and the inter-arrival time is exponentially distributed. We denote by $E, E = \lambda h$, the total amount of traffic offered in erlangs where λ is the call arrival rate and h denotes the call-holding time. The Erlang-b formula [KL75] is expressed by the probability: $P_b^{AC[i,j]} = \psi_b[E, T_{AC[i,j]}]$; this means that a new arriving call is blocked (meaning that all time-slots are busy) until a time-slot becomes free. Since the $\psi_b[E, T_{AC[i,j]}]$ formula assumes that the number of resources is known in advance (which is not the case in our model), we bound the Erlang-b to find an upper limit of the time-slots according to each service class. In our model, blocking threshold is considered to be known, while the number of time-slots is unknown. This can be better illustrated in UBTS (see Algorithm 4) where the probability of blocking is bounded by the required blocking probability $B_{AC[i,j]}$ and 1. $T_{AC[i,j]}$ is the upper bound of the number of resources (time-slots) and it is determined based on Erlang-b computation (see Algorithm 4).

Algorithm 4. Computing the number of time-slots for each service class: UBTS

Input: $AC[i,j]$: service class, E : total traffic, $B_{AC[i,j]}$: blocking probability threshold, CFP , τ : time slot duration, T :DTIM interval.

Output: $T_{AC[i,j]}$: Time-slots upper bound for each service class .

Hypothesis: No resources ($k=0$) condition $\psi_b[E,0]=1$;

we assume that $\psi_b[E, k-1] > \psi_b[E, k]$ for any k resources;

Step0: Bounded recurrence form of Erlang-b

$$B_{AC[i,j]} \leq P_b^{AC[i,j]} = \psi_b[E, T_{AC[i,j]}] = \frac{E^{T_{AC[i,j]}} / T_{AC[i,j]}!}{\sum_{k=0}^{T_{AC[i,j]}} E^k / k!} \leq 1;$$

For fixed k , we can express $P_b^{AC[i,j]}(k)$ as follows:

$$B_{AC[i,j]} \leq P_b^{AC[i,j]}(k) = \frac{E \times \psi_b[E, k-1]}{E \times \psi_b[E, k-1] + k} \leq 1;$$

then $B_{AC[i,j]} \leq P_b^{AC[i,j]}(k) = \frac{E \times P_b^{AC[i,j]}(k-1)}{E \times P_b^{AC[i,j]}(k-1) + k} \leq 1$;

Step1: Bounds conditions for k

We determine $0 \leq k = P_b^{AC[i,j]}(k-1) \left[\frac{E}{P_b^{AC[i,j]}(k)} - E \right] \leq E \times P_b^{AC[i,j]}(k-1) \left[\frac{1}{B_{AC[i,j]}} - 1 \right]$

Step2: Time-slots upper bound according to each service class

if ($i=1$ and $j=1$) **then** $T_{AC[1,1]} = \left\lceil \frac{CFP}{\tau} \right\rceil = \left\lceil \frac{\alpha^* T}{\tau} \right\rceil$;

else for $k \{ 1, 2, \dots, T_{AC[i,j]} \}$, $T_{AC[i,j]} = \left\lceil E \times P_b^{AC[i,j]}(k-1) \left[\frac{1}{B_{AC[i,j]}} - 1 \right] \right\rceil$;

where $(i=2, j \in [1,2]) \cup (i=1, j=2)$

where $B_{AC[i,j]}$ is a given blocking probability threshold for each service class; for example for voice call, $B_{AC[1,2]}$ is equal to 6%. This constraint is used to guarantee a smaller rate of blocking probability inside the network in order to meet users's expectations and to subsequently increase the operator's benefits.

By periodically applying steps (0-2) in UBTS, ED-MDA can compute an adequate upper bound of the number of time-slots to be reserved for each service class. Although this reservation scheme guarantees the finding of suitable time-slots, it may

cause interferences among time-slots that correspond to different users. Thus, a guard interval mechanism must be set to ensure that distinct transmissions belonging to different service classes do not interfere with each other.

4.4.2.2 Guard Computation

The class with highest priority in our case corresponds to HO voice calls; these calls could use up the CFP duration for voice packet transmission in the absence of other service classes, as shown in Fig. 28-(b). Additionally, a HO voice call takes up a minimum portion of the whole DTIM ($CFP - G_{AC[1.2]}$ denoted by X) in order to transmit its packets without sharing them with other classes as shown in Fig. 28-(b). ED-MDA manages the priorities among new and HO calls according to inequality (2); it operates under a time-slot reservation scheme expressed by the adjustment of CFP to maintain the HO priority (see Fig. 28-(b)) for video or voice calls.

The parameters $G_{AC[1.2]}, G_{AC[2.1]}, G_{AC[2.2]}$ are the respective guard intervals (computed according to service class). The guard interval threshold in ED-MDA is defined to be the maximum number of slots, during DTIM interval, that traffic of a given service class can use. The purpose of using guard intervals is to give higher priority to handoff calls over new calls and to protect against interferences among service classes. This is expressed in the following equations:

$$G_{AC[1.1]} = [0, \lceil CFP = \alpha^* T \rceil] \quad (32)$$

$$G_{AC[1.2]} = [\lceil X \rceil + 1, \frac{T_{AC[1.2]}}{K_{AC[1.2]}}] \quad (33)$$

where $K_{AC[1.2]}$ is the number of slots for voice

$$G_{AC[2.j]} = [\frac{T_{AC[i.2]}}{K_{AC[i.2]}} + 1, \frac{T_{AC[2.j]}}{K_{AC[2.j]}}] \quad (34)$$

where $K_{AC[2.j]}$ is the number of slots for video. If $i = 1$ then $j=1$ and if $i=2$ then $j=2$ in (34).

4.4.2.3 Average Traffic Measurement through N Observations

To get an approximately exact amount of total offered traffic load, we typically measure the traffic based on continuous observations over several periods of time $(\lambda_1, \lambda_2, \dots, \lambda_N)$, during which the instantaneous traffic λ_i is recorded at regular DTIM interval. The average offered traffic load, $\bar{\lambda}$, is computed as follows:

$$\bar{\lambda} = \frac{1}{M} \sum_{j=1}^N j \lambda_j \text{ where } M = \frac{N(N+1)}{2}. \quad (35)$$

4.4.2.4 MDA Formulation

In MDA, we characterize each MDAOP reservation request for call k by the triplet $\langle O^k, \pi^k, \delta^k \rangle_{k \in N}$ where O^k is the MDAOP offset from the DTIM start period, π^k is the MDAOP periodicity within the DTIM period, and $\delta^k = \left\lceil AIFS + \frac{L}{C} / \tau \right\rceil$ is the MDAOP duration in number of time-slots. π^k represents the number of times the specified MDAOPs repeat themselves equidistantly within a DTIM interval (T). In order to keep the advantages of a distributed MDA reservation protocol, a resource manager or an admission control has to be distributed among the links belonging to a given clique. In fact, due to the broadcast of MDA advertisement messages by the MDAOP requester node and the granter nodes [HM07], all nodes in the clique are aware of the reservation schedule $\langle O^k, \pi^k, \delta^k \rangle_{k \in N}$ in that clique.

4.4.2.5 ED-MDA Formulation

In per-flow ED-MDA scheduling, $2\delta_{AC[i,j]}^k$ is the number of time-slots reserved for each service classes (see (36)) in each of the $\pi_{AC[i,j]}^k$ (see (37)) sub-intervals that satisfies a hard constraint on a maximal delay $D_{AC[i,j]}^{\max}$ for maximum number of hops m in a path. In fact, we need to reserve enough time-slots during the sub-intervals for transmitting 2 voice or video conference (RTC) packets of size $L_{AC[i,j]}$, one upstream and the other downstream (bidirectional calls) over a clique having a transmission rate C . We note that this transmission occurs after duration $AIFS_{AC[i,j]}$. To prevent

exceeding the one-hop delay, the periodicity $\pi_{AC[i,j]}^k$ in the ED-MDA reservation request has to be sufficiently lower bounded by: $\pi_{AC[i,j]}^k \geq T/D_{AC[i,j]}^{\max}$. For the sake of simplicity, we consider a uniform distribution of $D_{AC[i,j]}^{\max}$ over interfering links even though a better repartition may take into account the non uniformity of traffic load over these links. Thus, the MDAOP duration (36) and periodicity (37) are expressed as follows:

$$2\delta_{AC[i,j]}^k = 2 \left\lceil \frac{AIFS_{AC[i,j]} + \frac{L_{AC[i,j]}}{C}}{\tau} \right\rceil \quad k \in N \text{ and} \quad (36)$$

$$\pi_{AC[i,j]}^k = \frac{DTIM}{D_{AC[i,j]}^{\max}} = \frac{T}{D_{AC[i,j]}^{\max}} \quad \text{where} \quad (37)$$

$$D_{AC[i,j]}^{\max} = \frac{D_{AC[i,j]}}{m} \quad (38)$$

The maximal delay is denoted by $D_{AC[i,j]}^{\max}$, i.e., the hard constraint on maximal delay for maximum number of hops m in a path and $D_{AC[i,j]}$ is the required delay by the service class $AC[i,j]$.

4.5 Admission Control Scheme EAC

The Call Admission Control (CAC) is an important task to ensure QoS requirements in multi-service WMNs. Therefore, in this study, ED-MDA is combined with a CAC for multimedia service networks, called EAC. The main objectives of EAC are: (1) differentiating between traffic with different QoS requirements; (2) providing an efficient adjustment of the CFP period to maintain high utilization of radio resources; and (3) providing higher priority to HO calls.

4.5.1 Possible CFP Values' Update Based on EAC Operation

To guarantee QoS to RTC in terms of delay and blocking probability, we need an efficient fine-tuning of the CFP duration according to service classes. EAC defines three variables: (1) α : holds the current fraction of the DTIM period reserved for MDA; (2) β : a multiplicative factor which is used to update CFP; and (3) VCS (Violation

Constraint State): assumes 0 when the state of the required blocking threshold is violated; otherwise assumes 1 (see Table 9). VCS-v is defined as a vector of VCSs (e.g., VCS_v="01"). The first variable is common in IEEE 802.11s whereas the latter two variables are EAC specific.

The parameter β is a multiplicative factor used to update the value of α and subsequently CFP according to the service class and VCS. A similar multiplicative factor (equal to 2) is implicitly defined in the original IEEE 802.11 binary exponential backoff; Contention Window (CW) is doubled upon each transmission failure. However, β in EAC can be either assigned statically, i.e. before runtime and remains constant or dynamically (adaptive).

EAC first stores the most recent S states of VCSs upon each call request where S indicates the length of VCS_v. For example, in the case of S=2, VCS-v equals to "10" means that VCS is equal to 1 followed by VCS is equal to 0. EAC re-adjusts α accordingly to the rules shown in Table 9. Besides, EAC checks the value of VCS_v; if the latter is equal to "11" and the current VCS is equal to 1 then α is divided by β . In other words, a slow decrease of α /CFP is preferred in this case since a QoS violation (for at least two states) is unlikely to be encountered. In the case where VCS_v is equal to "00" indicating a QoS violation and that the current VCS is equal to 1, α will revert to its current value. Furthermore, EAC multiplies the value of α by β when the current VCS is equal to 0.

Table 9. Alpha adjustment according to the blocking probability constraint state

Constraint Current state	State (S=2), VCS_v	Alpha/CFP adjustment
VCS=0 : Violation Constraint State	"00"	$\alpha_{AC[i,j]} = \alpha_{AC[i,j]} \times \beta$ (a)
	"01"	
	"10"	
	"11"	
VCS=1: No Violation Constraint State	"11"	$\alpha_{AC[i,j]} = \alpha_{AC[i,j]} / \beta$ (b)
	"01"	
	"10"	

	"00"	$\alpha_{AC[i,j]} = \alpha_{AC[i,j]} \quad (c)$
--	------	---

4.5.2 Time-Slots Reservation over ED-MDA by EAC Algorithm

EAC algorithm performs the necessary verification of the incoming traffic in terms of the DTIM utilization and decides whether the DTIM parameters (e.g., fractions of the DTIM parameter) should be modified according to the service class. If a violation in the blocking probability constraint ($B_{AC[i,j]}$) is encountered during a period T , EAC adjusts all the DTIM parameters in a way that may lead to a better utilization of the available resources. To realize DTIM parameters variation, we assume the existence of an online measurement mechanism that measures DTIM parameters variations and computes a suitable time interval, denoted by T , to capture periods with minimal DTIM parameters variation.

The algorithm starts by looking for RTC traffic (Step 0-1 of Algorithm 5); once found, EAC proceeds with the initialization phase (see step 2 of Algorithm 5). We apply UBTS algorithm to search for initial feasible CFP adjustments (i.e., computes adequately the upper bound of the number of time-slots for each $AC[i,j]$). EAC initializes this feasible solution rather than a randomly generated one. Then, according to our ED-MDA classification (see Table 8), EAC periodically checks whether DTIM exceeds a certain threshold ($\frac{CFP}{\tau}$) as well as the ($\frac{\alpha_{AC[i,j]}T}{\tau}$) threshold for the handoff voice $AC[1,1]$ and for the others service classes respectively. Based on the above, it then rejects or accepts the incoming calls (see Algorithm 5, line-7).

In case of a rejected call, EAC increments the rejection parameter (Algorithm 5, line-8) that can be used later for eventual adjustment of CFP parameters. Otherwise, (a call is accepted), EAC recalculates the relative blocking parameter $RB_{AC[i,j]}$ (31) and compares it to the given threshold blocking parameter; if $RB_{AC[i,j]}$ exceeds the threshold ($B_{AC[i,j]}$), then a violation of the constraint is detected; in this case, EAC re-adjusts the

parameters accordingly as depicted in row-0 of Table 9. Otherwise, it uses the information from row-1 of Table 9.

In fact, EAC makes use of blocking probability statistics, VCS and the VCS_v that are collected during the period of time T , to dynamically trigger or not CFP adjustment at the end of the time period (see Section 4.5.1).

Algorithm 5. Efficient Admission Control: EAC

Input: Call request: $k (AC[i,j], B_{AC[i,j]}, D_{AC[i,j]})$ in clique q_j where $B_{AC[i,j]}$ and $D_{AC[i,j]}$ are the blocking probability and the delay required for the $AC[i,j]$ respectively.

Output: Admission decision: accept or reject and CFP adjustment

- 1 Receive the call request from user with QoS requirements $(B_{AC[i,j]}, D_{AC[i,j]})$
- 2 **Step0: Traffic throughput-sensitive**
 if $(k \in NRTC)$ **then** the call is served in CP duration;
- 3 **Step1: Traffic delay and throughput-sensitive**
 if $(k \in RTC)$ {
- 4 **Step 2: Initialization**
 if $(k \in AC[1.1])$ **then** $i=Voice, j=HO$;
 if $(k \in AC[1.2])$ **then** $i=Voice, j=New$;
 if $(k \in AC[2.1])$ **then** $i=Video, j=HO$;
 if $(k \in AC[2.2])$ **then** $i=Video, j=New$;
 $\sigma = 0$;
- 5 **Step 3: Upper bound Time-slots computation**
 Call UBTS // Initial feasible CFP adjustment based time-slots computation for each $AC[i,j]$
- 6 **Step 4: DTIM utilization test**
 $U = \delta_{AC[i,j]} + \sigma$;
- 7 **if** $(U > (\frac{CFP}{\tau}) \ \&\& \ U > \frac{\alpha_{AC[i,j]}T}{\tau})$
- 8 **then** $R_{AC[i,j]}++$; exit; // The call k is rejected
- 9 $A_{AC[i,j]}++$; // The call k is admitted in CFP duration
- 10 $\sigma = \sigma + \delta_{AC[i,j]}$; // Update of DTIM utilization
- 11 **Step 5: CFP adjustment**

```

if ( $RB_{AC[i,j]} < B_{AC[i,j]}$ ) then VC.S =1;
    //No violation of the constraint: currently state
12 then if (VCS_v=="00")
     $\alpha_{AC[i,j]} = \alpha_{AC[i,j]}$ ; else  $\alpha_{AC[i,j]} = \alpha_{AC[i,j]}/\beta$ ;
13 else if (VCS==0) then  $\alpha_{AC[i,j]} = \alpha_{AC[i,j]} \times \beta$ ;
    //Violation of the constraint: currently state
Repeat Step 4 Every period of time  $T$  }
14 Update of the number of time-slots for each  $AC[i,j]$ ;

```

4.5.3 Impact of the State History Length: S

In the previous sections, for sake of clarity and simplicity, we considered $S=2$. We believe that considering a larger VCS_v could improve the QoS provided by ED-MDA; however, this improvement will come at a cost in terms of memory. In fact, for different values of S (e.g., $S=N$, $S=N/2$, etc), separate tables should be constructed to provide all possible state combinations, (i.e., 2^{N+1} , $2^{N/2+1}$ entries, etc). In our simulations, we studied our approach by considering bigger *history lengths* (e.g., $S \in [2..10]$). This experimentation is quite useful in practical deployment because it studies efficiently the channel behavior during the latest stored VCSs. Consequently, it helps taking the right decision concerning the acceptance or the rejection of the incoming calls. We note that the blocking probability constraint may be violated just two times during the time interval (e.g., $VCS_v="1001111111"$: 2/10); in this case, we adjust α according to equation (b) of Table 9, otherwise (e.g., $VCS_v="1111000000"$: 6/10) equation (a) of Table 9 is applied. In the case that the numbers of 0 and 1 are equal in the VCS_v (e.g., $VCS_v="1001010011"$: 5/10), ED-MDA takes into account current VCS to take decision. We did not consider the case where $S > 10$ because of the memory explosion problem. Besides, simulations results show there is no gain in increasing the value S beyond 6-10 (please refer to Section 4.6.2).

4.6 Simulation Results

In this section, we conduct a simulation study using ns-2 to evaluate and compare the performance of our proposed scheme, ED-MDA, with other existing

schemes. We evaluate several metrics: 1) end-to-end delay; 2) packet loss rate (%); and 3) call blocking/dropping probability. Note that the end-to-end delay is the sum of the access delays (the queuing delay, which is negligible in low load, and the contention delay) experienced in hops along the path's call from the MC to the gateway.

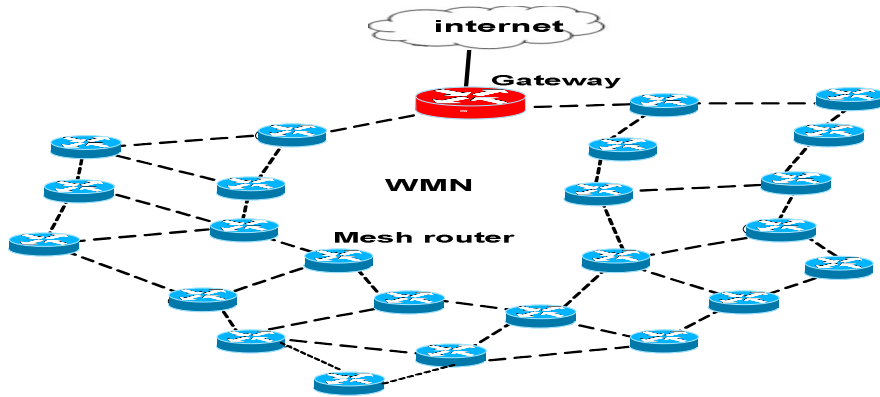


Fig. 30. The simulated topology

4.6.1 Simulation Configurations

The WMN topology used in simulations is arranged as a regular grid of 5x5 802.11 stations acting as MRs (see Fig. 30). More precisely, the node spacing is about 100 m and the position of each node is deviating from the regular grid by choosing a random perturbation of an angle in $[0, 2\pi]$ and a radius in $[0 \text{ m}, 25 \text{ m}]$. This kind of disturbance has been adopted in several existing WMN performance studies (e.g., [RH10d], [CL08]). We notice that in tests with MDA, all MRs apply the DCF and its MAC parameters after the scheduling of MDAOPs. The simulations parameters are presented in Table 10.

In the first scenario (scenario 1), simulations are performed under combination of two service classes (voice and video calls) while in the second scenario (scenario 2) four service classes (HO/new voice and HO/new video calls) are considered as described in sub-section 4.3.3.2. In the simulation results, we compare our proposed ED-MDA performance to: (1) WLAN 802.11 CSMA/CA for DCF [EG08]; (2) WLAN 802.11s MDA [HM08]; (3) WLAN 802.11s EDCA [HM07]; and (4) An improved version of MDA that considers dynamic relocation [CL08] of reserved MDAOPs two-hop away; we call this scheme MDA-DR.

Table 10: Simulation parameters

Parameter	Value	Parameter	Value
DTIM (ms)	32	P (bytes)	L-H
τ (μ s)	32 μ s	Topology	5x5 grid
Initial α	0.68	Routing protocol	AODV
H (bytes)	58	Rate/Packet size: video	0.01/1464
C (Mbps)	2	voice	G.729 codec
$D_{AC[1]}$ / $D_{AC[1.1]}$	150 ms/150ms	PHY radio model	SINR [CL08]
$D_{AC[2]}$ / $D_{AC[2.1]}$	300ms/300 ms	r: transmission range	250m
$B_{AC[1.1]}$ / $B_{AC[1.2]}$	4% / 6%	R: interferences range	550m
$B_{AC[2.1]}$ / $B_{AC[2.2]}$	10% /12%	EDCA	No MDA node [HM07]
β	1.5	MDA, MDA-DR	MDA node [CL08]
ED-MDA	<i>MDA node</i> based multi-service	DCF	No MDA node [EG08]

4.6.2 Results Analysis

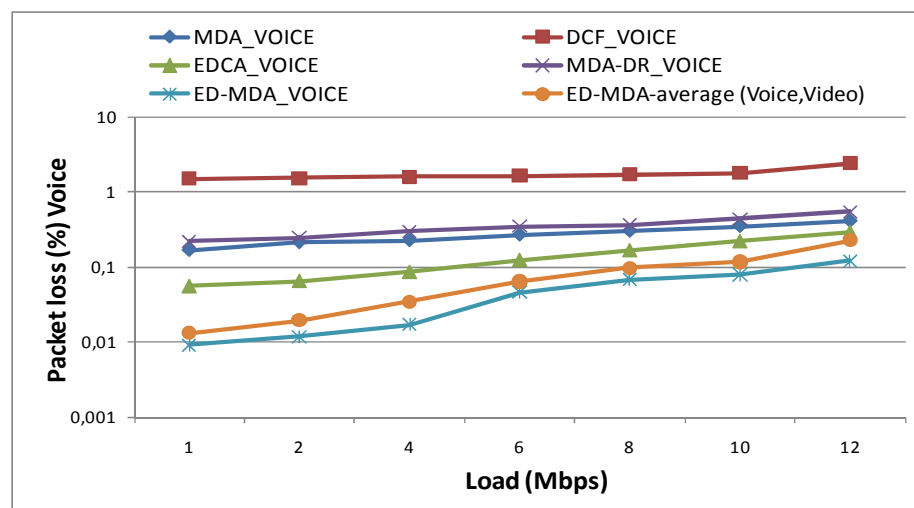


Fig. 31. Packet loss (%) for voice calls using different access Methods

MAC Access Method Analysis-DCF/MDA/EDCA/ED-MDA Using Scenario 1:

As presented in Section 4.2, the IEEE 802.11s draft 2.0 allows an optional contention-free MDA access method besides the well-known CSMA/CA contention-based access DCF method. We studied the performance of these access methods when transmitting data on the channel of each clique. Also, we study the performance of the mandatory MAC in IEEE 802.11s EDCA, and MDA-DR. In this paper, we are interested in studying the performance of these access methods compared to our proposed ED-MDA, which is based on MDA while taking into account multi-service traffic. Moreover, ED-MDA prioritizes HO calls over new calls; this feature will be simulated in scenario 2.

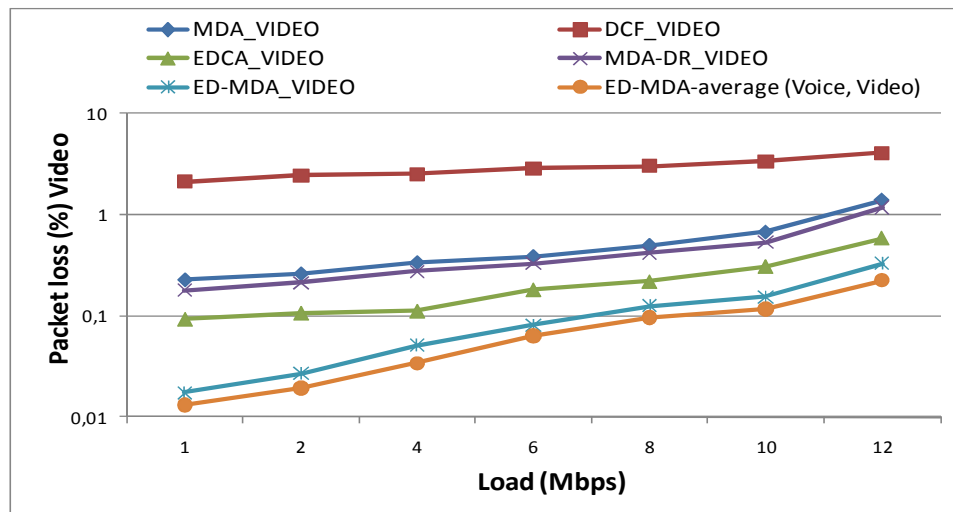


Fig. 32. Packet loss (%) for video calls using different access methods

Note that MDA allows less contentions to happen; this results in much lower packet losses (see Fig. 31 and Fig. 32) for MDA/MDA-DR compared to DCF. Fig. 31 also shows a reduction of packet losses for voice calls with EDCA; this can be explained by the fact that EDCA uses the MAC parameters set for the voice access category, giving higher priority to voice calls. By comparing the performance of ED-MDA to all other access methods, the packets losses are observed to be lower with ED-MDA. This can be explained by the fact that ED-MDA efficiently adjusts the CFP duration to serve, with high priority, voice calls while avoiding the starvation of video calls connections (see Fig. 31 and Fig. 32). Comparing ED-MDA voice, video and

average (packet losses) of the two classes, we clearly see that voice calls have lower losses in the network than video calls.

It is important to note that with EDCA, we give higher priority to voice calls over video calls; however, with ED-MDA, we adjust the CFP duration to guarantee efficient time-slot utilization for all classes. In fact, with this adjustment, we also prioritize voice calls over video calls by reserving a number of slots exclusively to voice calls as shown in Fig. 28-(b). This being said, ED-MDA and EDCA allow traffic differentiation but the way they apply this differentiation is a key factor making our approach most suitable in ensuring better network performance.

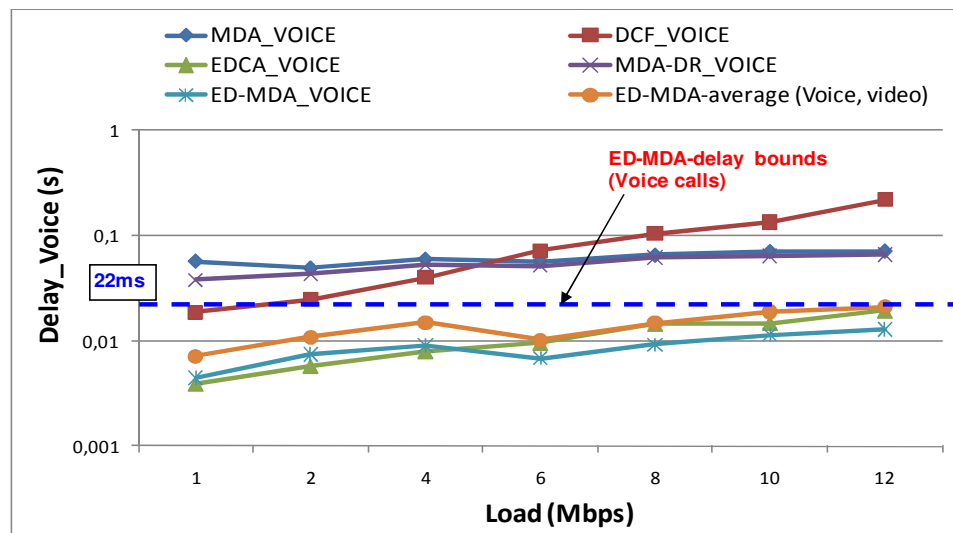


Fig. 33. Delay for voice calls using different access methods

Fig. 33 presents the average end-to-end delay experienced by network voice calls when using the various simulated schemes. In the case of low offered load (1 to 4 Mbps), where collisions are very rare, the DCF method provides lower delays than MDA and MDA-DR. We note that DCF/MDA/MDA-DR do not grant higher priority to voice calls compared with video calls; this explains their poor performance with respect to voice delay. Besides, due to the scheduling of MDAOPs although the time-slots are available in a low offered load condition, the average delay is higher with MDA and MDA-DR compared to DCF. In highly offered load situation (e.g., 10 Mbps), the average delay with MDA or MDA-DR does not exceed 224 ms; it is bounded by the

DTIM interval, which is equal to 32ms, multiplied by the maximum number of hops, which is equal to 7 in our topology. Fig. 33 shows that EDCA and ED-MDA schemes present lower delays, which is expected since these two schemes consider traffic differentiation and give priority to voice calls.

In the case of low offered load, the methods based on MDA wait for longer periods of time before transmitting in specific reserved contiguous time-slots even when collisions are rare and time-slots are available. However, in high offered load situation the delay in these schemes is bounded as explained above. Nevertheless, the delay provided by DCF and EDCA increases boundlessly with the increase of the offered load. We observe that the link delay with ED-MDA experienced by voice packets never exceeds the hard constraint on the maximal delay of $D_{AC[1]}/7=150/7=21.42$ ms.

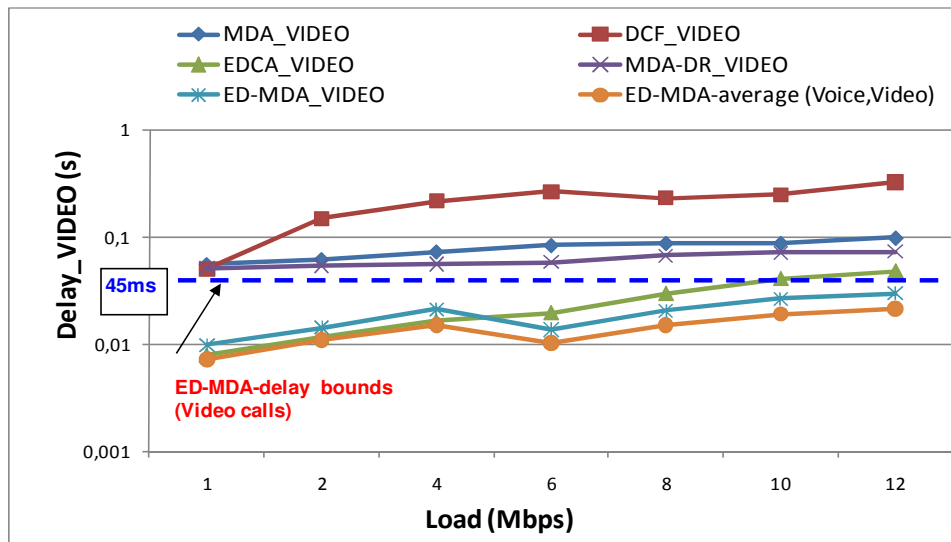


Fig. 34. Delay for video calls using different access methods

Fig. 34 shows that ED-MDA delay, experienced by video packets never exceeds the hard constraint on the maximal delay of $D_{AC[2]}/7=300/7=42.85$ ms.

MAC Access Method Analysis-MDA/ED-MDA Using Scenario 2: We take into account new and HO calls in this scenario in addition to video and voice calls.

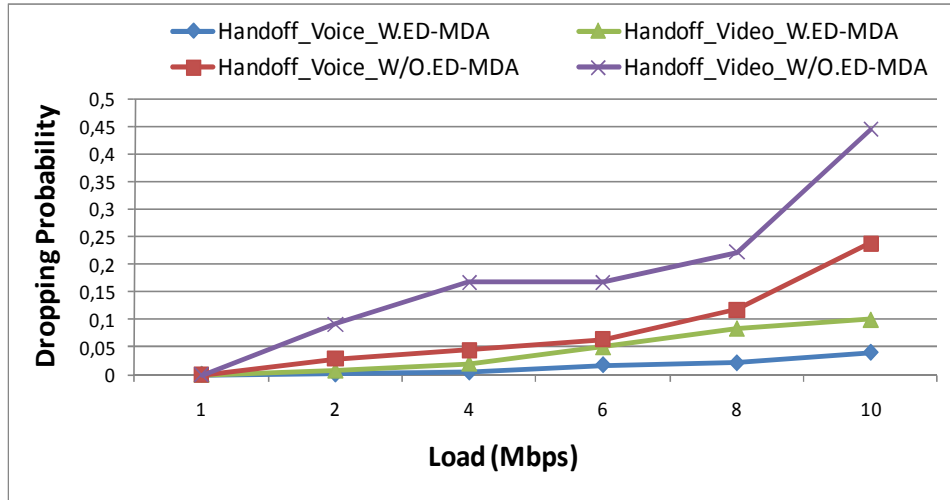


Fig. 35. HO calls dropping probability with and without ED-MDA

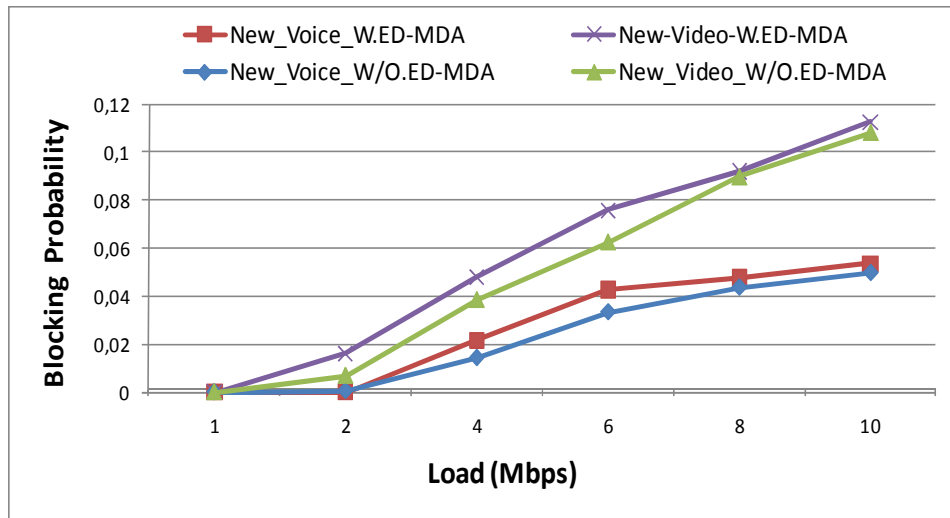


Fig. 36. New calls blocking probability with and without ED-MDA

In order to study the impact of the reserved guard $CFP-G_{AC[1,2]}$ on the HO voice calls dropping probability (see Fig. 35), we simulated the network behavior with the same parameters as in ED-MDA (called W.ED-MDA) and without ED-MDA (called W/O.ED-MDA). Also similar tests are done to investigate the impact of the guard threshold for HO video calls. We note that the admission control policy of EAC, with guard thresholds, offers HO calls dropping rate much lower than without guard thresholds (i.e., MDA) which varies from 2 to 12 times for an offered load ranging

from 2 to 10 Mbps. Indeed, up to 2 Mbps, all HO voice or video calls with ED-MDA are admitted but over 4Mbps, the dropping probability ranges from 0.4% to 2% for HO voice and video respectively which is very acceptable for users. However, when offered load is equal to 10 Mbps, the dropping probability does not exceed 4% which is still acceptable for voice traffic.

Fig. 36 shows that new connections blocking probability increases with the offered load. We observe that for offered loads varying between 1 and 8Mbps using ED-MDA scheme, we get an acceptable blocking rate (0 to 5% for new voice calls and 0 to 9% for new video calls); beyond these values, the blocking rate for video calls exceeds 11% when the offered load exceeds 8Mbps. Fig. 36 also shows that the lowest blocking rate is offered when using W/O.ED-MDA; this is expected since the new and HO calls are treated similarly; as long as time slots are available, these calls are admitted. However, when a guard threshold is used for new calls using ED-MDA, a part of the CFP (i.e., a number of time-slots) is reserved exclusively for HO voice calls and the guard threshold for HO video calls is higher than the guard threshold for new video calls; the new calls have less chance to be admitted which causes the increase of call blocking probability.

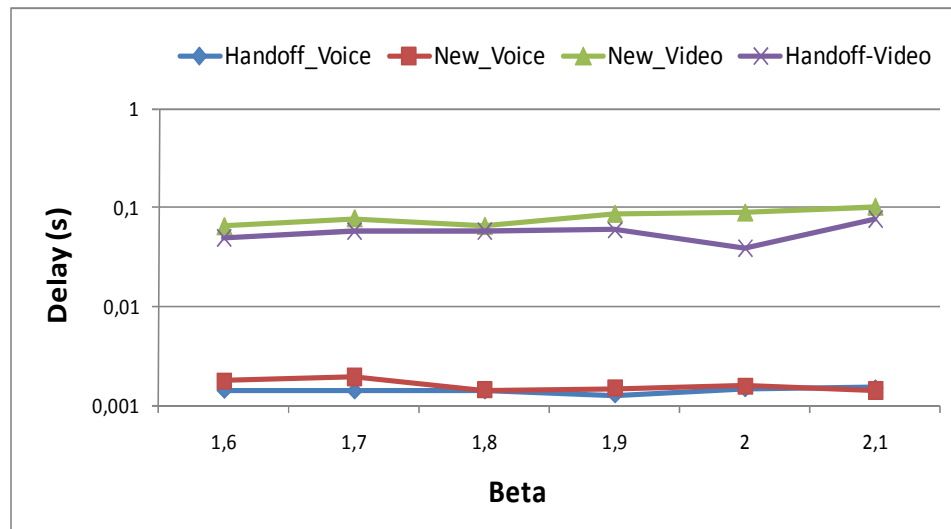


Fig. 37. Impact of the beta value on the HO and new calls delay for an offered load equal to 10 Mbps with ED-MDA (S=2)

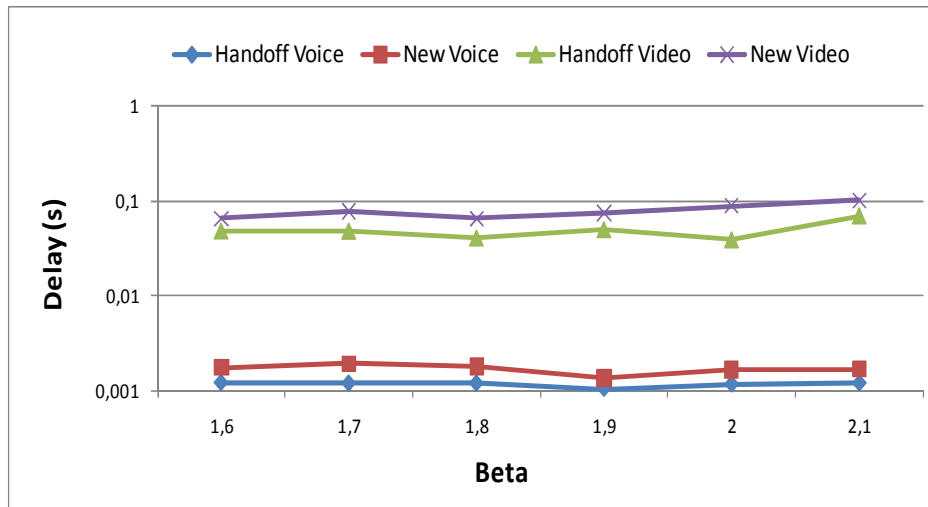


Fig. 38. Impact of β on the HO and new calls delay for an offered load equal to 10 Mbps with ED-MDA ($S=10$)

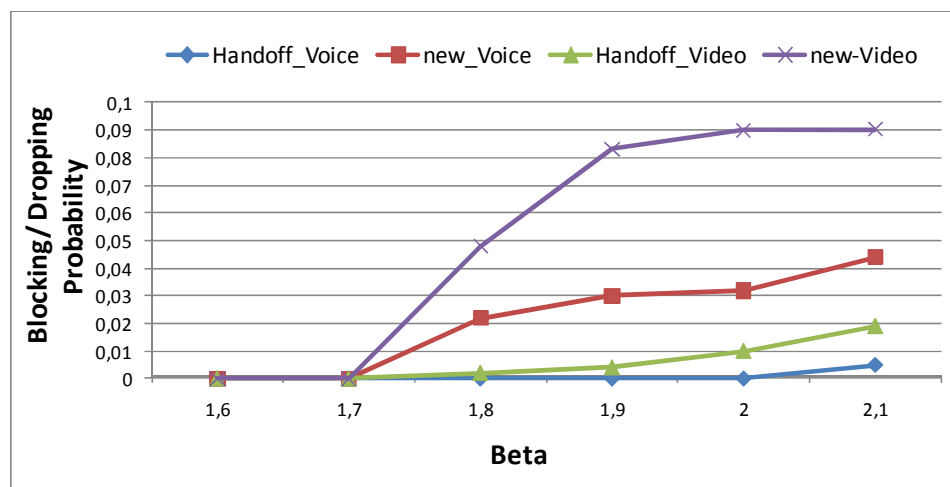


Fig. 39. Impact of β on the HO and new calls Blocking/Dropping probability for an offered load equal to 10 Mbps with ED-MDA ($S=2$)

Fig. 37, Fig. 38 and Fig. 39 show the impact of β on the HO and new calls with ED-MDA and with different history lengths. The delay of HO or new voice calls is much lower than the delay of video calls which is expected since ED-MDA uses reserved guard for voice calls. But, delays with HO and new voice do not vary a lot when varying β ; this is due to the fact that S is equal to 2 (see Fig. 37). Indeed, we do not have enough history to make an accurate adjustment; the bigger the history length,

the more accurate the adjustment and thus the better the performance of our approach improvement is expected.

In Fig. 38, we consider larger history length ($S=10$). We observe that with $S=10$, when we vary β especially starting from $\beta=1.7$, handoff voice/video calls have smaller delays than new calls.

Fig. 39 shows the blocking and dropping probability while varying β . We observe that the dropping of handoff calls is much lower (starting from β equal to 1.8) than new calls regardless of the nature of the call (voice or video). An interrupted communication is a very frustrating phenomenon that may happen to MCs. Thus, our proposed admission control protocol EAC coupled with ED-MDA avoids the forced termination of an ongoing call at the expense of slightly higher blocking of new calls. This is a price worth paying to provide low delays and loss rates while satisfying QoS requirements of multi-service WMNs.

For the next evaluation, we choose to vary the history length, S from 2 to 10 (see Fig. 40 and Fig. 41) while the offered load is kept equal to 10 Mbps and β is set to 1.9.

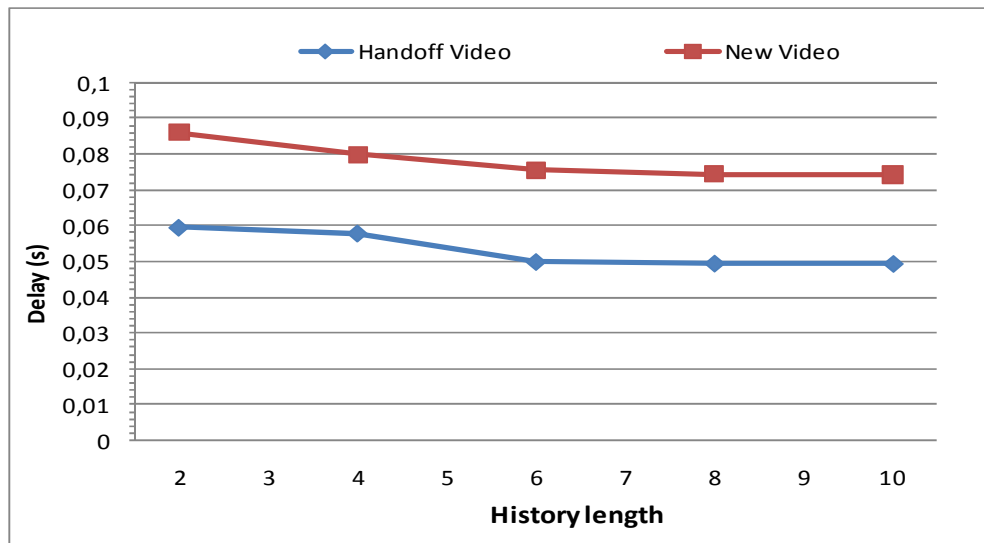


Fig. 40. Impact of the history length on the HO and new video calls delay for an offered load equal to 10 Mbps and β equal to 1.9 with ED-MDA

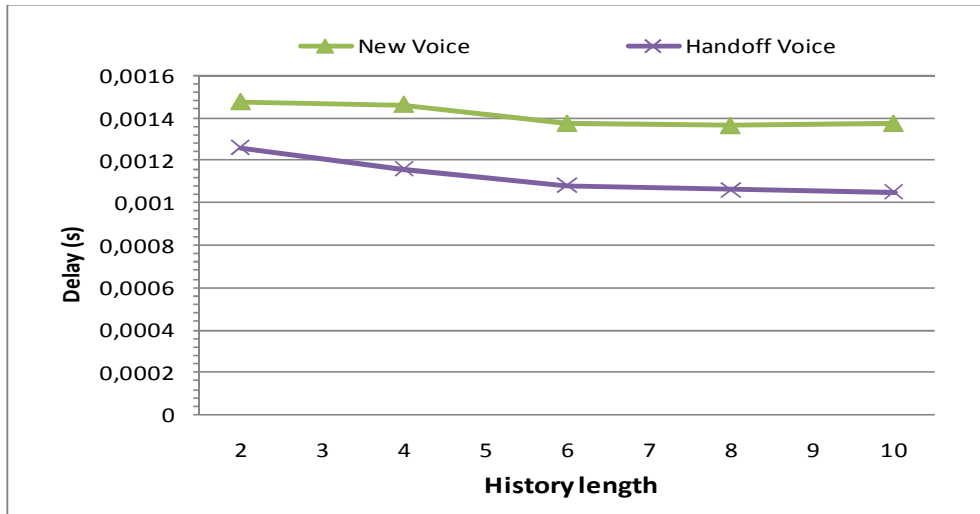


Fig. 41. Impact of the history length on the HO and new voice calls delay for an offered load equal to 10 Mbps and β equal to 1.9 with ED-MDA

Fig. 40 and Fig. 41 show that when the history length varies from $S=2$ to $S=6$, the delay for handoff video calls and voice calls drops as low as 14% and 16% respectively. However, starting from $S=6$, a slight delay improvement of 1% and 3.7% is observed for handoff video calls and voice calls respectively. Thus, there is no gain in increasing the value of S beyond 6-10.

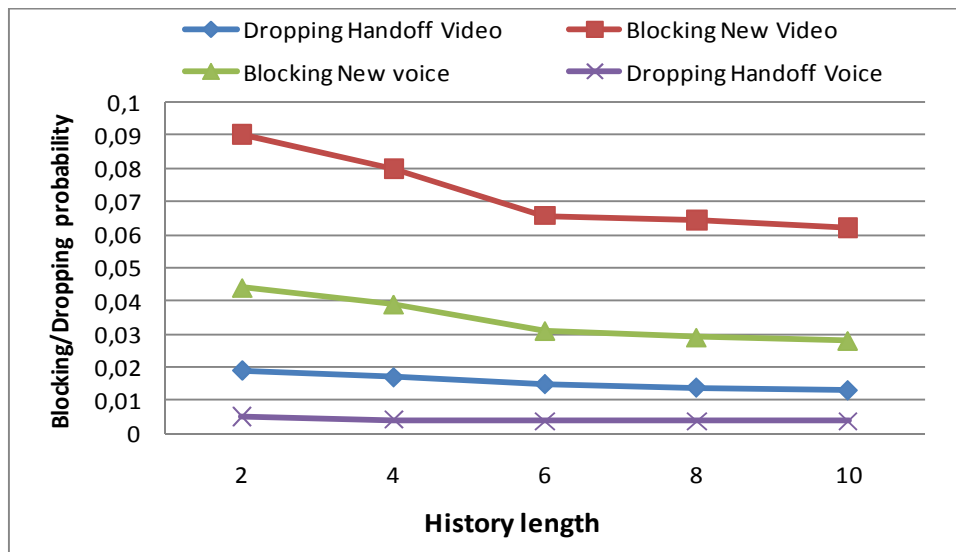


Fig. 42. Impact of the history length on the HO and new calls blocking/dropping probability for an offered load equal to 10 Mbps and β equal to 2.1 with ED-MDA

Fig. 42 shows the blocking and dropping probability while varying S . We observe that the dropping of handoff calls is much smaller (from S equal 2 to 6) than new calls regardless of the nature of the call (voice or video). Starting from $S=6$, a slight dropping probability improvement of 3% and 12% is observed for handoff video calls and voice calls respectively. However, from 2 to 6, the dropping probability drops as low as 20% for handoff video calls and 23% for handoff voice calls respectively. We observe also a significant improvement of the blocking probability for voice and video calls when the history length varies from 2 to 6.

4.7 Conclusions and Future Work

In this paper, we have shown that MDA that carefully reserves dimensioned CFP according to each service class (ED-MDA) provides better voice/video QoS in terms of packet loss, delay and dropping rate. Furthermore, we proposed an admission control algorithm EAC that takes the HO and multiple service classes into account. Our proposed scheme ED-MDA combined with EAC is based on the guard interval principle which reserves necessary time-slots (i.e., resources) to grant high priority to HO calls. Compared to MDA mechanism without guard interval, the simulation results show that our proposed solution outperforms the current MDA and many others existing schemes regarding the management of HO calls. Indeed, it offers a lower dropping probability and a lower delay for voice/video HO calls at the cost of a slightly higher blocking rate of new calls.

Currently, we plan to extend our approach to take into account interferences outside the two-hop neighborhood that can still occur with multi-service MDA.

4.8 Acknowledgment

We would like to thank Dr. F. Bastin for his valuable comments. The research reported in this manuscript has been supported in part by Natural Sciences and Engineering Research Council of Canada (NSERC) and Bell Canada.

Chapitre 5: Optimization Model for Handoff-Aware Channel Assignment Problem for Multi-Radio Wireless Mesh Networks

Jihene Rezgui, Abdelhakim Hafid, Racha Ben Ali, Michel Gendreau

Abstract

Optimal channel assignment (CA) in Multi-Radio Wireless Mesh Networks is an NP-Hard problem for which solutions usually leave several links interfering. Most of these solutions usually consider the overall throughput as the main optimization objective. However, other objectives have to be considered in order to provide better quality wireless connections to non-stationary users. In this paper, we propose a multi-objective optimization model that, besides maximizing throughput, improves fairness and handoff experience of mesh clients. In this model, we use the Jain's index to maximize users' fairness and we allow same-channel assignments to links involved in the same high handoff traffic, thus reducing handoff-triggered re-routing characterized by its high latency. Then, we propose a centralized variable neighborhood search and a Tabu search heuristics to efficiently solve our model as an offline CA process. Moreover, in order to adapt to traffic dynamics caused especially by user handoffs, we propose an online CA scheme that carefully re-assigns channels to interfaces with the purpose of continuously minimizing the re-routing overhead/latency during user handoffs. We further improve this online scheme using load balancing. Simulation results show the good performance of our proposed approach in terms of delay, loss rate, overall throughput and fairness. Particularly, performance results of our online handoff-aware CA show the effectiveness of handoffs not involving path re-routing in decreasing the delay, especially when considering load balancing.

Keywords: Channel assignment, Handoff, Network management, Wireless communication, Wireless mesh network.

Status: This paper is submitted to IEEE Journal Transactions on Vehicular Technology [RH10g]. The ideas presented in this paper are largely based on the following published papers:

- Handoff-Aware Channel Assignment for Multi-Radio Wireless Mesh Networks, IEEE ICC, South Africa, 2010 [RH10b].
- Meta-heuristics for Channel (Re-) Assignment Problem for Multi-Radio Wireless Mesh Networks, INFORMS Telecom, Montreal, Canada, 2010 [RH10c].
- Variable Neighborhood Search Method for Channel Assignment Problem in Multi-Radio Wireless Mesh Networks, IEEE LCN, Denver, Colorado, U.S.A, 2010 [RH10f].

5.1 Introduction

Mesh Networks (WMNs) consisting of single radio mesh routers (MRs) operating on a single channel suffer from a dramatic decrease in network capacity as the number of wireless hops increases. Equipping MRs with multiple radios on non-overlapping or partially-overlapping IEEE 802.11 channels can significantly increase the WMN capacity if a special care is taken in the assignment of channels to the radio interfaces. However, a channel assignment (CA) in Multi-Radio WMN (MR-WMN) that achieves good performance presents considerable challenges. In fact, the problem of optimal CA in maximizing performance by minimizing interferences has been proven to be NP-hard since it can be mapped into a graph-coloring problem in arbitrary mesh topology [RG04]. Therefore, several heuristic algorithms providing more or less good CA solutions for MR-WMN were proposed and evaluated in the literature [RG04], [MD05], [RW07], [SG08].

In this paper, we propose a novel multi-objective CA optimization model in opposition to existing models/schemes that usually tackle the optimization problem of a single or two objectives tightly related to interferences [MD05], [RW07], [SG08]. In particular, our proposed model, besides maximizing throughput and fairness, minimizes handoff overhead and minimizes traffic load variance to achieve load balancing. Fairness can be seen as the capability of the network to always provide each client its

fair share of the maximal achieved overall throughput in the WMN. Thus, when the WMN is saturated, the maximal throughput is equally divided between demanding clients. The minimization of handoff overhead is implemented by determining solutions that build MR-WMN logical topologies with less re-routing overhead; in fact, it allows most of the re-routed traffic (because of handoff) to change only first-hop MRs along the pre-handoff paths. This will significantly minimize service disruption without the resources and overhead generated by existing approaches [SX08].

Our contributions, in this paper, can be summarized as follows: (1) We propose a novel multi-objective optimization model that assigns channels to interfaces in MR-WMNs; the proposed model includes 4 objective functions and therefore provides a better flexibility compared to existent CA schemes; moreover, it is the first model to consider handoff traffic in CA; (2) the CA problem being NP-hard, we propose two Meta-Heuristics, called MHALB and MTABU, to search for a near-optimal CA solution of the proposed model. MHALB is based on Variable Neighborhood Search (VNS) [HM03] which is used for solving combinatorial optimization problems in which a systematic change of neighborhood within a local search is carried. MTABU is based on Tabu search which keeps a history of the search process and prohibits comebacks to previous solutions [Ge02]. We also propose a greedy search Heuristic (Algorithm 6) to search for an initial feasible CA solution for the Load Balancing and Throughput maximization problem, called HALBTH; thus, starting with this good initial solution rather than a randomly generated one, MHALB and MTABU provide a better final solution; (3) we propose a distributed handoff-aware channel re-assignment, called HA, that improves handoff performance by limiting most of the flow path re-routings to first-hop MRs. It dynamically re-assigns channels whenever the handoff rate, which is measured at local MRs, exceeds a predefined threshold that significantly degrades the performance. We also extend HA to take into account load balancing in a scheme called HALB with the purpose of efficient utilization of network resources.

While in the first and second contributions we provide centralized static CA heuristics (MHALB/MTABU) that compute near-optimal solutions by considering several objectives, in the third contribution we provide a distributed dynamic CA

algorithm that continuously improves/adapts the near-optimal solution produced by the static heuristics to the dynamic variation of traffic load and handoff rates. The process of computing a CA solution by HALBTH, MHALB and MTABU combined with the continuous adaptation by HA and HALB is called HALBTH⁺ (i.e., HALBTH + HALB), MHALB⁺ and MTABU⁺ respectively.

The remainder of the paper is organized as follows. Section 5.2 provides a brief overview of related research. Section 5.3 defines the WMN CA problem and presents a mathematical formulation of the problem solution. Section 5.4 proposes a resolution of the proposed model using a greedy search algorithm, variable neighborhood search algorithm and Tabu search algorithm. Section 5.5 proposes the on-line CA scheme. Section 5.6 evaluates the proposed solution via simulations. Finally, Section 5.7 concludes the paper.

5.2 Related Work

There exists a vast literature on CA schemes in wireless networks; a thorough study can be found in [SV04]. These schemes can be broadly classified in three categories: (1) fixed CA (FCA): it assigns channels once; indeed, channels assignment does not change over time [RG04], [MD05], [SG07]; (2) dynamic CA (DCA): it continuously updates the assignment of channels to improve performance [RW07], [SG08], [SV04], [BC04], [RC04], [RC05]; and (3) hybrid CA (HCA): it applies a fixed CA scheme for some interfaces and a dynamic CA scheme for others [KV05], [KV06], [RB06].

In FCA, a set of channels are permanently allocated to interfaces with respect to the interface switching time; they can be reused in other MRs, sufficiently distant, such that co-channel interference is tolerable. The advantage of FCA is its simplicity; its drawback is that it does not adapt its channel reuse to accommodate the varying offered load. Such schemes can be further subdivided into (a) common channel assignment (CCA) [RG04]: the radio interfaces of each node are all assigned the same set of channels; and (b) varying channel assignment (VCA) [MD05, SG07]: the interfaces of different nodes may be assigned different sets of channels, such as in CLICA

(Connected Low Interference Channel Assignment) [MD05] and in MesTiC (mesh-based traffic and interference aware CA) [SG07].

In DCA any interface can be assigned any channel, and interfaces can frequently switch from one channel to another. Therefore, when nodes need to communicate with each other, a coordination mechanism [RG04], [MD05], [SV05], [SG07], [BC04] has to ensure they are on a common channel. The benefit of DCA is the ability to switch an interface to any channel, thereby offering the potential to use many channels with few interfaces. However, the key challenges involve channel switching delays (typically on the order of milliseconds in commodity 802.11 wireless cards), and the need for coordination mechanisms for channel switching between nodes. Distributed-HYAcinth (D-HYA) which is a distributed version of Centralized-HYAcinth (C-HYA) [RW07, RC05] builds a spanning tree network topology and reacts to traffic load changes to improve both load balancing and aggregate throughput. However, it poses a potential problem to multipath routing in mesh networks and does not consider the network connectivity.

In HCA the concepts of fixed and dynamic strategies are combined by applying FCA to some interfaces and DCA to others [KV06]. HCA can be further classified based on whether the fixed interfaces use a CCA or VCA approach. HCA schemes are attractive because, as with FCA, they allow for simple coordination algorithms, while still retaining the flexibility of DCA. Link Layer Protocol (LLP) [KV06] categorizes available interfaces into fixed and switchable interfaces, i.e., can be switched over short time scales among non-fixed channels based on the amount of data traffic. LLP assigns different channels to fixed interfaces of different nodes; then, all channels can be used; the switchable NICs are used for connectivity only. However, it does not take into account the traffic load in assigning the fixed channels. Interference-Aware Channel Assignment (IACA) [RB06] is based on the multi-radio conflict graph (MCG) and assigns one radio in each node to operate on a default common channel. IACA provides alternate fallback routes, and avoids flow disruption by traffic redirection over a default channel. Nevertheless, it does not consider the traffic load in assigning channels.

In several contributions, the CA problem is combined with other performance problems in MR-WMN and is modeled as an optimization problem. In [RW07] authors formulate the joint CA, interface assignment, and MAC fairness problem as a cross-layer non-linear mixed-integer network utility maximization problem. The authors choose to maximize a utility function that reflects the proportional fairness among the MAC flow rates. However, they neglect optimizing the overall throughput since they show its absolute value without showing how far it is from the optimal one and if a tradeoff between throughput and fairness can be achieved. Besides, they do not consider the non-uniformity of the traffic on logical links within a clique (i.e., a clique is a graph that has all its vertices adjacent to each other), which is generally affected by traffic demands and routing protocol. Subramanian et al. [SG08] consider the problem of CA with the objective of minimizing the overall network interference. They formulate the problem using semi-definite and linear programs. They extend their model to non-uniform traffic through weighting the interferences by traffic loads. Then, using a distributed Tabu search CA and a centralized greedy CA, they obtain tight lower bound solutions to this NP-hard problem. However, the authors in [SG08] do not consider the MAC fairness in their problem. Further-more, due to the centralized nature of the Tabu search CA approach; it is adequate only for ‘managed’ WMNs. Indeed, their approach is run as an offline CA that does not adapt to dynamic traffic variation.

Several CA schemes in the literature [RG07], [RC04] and [RC05] take into consideration the load in establishing logical links in MR-WMNs. In our CA scheme, we take into account both (1) the non-uniformity of handoff traffic between different MRs; and (2) the non-uniformity of traffic load over the MR-WMN. In fact, due to the mobility of a fraction of mesh clients (MCs), the load on each MR is usually dynamic. Therefore, a standard load-aware CA scheme taking into account only the heterogeneity of the offered load has to be reconsidered/reevaluated in the case of dynamic load incurred by handoff traffic. The first issue related to MC handoff is that a handoff on a different channel causes channel scanning and re-association (latency) that may disrupt real-time services; this is usually resolved in the literature using proactive scanning [WT07]. The second issue, which is tackled by our scheme, is that handoff may also

involve re-routings on other MRs (e.g., second-hop MRs) in addition to first-hop MRs; this may result in much more latency and therefore much more service disruption. To the best of our knowledge, the second issue has not been considered, in the literature, using a dynamic CA scheme. Indeed, most existing approaches [AD06, SC06] make use of proactive multipath routing schemes that establish/reserve alternate paths a priori to handoffs; this may be very inefficient in terms of resource utilization and overhead. Thus, we propose a handoff-aware dynamic CA scheme that re-assigns channels to MR interfaces with the objective to reroute most of the handoff traffic by changing only first-hop MRs along the pre-handoff flows paths. This will significantly minimize service disruption without the resources and overhead generated by existing approaches. Furthermore, in order to increase the overall throughput of the WMN [SH05], we take into account load balancing when choosing among different handoff-aware assignment solutions.

5.3 Formulation

5.3.1 Problem Description

We consider the problem of CA in MR-WMN by exploiting the tradeoffs among throughput, fairness, traffic load balancing and handoff overhead. Indeed, maximizing fairness will significantly reduce the overall throughput (as observed in 802.11 networks [JL05]); in opposition, if load balancing is maximized, the overall throughput is indirectly increased [DV06]. Assigning a same channel to interfaces of neighboring MRs minimizes handoff overhead (compared to handoff on different channels); however, this may cause high interference levels and thus degrade network performance in terms of throughput. Therefore, optimizing one of these criteria will affect/undermine other criteria; thus, a multi-objective approach is definitively recommended for this kind of problems. In this Section, we propose a multi-objective optimization model that offers the possibility to optimize a combination of several criteria. Further the NP-Hard property of the proposed CA model (property of CA problems in general), the evaluation cost of multi-objective functions is very high

compared to single objective functions; therefore, appropriate heuristics must be carefully designed to resolve the problem.

The objective of our proposed CA scheme is four fold: (1) Perform a static initial near-optimal CA over the whole network while considering throughput, fairness, load balancing and handoff; (2) perform a dynamic CA only if necessary to minimize packets losses and delay degradation caused by handoffs; while considering (a) handoff overhead reduction and (b) traffic load balancing to improve the performance of the MR-WMN for a dynamic handoff traffic; (3) increase the capacity of MR-WMNs by accepting more users through throughput maximization; and (4) ensure users fairness in utilizing the shared capacity, i.e., each user has its fair share of throughput.

5.3.2 Network Model

We consider a MR-WMN as illustrated in Fig. 43 MRs forward traffic from/to MCs that are associated to them. MRs also forward traffic between each-other while forming a meshed multi-hop wireless backbone network. This network forwards the user traffic, i.e. MC's traffic, to/from the gateways, which are connected to the Internet.

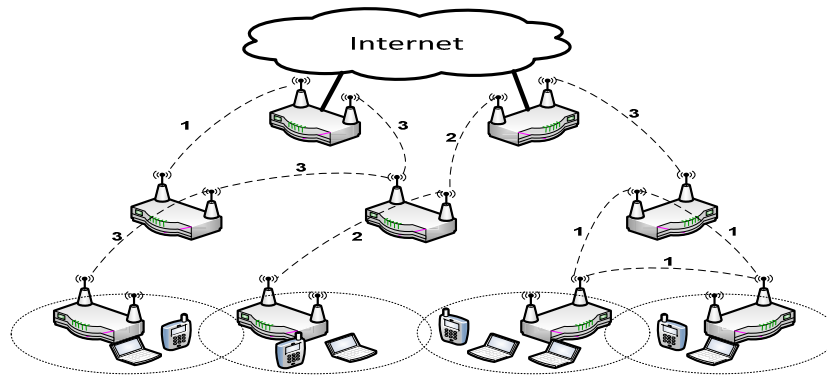


Fig. 43. A 2-radius MR-WMN with 3 orthogonal 802.11 frequencies assignment

We formally model the wireless meshed backbone of MRs as an undirected graph, called connectivity graph, $G = (V, E)$ where V represents the set of mesh nodes and E the set of edges between these nodes. Among these nodes, $P \subset V$ represents the set of the gateways that connect to the Internet. $\forall (u, v) \in V$, an edge $e = (u, v) \in E$ if the distance between u and v , denoted $d(u, v)$, is smaller than the minimum range,

denoted $\min(r_u, r_v)$, of u and v (i.e., $d(u, v) \leq \min(r_u, r_v)$) where r_u and r_v represent the radio transmission ranges of nodes u and v respectively. The connectivity graph after channel assignment is denoted $G_A = (V, E, A_G)$ where $A_G = \{A_G(u), \forall u \in V\}$ and $A_G(u)$ is the set of channels assigned to u . We denote NC the number of channels per radio, and R the number of radios per node; typically, we have $R < NC$.

5.3.3 Cliques

A good CA avoids, as much as possible, situations where adjacent links use the same channel; this allows achieving efficient channel utilization and minimizing interferences. However, in most cases, heuristics [RG04] leave some adjacent links using the same channel. Channel conflicts will happen if these adjacent links transmit the data at the same time. To solve this problem, we apply the concept of clique from graph theory to MR-WMNs [DK88]. The motivation behind cliques is to identify the conflicting links (see Fig. 44).

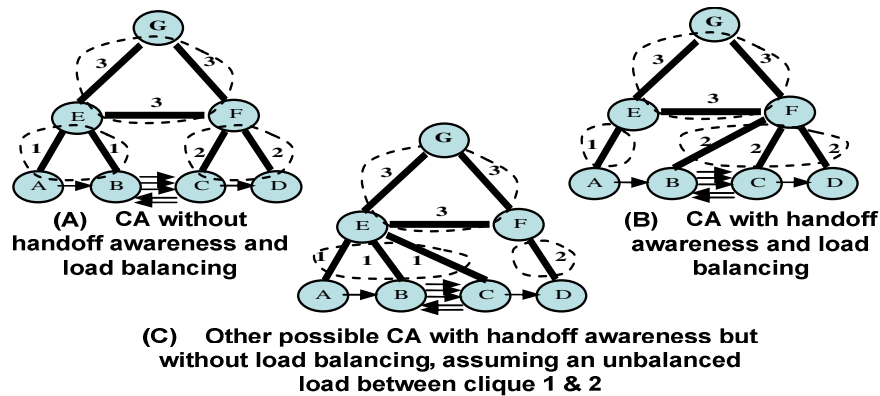


Fig. 44. An example of MR-WMN CA schemes

Therefore, two logical links (i, j) and (p, q) that belong to the same clique must not be active simultaneously. In this paper, a clique is defined as a set of links sharing the same channel and having a pairwise distance smaller than or equal to the minimum radio transmission range of the pair links. A maximal clique is a clique that belongs to no other clique. We use maximal cliques to determine the nodes which compete to access the same channel. In this paper, clique and maximal clique are used interchangeably. Obviously, a handoff user is better off using the same clique during

the handoff process; in this case, no path re-routing involving MRs beyond the first hop is performed. This is desired because path re-routing involving other hops beyond the first hop MR generates considerable overhead. This can be critical to real-time flows, such as voice over IP.

Fig. 44-A shows a sample MR-WMN that consists of 3 cliques where a number of MCs are performing handoffs (due to mobility) between B and C. We use the algorithm reported in [DK88] to compute cliques for the 7-nodes MR-WMN. A CA scheme without handoff awareness and without load balancing, called W.O.HALB, has no impact on the clique's formations regarding handoff and load dynamics (Fig. 44-A). In Fig. 44-B, we illustrate clique formation that provides best load balancing between clique 1 and clique 2; the load of a clique is the aggregate traffic load generated by MRs composing the clique. Fig. 44-B illustrates the clique formation (using HALB) that places the highest inter-MR handoff traffic inside one of the cliques. Fig. 44-C shows another clique's formation under HA, i.e. HALB without load balancing.

5.3.4 Problem Formulation

We formulate the CA problem as a multi-objective optimization model. Table 11 shows the notations used to describe the model.

Table 11. Notations parameters and variables

N	The set of MRs in the network
R	The set of radio interfaces on each MR
NC	$NC = \{1..K\}$, the set of κ channels, $ NC = K$
TS	$TS = \{TS^1, \dots, TS^s\}$, $s < N $, the set of traffic spots (positions of traffic concentrations in the service area)
G	The set of gateways
λ_i	The traffic generated by TS^i to the Internet
V_m	$V_m = \{MR^1, \dots, MR^m\}$, $m < N $, the set of MRs with access point functionality
ψ^h	The handoff threshold which is measured by MR to trigger the CA.
Q_k	$Q_k = \{q_1^1 \dots q_n^K\}$, the set of cliques with K channels
h_{ij}	The handoff traffic from MR^i to MR^j , $MR^i, MR^j \in V_m, i \neq j$
δ_{ik}	The traffic load on channel k of $MR^i \in V_m$
t_i	The traffic load of $MR^i \in V_m$

b_{ij}	A binary connectivity parameter that assumes 1 if two MRs, $MR^i, MR^j \in N, i \neq j$ are connected via a wireless link; 0 otherwise
z_{ij}	A binary activation parameter that assumes 1 if a traffic flow exists between $MR^i, MR^j \in N, i \neq j$; 0 otherwise
r_i	The transmission range of MR^i where $(d(i, j) \leq \min(r_i, r_j) \Rightarrow b_{ij} = 1)$
$x_i^{q_j^k}$	A binary variable = 1 if $MR^i \in q_j^k$; 0 otherwise
$\varphi_i^{q_l^k}$	A binary variable = 1 if channel k is assigned to clique q_l^k
$N(q_i^k)$	The set of the neighboring cliques $q_j^l \in Q_k$ in the transmission range of q_i^k .
y_i^k	A binary variable that assumes 1 whenever $MR^i \in V_m$ has interface radio with $MR^j \in N$ which is assigned to channel k ; 0 otherwise
l_i^k	A binary variable that assumes 1 whenever $MR^i \in V_m$ has interface radio with TS^i which is assigned to channel k ; 0 otherwise
f_{ij}^k	The traffic flow routed from MR^i to MR^j on channel $k, k \in NC, MR^i, MR^j \in N, i \neq j$
η_i	The traffic load of clique $q_i^k, k \in NC$
C_{ij}	The capacity of link (i, j)

We suppose initially that the MRs operate using the same number of radios u , each with K channels ($K > u$). K can be 12 non-overlapped channels when using IEEE 802.11 a and 3 non-overlapped channels when using IEEE 802.11 b.

The objective of our model is to minimize the inter-cliques handoff, i.e. handoffs occurring between different neighboring cliques which generally operate on different channels, while minimizing the traffic load variation between them. Furthermore, our proposed model maximizes the overall throughput on the network and ensures fair share of throughput among MCs. Accordingly, our CA model is formulated as follows:

Objective functions

$$\text{Min} [\sum_{q_i^k \in Q_k, q_j^k \in N(q_i^k), q_i^k \neq q_j^k} |\eta_{q_i^k} - \eta_{q_j^k}|] \quad (39)$$

$$\text{Min} [\sum_{\substack{i \in q_m^k, j \in q_l^k \\ q_m^k \neq q_l^k}} h_{ij} x_i^{q_m^k} x_j^{q_l^k}] \quad (40)$$

$$\text{max} (\sum_{i \in N \setminus \{V_m\}} \frac{f_{ig}^k}{C_{ig}}), \forall k \in NC, \forall g \in G \quad (41)$$

$$\max(\mu) \tag{42}$$

Subject to constraints

$$\sum_{j \in V_m} f_{ij}^k \leq \lambda_i, \forall i \in TS, \forall k \in NC \tag{43}$$

$$\sum_{j \in TS} f_{ji}^k + \sum_{k' \neq k} \delta_{ik'} + \sum_{v \in N} f_{vi}^{k'} = \sum_{v \in N} f_{iv}^{k'}, \forall i \in V_m, \forall k, \tag{44}$$

$$\sum_{j \in N-1} f_{ji}^k = \sum_{j \in N-1} f_{ij}^k, \forall i \in N \setminus \{V_m\}, \forall k \in NC \tag{45}$$

$$\sum_{j \in N-1} f_{jg}^k \geq \sum_{i \in TS} \lambda_i, \forall k \in NC, \forall g \in G \tag{46}$$

$$\sum_{k \in NC} y_i^k = 1, \forall i \in N-1 \tag{47}$$

$$\sum_{k \in NC} l_i^k = 1, \forall i \in N-1 \tag{48}$$

$$\sum_{q_j^k \in Q_k} x_i^{q_j^k} = 1, \forall i \in N-1 \tag{49}$$

$$y_i^k \geq x_i^{q_j^k}, \forall q_j^k \in Q_k, \forall i \in V_m \tag{50}$$

$$l_i^k + y_i^k \leq 1, \forall k \in NC, \forall i \in V_m \tag{51}$$

$$x_i^{q_l^k} + x_j^{q_l^k} + b_{ij}^{q_l^k} \leq 2 * z_{ij}^{q_l^k} + 1, \forall q_l^k \in Q_k, \forall i, j \in \tag{52}$$

$$x_i^{q_l^k} + x_j^{q_l^k} + b_{ij}^{q_l^k} \geq 3 * z_{ij}^{q_l^k}, \forall q_l^k \in Q_k, \forall i, j \in V_m \tag{53}$$

$$\sum_{k \in NC} \sum_{i, j \in N} z_{ij}^{q_l^k} = 1, \forall q_l^k \in Q_k \tag{54}$$

$$z_{ij}^{q_l^k} \leq M \times f_{ij}^k, \forall i, j \in N, \forall M \in R^+ \tag{55}$$

$$M \times z_{ij}^{q_l^k} \geq f_{ij}^k, \forall i, j \in N, \forall M \in R^+ \tag{56}$$

$$\sum_{j \in N-1} b_{ij}^{q_l^k} \geq \sum_{k \in NC} x_i^{q_j^k}, \forall i \in V_m, \forall q_l^k \in Q_k \tag{57}$$

$$\sum_{j \in N-1} b_{ij}^{q_l^k} \leq \sum_{j \in N-1} \sum_{k \in NC} x_j^{q_l^k}, \forall i \in V_m, \forall q_l^k \in Q_k \tag{58}$$

$$\sum_{k \in NC} z_{ij}^{q_l^k} \leq b_{ij}^{q_l^k}, \forall i, j \in N, \forall q_l^k \in Q_k \tag{59}$$

$$\eta_{q_j^k} = \sum_{i \in V_m} t_i x_i^{q_j^k} + \sum_{j \in TS_i \in V_m} f_{ij}^k, \forall k \in NC \quad (60)$$

$$\sum_{k \in NC} \delta_{ik} = t_i, \forall i \in V_m \quad (61)$$

$$\sum_{k \in NC} \varphi^{q_i^k} = 1, \forall q_i^k \in Q_k \quad (62)$$

$$x_i^{q_i^k} \leq \varphi^{q_i^k}, \forall q_i^k \in Q_k, \forall k \in NC, \forall i \in N \quad (63)$$

$$\sum_{j \in N \setminus \{V_m\}} \sum_{q_i^k} b_{ij}^{q_i^k} = 1, \forall i \in V_m \quad (64)$$

In our model, the objective function (39) minimizes the traffic load variance between all neighboring cliques to balance the load. The objective function (40) minimizes the handoff frequency between two MRs belonging to different cliques. We assume that the handoff traffic between MRs is known a priori when MHALB/MTABU/HALBTH are performed; we set the handoff traffic values (h_{ij}) according to extensive simulations performed using ns-2. In Section 5.5, we consider the case where the handoff traffic is not known a priori and changes dynamically. The objective function (41) maximizes the total throughput by maximizing the sum of all flow-capacity ratios over all wireless links in the MR-WMN. The objective function (42) aims to fairly supply all MCs (TSs) by maximizing the Jain's fairness index (μ) [JC84]. We express μ as follows:

$$\mu = \left(\sum_{i=1}^s TH_i \right)^2 / \left(s \sum_{i=1}^s TH_i^2 \right) \quad (65)$$

where TH_i is the throughput generated by TS^i and s is the number of TSs and $0 \leq \mu \leq 1$. We express TH_i as follows:

$$TH_i = \sum_{j \in N} f_{ij}^k, \forall i \in s, \forall k \in NC \quad (66)$$

Ideally, all TSs have access to the same throughput; in this case, μ is equal to 1. In the worst case, μ becomes $1/s$, i.e. the inverse of the number of TSs.

Constraints (43)-(46) define the flow balance equations for each MR. Constraints (47) and (48) prevent a MR from selecting the same channel more than once to assign it to its interfaces. Constraint (49) states that the MR with IEEE 802.11 access point functionality (i.e., MR interfacing with MCs) belongs to one clique; this constraint ensures that every node, except gateways, has to belong to one or more cliques containing nodes with access point functionality. Constraint (50) states that node $j \in V_m$ has an interface configured with the same channel as the new clique to which it will move. Constraint (51) states that the number of links from a node is limited by the number of its radio interfaces; it also states that the channel assigned to the interface interfacing with MCs (TSSs) must be different from the channels assigned to the other interfaces of the MR. Constraints (52) and (54) prevent link interferences and allow only one link to be active in the same clique. Constraints (55) and (56) ensure that if a traffic flow exists between two nodes, then the link, between these 2 nodes, is active; M is a positive big value. Constraints (57) and (58) ensure that the new node added to clique q_i^k is connected to all nodes in this clique. Constraint (59) ensures that a link is active only if it exists. Constraint (60) computes traffic load in each clique. Constraint (61) sums the total traffic load for all the MR's channels. Constraints (62) and (63) ensure that all nodes in a clique use the same channel. Constraint (64) forces each node with access functionality to connect to at least one MR located on a higher level (e.g. MRs A, B, C and D must each be connected to E or/and F as shown in Fig. 45).

The objectives functions (39), (40) and (42) are not linear. In the following, we propose a method to linearize them. We modify the objective function (39) to

$$\text{Min } w \quad (67)$$

Subject to the following constraint:

$$|\eta_{q_i^k} - \eta_{q_i^{k'}}| \leq w \Rightarrow \begin{cases} w \geq \eta_{q_i^k} - \eta_{q_i^{k'}} \\ \text{and} \\ w \geq -\eta_{q_i^k} + \eta_{q_i^{k'}} \end{cases}, \forall q_i^k \in Q_K, q_i^{k'} \in N(q_i^k) \quad (68)$$

In order to linearize the objective function (40), we define a new variable

$$T_{ij}^{q_m^k q_l^k} = \begin{cases} 1, & \text{if } X_i^{q_m^k} = 1 \text{ and } X_j^{q_l^k} = 1 \\ 0, & \text{otherwise} \end{cases} \quad ; \text{ the objective function (40) becomes:}$$

$$\text{Min}([\sum_{\substack{i \in q_m^k, j \in q_l^k \\ q_m^k \neq q_l^k}} h_{ij} T_{ij}^{q_m^k q_l^k}]) \quad (69)$$

Subject to the following constraints:

$$T_{ij}^{q_m^k q_l^k} + 1 \geq X_i^{q_m^k} + X_j^{q_l^k}, i, j \in V_m, q_m^k, q_l^k \in Q_k \quad (70)$$

$$T_{ij}^{q_m^k q_l^k} \leq X_i^{q_m^k}, i, j \in V_m, q_m^k, q_l^k \in Q_k \quad (71)$$

$$T_{ij}^{q_m^k q_l^k} \leq X_j^{q_l^k}, i, j \in V_m, q_m^k, q_l^k \in Q_k \quad (72)$$

Since it is difficult to linearize the objective function (42), we approximate μ as follows:

$$\mu = \max TH_i - \min TH_i, \quad (73)$$

where

$$\max TH_i \geq TH_i \text{ and } \min TH_i \leq TH_i \quad (74)$$

In this case, the objective function (73) consists of minimizing μ .

Constraints (68), (70)-(74) ensure that the objective functions (39), (40) and (74) are linear and could be used during the resolution of the proposed model.

5.4 Solving The Model

In the previous section, we proposed a multi-objective model to solve the CA problem. However, for sake of simplicity (solving the problem using pure multi-objective optimization methods is for future work), we convert it to an aggregated form using a single objective function defined in equation (75):

$$\begin{aligned}
& \text{Min } (\alpha_1 w + \alpha_2 [\sum_{\substack{i \in Q_m^k, j \in Q_l^{k'} \\ Q_m^k \neq Q_l^{k'}}} h_{ij} T_{ij}^{q_m^k q_l^{k'}}] - \alpha_3 [\sum_{i \in N} \frac{f_{ig}^k}{C_{ig}}] \\
& + \alpha_4 [\max TH_i - \min TH_i]) \quad (75)
\end{aligned}$$

where $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$, $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are positive weight coefficients and $w = \max_{q_i^k \in Q_K, q_j^{k'} \in N(q_i^k), q_i^k \neq q_j^{k'}} |\eta_{q_i^k} - \eta_{q_j^{k'}}|$.

The determination of the ‘‘optimal’’ values of $\alpha_1, \alpha_2, \alpha_3$ and α_4 is out of scope of this paper; in future work, we will investigate the use of the analytical model proposed in [Ne08] to determine these values.

The problem of optimal channel CA is known to be NP-hard [RG04]. Preliminary experiments showed that we could not obtain exact solutions for this problem when dealing with real-size networks. We therefore decided to use heuristic algorithms to obtain near-optimal solutions for larger instances. Due to the high cost of evaluating a multi-objective function, heuristics with fast descent and local minima escape strategy are preferred to accelerate the search of near-optimal CA solutions. For that purpose, we use variable neighborhood search (VNS)-based heuristic to solve our CA multi-objective problem. Our proposed algorithm, MHALB, illustrated in Algorithm 7 operates in two phases: (1) It first produces an initial/feasible solution using a simple greedy search algorithm, called HALBTH (see Algorithm 6); and then (2) it improves the initial/feasible solution using a VNS meta-heuristic.

5.4.1 Heuristic Based on Greedy Algorithm

HALBTH is used to provide an initial CA solution to the proposed optimization model. First, we determine the connectivity graph of the MR-WMN, and arbitrarily assign channels to interfaces of MRs. Then, we compute cliques [22]. Finally, we run the greedy search algorithm illustrated in Algorithm 6. The algorithm, carefully re-assigns channels to interfaces with the objective of confining most of the user handoff traffic inside a clique, thus preventing re-routings involving MRs beyond first hop MRs (of pre-handoff paths) and consequently limiting handoff overhead. The greedy algorithm determines an initial solution that is certainly sub-optimal; in fact, it does not

consider all objectives of the proposed model. In the next section, we present a meta-heuristic that improves the initial solution to produce a near-optimal solution. The pseudo code of HALBTH is shown in Algorithm 6.

Algorithm 6. Heuristic HALBTH algorithm

Input: connectivity graph, static CA, cliques

Output: new topology with new CA

Step0: initialization

Build an initial/feasible solution as follows: (1) Assign channels statically; (2) determine connectivity graph; and (3) Compute cliques; handoff=1;

While ($handoff = 1$) /* user handoff exists between $i, j \in V_m$ */

Step1: building a sorted list of handoff rates

For every $i, j \in V_m$ and $i \in q_l^k, j \in q_m^{k'}, k \neq k'$

 Compute h_{ij} ; Create a set $S = \{h_{ij} \mid h_{ij} \geq \psi^h\}$;

If $S == \emptyset$ **then** $handoff == 0$;

Else Sort elements of s in descendent order;

Step2: greedy search heuristic for near optimal solution

 Take first value of S , h_{ij} associated with $i, j \in V_m$;

 Generate two solutions as follows:

 (1) Move i to clique $q_m^{k'}$; Switch node i to channel k' ;

 (2) Move j to clique q_l^k ; Switch node j to channel k ;

 Evaluate the two solutions according to the objective function (75) and choose the optimal solution.

EndWhile

5.4.2 A Variable Neighborhood Search Method

To improve the initial solution produced by HALBTH, we propose to use a VNS meta-heuristic, called MHALB. Each neighborhood is explored using first-improvement local search; this means that we stop to search for better solutions in the current neighborhood when the first-improvement is found. The basic idea behind the VNS meta-heuristic is a systematic change of neighborhoods performed by the local

search algorithm. Indeed, the frequent change of neighborhoods as soon as a first-improvement is found accelerates the descent to local minimums regardless of its quality in terms of improvements. This is desirable in our case since the set of the neighborhood solutions generated by each elementary move is very large and evaluating all these solutions is very expensive in terms of time (see Section 5.4.2.4). Generally, VNS proceeds by a descent method to a local minimum, and then explores, systematically, increasingly distant neighborhoods of this solution. The best order of applying different moves is out scope of this paper; we will investigate this issue in future work. In the proposed meta-heuristic, we consider two types of movements (1) node move; and (2) cliques exchanges.

5.4.2.1 Node move

A node is removed from its current clique and inserted in another clique that contains at least one mesh node interfacing with a MC. All possibilities are systematically tried until an improvement is found; an improvement corresponds to a smaller value for the objective function (75). We repeat this procedure for all MRs until a local optimum is reached. Our VNS-based heuristic jumps from the current neighborhood to a new neighborhood if and only if a better solution has been found.

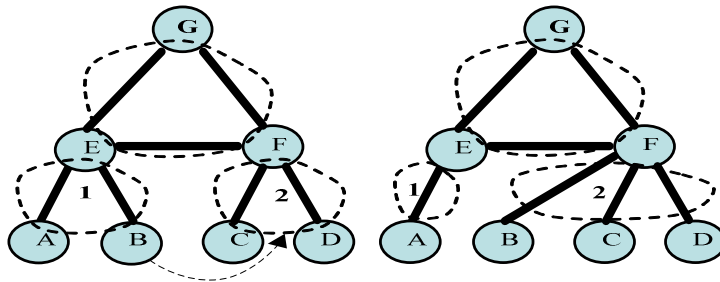


Fig. 45. Node move

A sample node move is illustrated in Fig. 45; node B moves to clique (FCD) by re-assigning it channel (2); this move produces two new cliques: (EA) and (FBCD). We repeat this procedure for all nodes in different cliques; for example a next move, in our example, may be moving node A to clique (FCD) by re-assigning it channel (2); this produces two new cliques: (EB) and (FACD).

5.4.2.2 Cliques exchange

Given a pair of cliques that contains mesh nodes interfacing with MCs, two of these nodes are exchanged. All possible exchanges on every pair of cliques are considered until an improvement is found; an improvement corresponds to a smaller value for the objective function (75). We repeat this procedure for all possible cliques until a local optimum is reached.

A sample exchange is illustrated in Fig. 46; we do two movements at the same time (two nodes are re-assigned): node B switches to channel (2) and node (D) switches to channel (1).

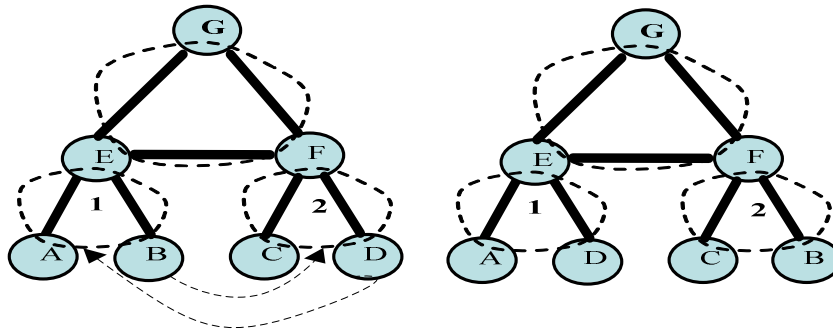


Fig. 46. Cliques exchanges

5.4.2.3 MHALB

The pseudo of MHALB, using the two types of moves (see Section 5.4.2), is shown in Algorithm 7 (step 0-2) where: (1) N1 is the node move neighborhood; (2) N2 is the cliques exchanges neighborhood; and (3) S^j is a local optimum solution based on neighborhood j . The neighborhood moves in MHALB correspond to moves from N1 to N2 and vice-versa. MHALB stops when it improves the initial solution by $\beta\%$ (or a predefined timer expires or a maximum number of iterations is reached). The value of β is predefined by the network operator.

Algorithm 7. Meta-heuristic: MHALB algorithm

Step0: initialization

Create an initial solution S^0 using the greedy algorithm HALBTH (Algorithm 6);

$$f^0 \leftarrow f(S^0);$$

Step1: node move

Perform a local descent based on N1;

Let S^1 be the local optimum obtained;

If S^1 is different from S^0 **then** $S^0 \leftarrow S^1$;

Step2: cliques exchanges

Perform a local descent based on N2;

Let S^2 be the local optimum obtained;

If S^2 is different from S^0 **then** $S^0 \leftarrow S^2$;

If $f(S^2) < (1-\beta)f^0$ **then** exit;

go to step 1;

5.4.2.4 Algorithm Complexity

$O([h \times m \log m + (R-1) \times V_m + 2(R-1) \times V_m] \times C_F)$ is the complexity of MHALB where h is the number of handoffs in the network, m is the size of the list S , C is the number of cliques in the network, s is the number of TSs, and $C_F = O(\alpha_1 C^2 + \alpha_2 C^2 + \alpha_1 N^2 V_m^2 + \alpha_4 s N^2)$ is the complexity of the multi-objective function (Eq. (75)).

5.4.3 A Tabu Search Algorithm

Our proposed Tabu search meta-heuristic called MTABU, uses two types of moves: N1 and N2 (see Section 5.4.2). The pseudo-code of MTABU is shown in Algorithm 8. In MTABU, we start with an initial solution S^0 using the greedy algorithm HALBTH ($f^0 \leftarrow f(S^0)$). Starting from S^0 , we create a sequence of solutions $S^0, S^1 \dots S^i, \dots$, in attempt to reach a near-optimal solution according to the objective function (Eq. (75)).

The i^{th} iteration: In the i^{th} iteration ($i \geq 0$) of this algorithm, we create the next S^{i+1} in the sequence from S^i ; more specifically, we generate all N1 neighborhood solutions of S^i and we pick the solution S^{i+1} that has the lowest objective function (Eq. (75)) $f(S^{i+1})$. In order to escape local minima, the algorithm does not terminate even if $f(S^{i+1})$ is bigger than $f(S^i)$.

After a maximum number of allowed iterations, denoted by i_{max} , without improvement exploring N1, we jump to N2 neighborhood. The exploration of N2 is performed using search intensification based on N2. This search is done by a simple descent as shown in Algorithm 8 (step 2). The idea behind the concept of search intensification is that, as an intelligent human being would probably do, one should explore more thoroughly the portions of the search space that seem promising to make sure that the best solutions in these areas are indeed found.

Therefore, to make MTABU more effective, it is important to add throughout the search an intensification step (i.e, Algorithm 8-step2) without sampling; thus, the best solutions available in the various regions of the search space explored by MTABU will be found and recorded.

Tabu-List: To prevent cycling, i.e. exploring the same movements, we maintain a FIFO tabu-list Ω of limited size. If solution S^{i+1} is created from S^i using (movement, S^i) by assigning a new channel to the MR interface (applying N1), then we will add (movement, S^i) to Ω . When we generate solutions based on N1 movement, we ignore solutions computed using the same movement to get S^i in Ω . The size of Ω is denoted by SI. Each movement is kept in the Tabu list during SI iterations before being reconsidered again in the N1 exploration. It is also possible to assign a different number of iterations, chosen randomly, to each movement added to Ω .

Termination: We keep track of the best (i.e., with lowest objective function) solution (S^{best}) seen so far by the algorithm. The step 1 (the exploration of

neighborhood N1) terminates when a maximum number of allowed iterations i_{\max} was reached without any improvement in $f(S^{best})$ (the most commonly used criterion).

Algorithm 8. Meta-heuristic: MTABU algorithm

Step0: initialization

Create an initial solution S^0 using the greedy algorithm HALBTH (Algorithm 6);
 $f^0 \leftarrow f(S^0); S^{best} \leftarrow S^0; f^{best} \leftarrow f^0; \Omega = null; i \leftarrow 0; k \leftarrow 0;$

Step1: Tabu search based on N1 : node move

While $i \leq i_{\max}$ **Do**

Generate solutions of f^k based on N1;

Let S^{k+1} be the next solution with lowest objective function;

Add *movement*, S^k to Ω ; **If** Ω is full **then** delete its oldest entry;

If $f(S^{k+1}) < (1-\beta) f^{best}$ **then** $f^{best} = f^{k+1}; i \leftarrow 0;$

Else $i \leftarrow i + 1;$ **Endif**;

$k \leftarrow k + 1;$

End While; **return** f^{best} ;

Step2: Search intensification based on N2 :cliques exchange

A simple descent;

If $f(\tilde{S}) < (1-\beta) f^{best}$ **then** $f^{best} = f(\tilde{S});$

return f^{best} ;

5.5 Proposed On-Line Channel Assignment Scheme

In the previous section, we presented CA heuristics which are CPU intensive since they apply to the whole WMN using centralized offline algorithms. In this section, we present a distributed online CA algorithm that improves the offline CA with respect to traffic dynamics and handoff traffic variations. To deal with these traffic variations, we assume the existence of an online measurement mechanism that measures traffic/handoff variations and computes a suitable time interval, denoted by T , to capture periods with minimal traffic/handoff variations.

5.5.1 Channel Assignment Trigger

The proposed algorithm re-assigns channels (i.e., reconfigures radio interfaces) in real-time while traffic may be flowing through these interfaces; this may cause

service disruption (data losses) during CA. To minimize disruption, our proposed scheme prioritizes intra-clique handoffs over inter-clique handoffs. Therefore, when a MR experiences an inter-clique handoff rate exceeding a fraction ψ^h of its traffic load, the online CA is triggered. The same channel, i.e. and thus the same clique, is assigned to MRs interfaces involved in the frequent handoff rate: we assign the clique's channel of one's interface to the other's interface. However, the choice of which MR will change its clique is not trivial since the network performance can be impacted by the choice of the re-assignment. Our scheme chooses the re-assignment option that minimizes the load variance among neighboring cliques in a distributed manner and the traffic load in each clique and its neighboring cliques are computed as in [RH08c]. Note that our scheme dynamically adjusts the handoff rate threshold ψ^h , i.e. the handoff rate above which the online CA is triggered. The adjustment is triggered when a measured performance parameter, such as packet loss, exceeds a given threshold. For instance, if measured packet loss, collected from interface statistics during one or more time intervals (i.e., one or more T), exceeds a given QoS level packet loss, then the handoff rate threshold is decreased and it can be regularly adjusted (increased/decreased) at the end of each time interval.

5.5.2 Online Channel Assignment Algorithm

Algorithm 9 details the assignment of channels and thus the adjustment of cliques depending on the inter-clique handoff rate. We assume that the online CA is re-executed periodically each period of time T indicating a considerable traffic variation; during T no noticeable traffic variation occurs. In this algorithm, the handoff rates between two MRs is denoted by h_{ij} which is equal to the number of handoffs occurring between nodes i and j during the period of time T . The parameter traffic load, which is continuously measured on each MR, is used as criterion to select a unique CA solution when several ones are possible. This is done when the cliques involved in the online CA, i.e. cliques having a handoff rate between their respective MRs exceeding the threshold ψ^h , have different variances of traffic load.

Algorithm 9. Dynamic distributed CA algorithm : HALB

Input: handoff rate h_{ij} , traffic load η_i of clique q_i^k

Output: New topology with new CA

- Compute traffic load in each clique [RH10d];
 - Exchange traffic load information between neighboring cliques [RH10d];
 - Compute the variation of traffic load in clique q_i^k and the other neighboring cliques for $\forall j \in q_m^{k'}$;
 - Assign the channel of the least loaded neighboring clique to the interface of the MR belonging to the most loaded neighboring clique;
 - The handoff statistics h_{ij} are recorded during time period for each interface;
-

The pseudo-code of the MHALB⁺ (resp. MTABU⁺), which includes MHALB, HALBTH and HALB (resp. MTABU, HALBTH and HALB), is shown in Algorithm 10.

MHALB⁺, MTABU⁺ and HALBTH⁺ are distributed CA algorithms that continuously improve the near optimal solution produced by the static heuristics MHALB/MTABU/ HALTBTH (step 0). These algorithms adapt to the dynamic variation of the traffic load and the handoff rates. In fact, they make use of handoff statistics, that are collected during the period of time T , to trigger or not CA (see Section 5.5.1) at the end of the time period.

Algorithm 10. Distributed dynamic CA algorithm HALBTH⁺/MHALB⁺ /MTABU⁺

Step 0:

Apply the heuristics HALBTH/MHALB/MTABU to get near optimal solution.

Step1 : Online CA improvement regarding dynamic traffic variation

If $\sum_{j \in V_m} h_{ij} \gg \psi^h, i \in q_l^k, \forall j \in q_m^{k'}, k \neq k'$ **then**

Call Algorithm 9 : HALB at each Mesh Router

End if

Repeat Step 1 Every period of time T

5.6 Simulation Results

We implemented MHALB, MTABU and HALBTH using C++ language and we executed them on a machine having a Pentium 2.1 GHz CPU, 2.0 GB of RAM and Fedora 11. To assess the quality of the solutions returned by MHALB and HALBTH, we compare them to CPLEX solutions [CP]; CPLEX provides exact solution (i.e., the optimal solution) of the proposed model but only for small size networks; for realistic size networks, it does not return a solution. Using the solution produced by MHALB and the optimal solutions returned by CPLEX, we set up and configured a simulation model (see below) and conducted simulations using ns-2 [NS] in order to evaluate the performance of MR-WMN in terms of delay, packet loss, throughput and fairness under the proposed CA scheme. We use a topology composed of 10/13/19 MRs respectively with our proposed schemes. The radio transmission range r takes one of the following values: 150 meters, 200 meters and 250 meters and the transmission interference R of each wireless MR is 550 meters. Also, we fix the MR-WMN packet size to 1000 KB. The voice call arrival process follows a Poisson distribution and the packet arrival process follows a deterministic distribution with a constant rate (CBR) with rate of 1Mbps and interval of 0.01 sec. The average number of simulation runs performed for each result, shown in the figures below, is 10. The throughput is defined as the average rate of successful packet delivery over a communication channel. The delay is measured as the mean MC-to-Gateway delay over all flows in the network. We used a maximum number of iterations equal to 10 for the meta-heuristic based on VNS to move to/from N_1 or N_2 resp. The same number of iterations is used for Tabu search to move from step1 to the search intensification step. The handoff threshold which is measured by MR to trigger the CA is equal to 55% (for our simulation setup). These values are chosen based on extensive simulations using ns-2 (e.g, we found that a number of iterations bigger than 10 does not improve the results).

In our performance evaluations (case I-II-III) of the MR-WMN under different CA schemes, we choose different combinations of objectives in the aggregate function (Eq. (75)). This will help evaluate the impact of the weights given to different objectives on the real performance metrics and then prove or disprove the assumed relations between them.

5.6.1 Case I: Known Handoff traffic

In our first simulation scenario, we evaluate the performance (delay, packet loss and throughput) of the MR-WMN using (1) HALBTH; (2) MHALB; and (3) CPLEX. In this scenario, using weights $(1/3, 1/3, 1/3, 0)$, we are interested in optimizing all objectives except fairness which will be evaluated later.

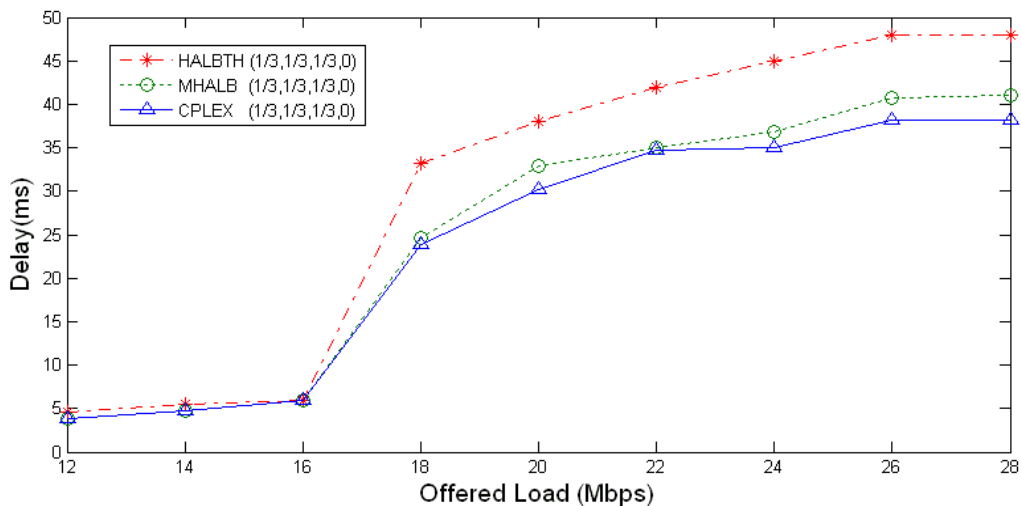


Fig. 47. Delay with HALBTH, MHALB and optimal solution

Fig. 47 shows that MHALB is able to reduce the delay by an average of 17% compared to HALBTH. The optimal solution (returned by CPLEX) outperforms, in terms of delay, MHALB and HALBTH solutions by 5% (i.e., $[\text{average_delay}(\text{MHALB}) - \text{average_delay}(\text{CPLEX})] / (\text{average_delay}(\text{MHALB}))$) and 20% (i.e., $[\text{average_delay}(\text{HALBTH}) - \text{average_delay}(\text{CPLEX})] / (\text{average_delay}(\text{HALBTH}))$) respectively.

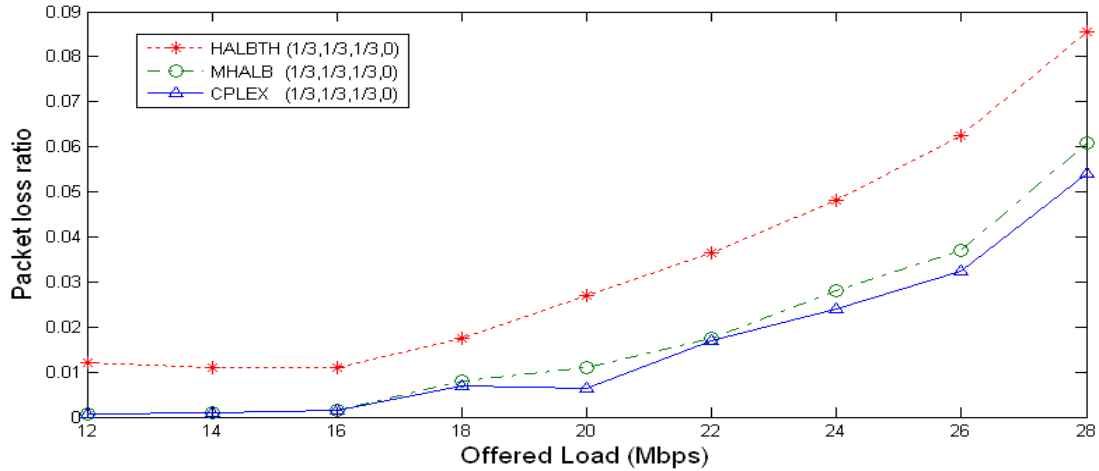


Fig. 48. Packet loss rate with HALBTH, MHALB and optimal solution

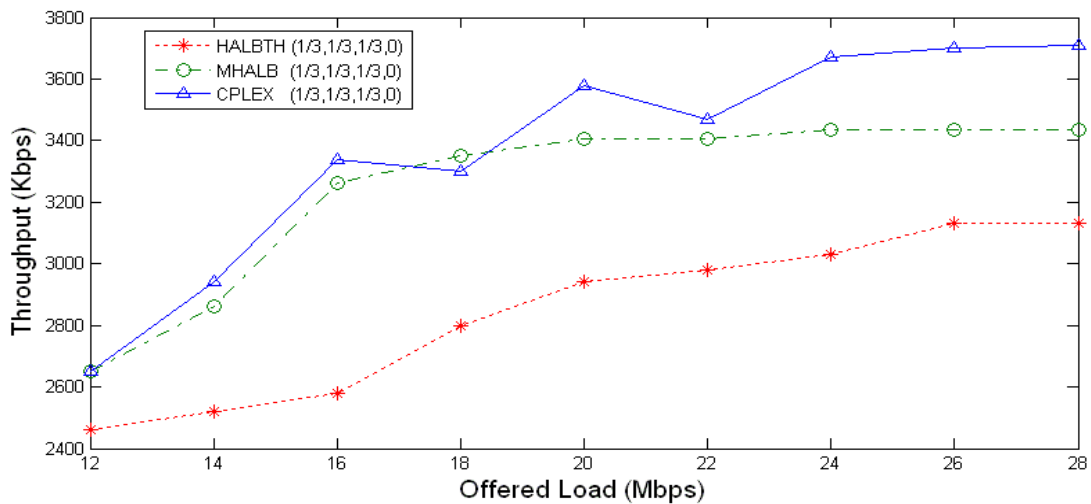


Fig. 49. Throughput with HALBTH, MHALB and optimal solution

Fig. 48 shows that MHALB significantly reduces the loss rate compared to HALBTH; in fact, MHALB reduces the mean loss rate (over all loads) by 47% compared to HALBTH. The optimal solution outperforms, in terms of loss rate, MHALB and HALBTH solutions by 12% and 54% respectively.

Fig. 49 shows that MHALB provides better WMN throughput than HALBTH; its outperformance is even more marked starting from load 16 Mbps. This can be explained by the fact that good load balancing capability leads to efficient usage of network resources and results in higher throughput in the network. We observe, for example, that the improvement provided by MHALB over HALBTH at load 16 is

about 20%. Fig. 49 shows also that the optimal solution outperforms MHALB by merely 4%.

Fig. 47, Fig. 48 and Fig. 49 respectively show that the MHALB provides an average performance, in terms of delay, packet losses and throughput, very similar to the performance provided by the optimal CA, i.e. the exact solution of the proposed model given by CPLEX; we conclude that MHALB provides high quality solutions (i.e., near optimal solutions). Note that MHALB returns a solution 50 times faster than CPLEX (see Table 12). For networks with more than 10 nodes, CPLEX does not return a solution in a reasonable time (starting from 13 nodes, it is more than 3 days).

Table 12. CPU time (sec) with HALBTH, MHALB and CPLEX: 10-nodes network

Network size: 10 nodes			
Load (Mbps)	HALBTH	MHALB	CPLEX
12	3	119	> 10000
14	12	133	> 10000
16	12	140	> 10000
18	12	137	> 10000
20	13	146	> 10000
22	14	149	> 10000
24	14	146	> 10000
26	15	282	> 10000
28	15	200	> 10000

In the second and the third simulation scenarios, we consider a network with more than 10 nodes. Similarly to the first scenario, CPLEX does not return a solution in

a reasonable time. Thus, we use exclusively MHALB (which returns near-optimal solutions as shown in Fig. 47, Fig. 48 and Fig. 49 to resolve the proposed model. In the second simulation scenario, we vary the weights assigned to the four objectives and we evaluate the performance of the MR-WMN in terms of delay, packet loss and throughput while varying the offered load.

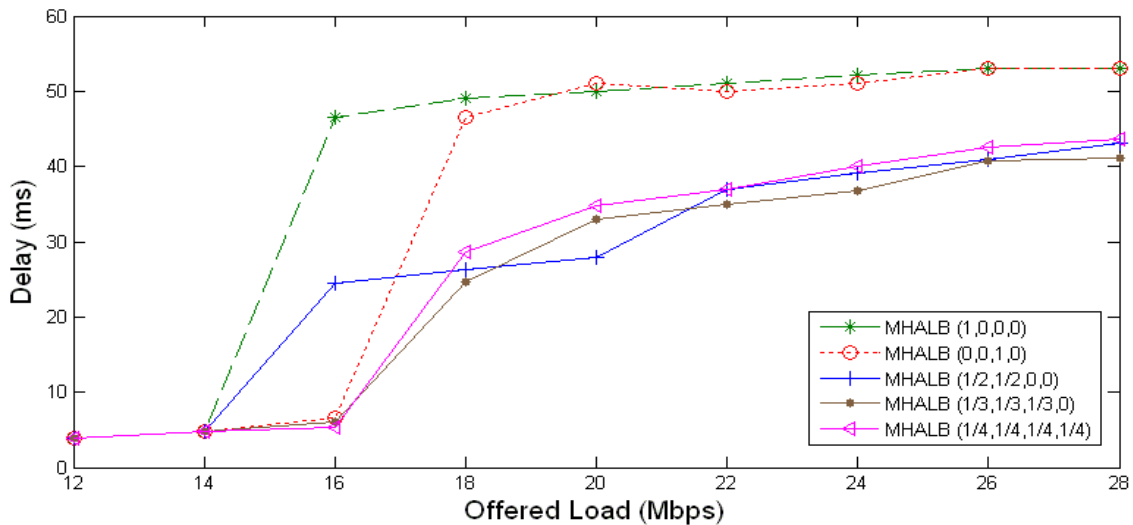


Fig. 50. Delay with MHALB

Fig. 50 shows that, compared to other multi-objective combinations not considering the handoff overhead, a combination that uses a non null value of α_2 ($1/4$ or $1/3$ or $1/2$) to minimize the handoff overhead decreases the MR-WMN delay by 20-25% in average under a moderate to high offered load. This can be explained by the fact that our handoff-aware scheme ($\alpha_2 > 0$) re-assigns carefully channels to MR interfaces so that most of the handoff traffic is re-routed by only changing first-hop MRs of original paths; this will definitely reduce delay compared to arbitrary re-routing that may involve second and third hop MRs.

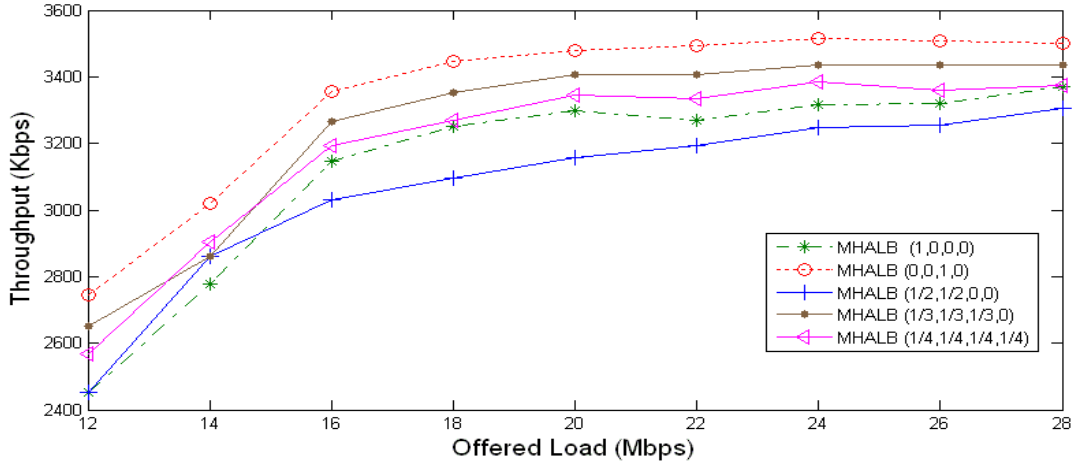


Fig. 51. Throughput with MHALB

Fig. 51 and Fig. 52 show that, considering throughput maximization ($\alpha_3 > 0$), load balancing ($\alpha_1 > 0$) and fairness ($\alpha_4 > 0$) in MHALB, we significantly decrease the packet loss (over 73% improvement in average). This can be explained by the tight relationship between all of these performance metrics. In fact, increasing the throughput, either directly or indirectly through load balancing, will automatically decrease packet losses.

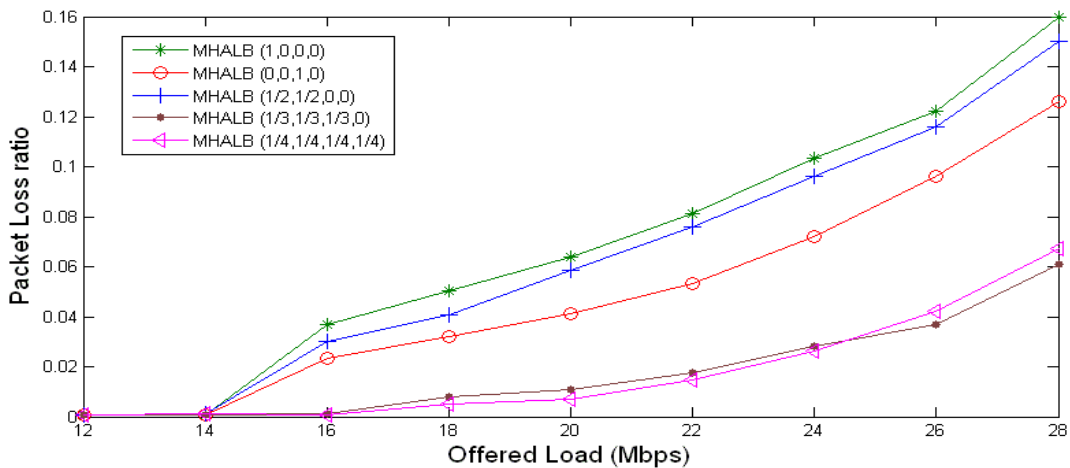


Fig. 52. Loss with MHALB

In the third simulation scenario, we vary the weights assigned to the four objectives and we observe the performance of the MR-WMN in terms of fairness and throughput while varying the offered load.

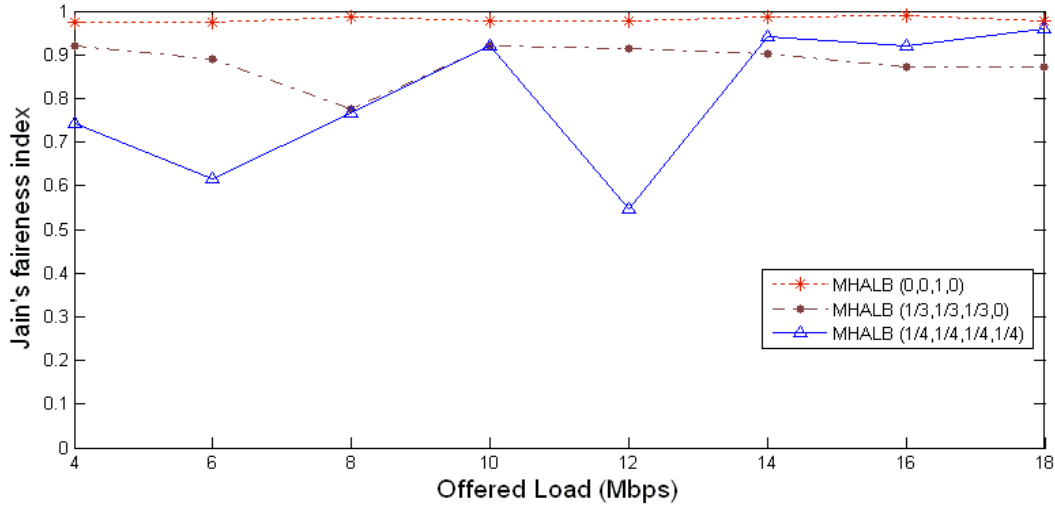


Fig. 53. Jain's fairness index

Fig. 53 shows that the Jain's fairness index (μ) is almost equal to 1 for MHALB (0, 0, 0, 1); this is desired to share the WMN throughput between the MCs in a fair manner without starving some of them. In fact, MHALB (0, 0, 0, 1) improves the mean μ (over all loads) by 19% and 10% compared to MHALB (1/3, 1/3, 1/3, 0) and MHALB (1/4, 1/4, 1/4, 1/4) respectively.

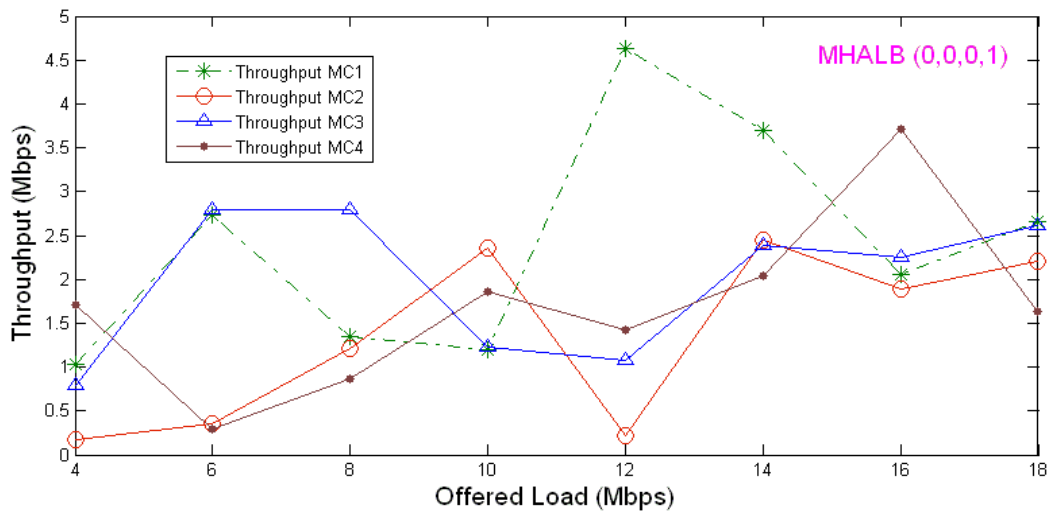


Fig. 54. Throughput achieved by MCs (0,0,0,1)

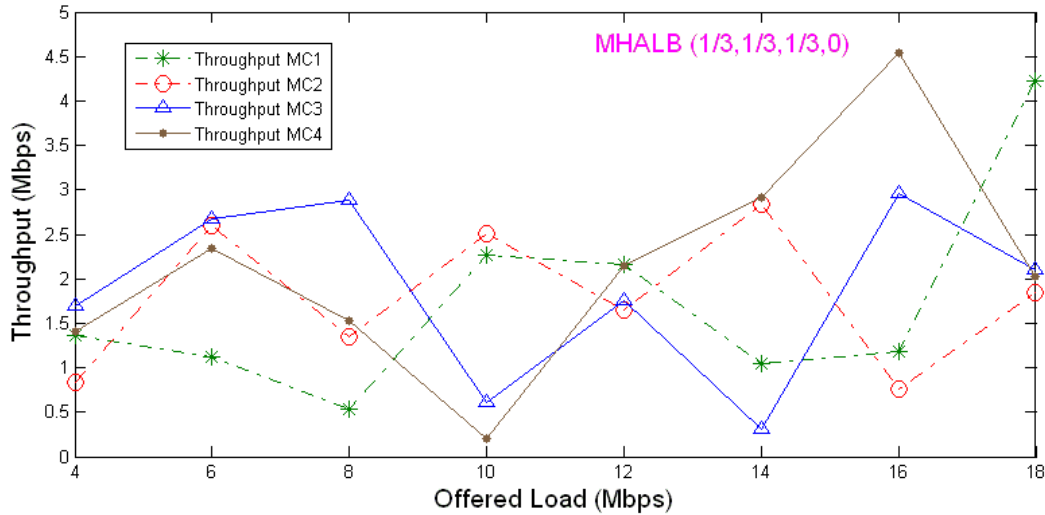


Fig. 55. Throughput achieved by MCs (1/3,1/3,1/3,0)

Fig. 54 and Fig. 55 show the throughput achieved by each MC in the MR-WMN. In Fig. 54, we observe that MCs throughput exhibits nearly the same trend when the offered load increases while in Fig. 55 the MCs throughput shows an irregular behavior; this proves the impact of the fairness on the underlying throughput which is also consistent with the results obtained in Fig. 53.

5.6.2 Case II: Dynamic CA with and without HALBTH

We compare first three schemes: (1) W.O.HALB: an existing CA in MR-WMN, such as in [SV04]; (2) HALB is our proposed CA scheme that dynamically re-assigns channels with one of the purposes of confining most of the handoffs in the same clique and balances the load among cliques as shown in Algorithm 9; (3) HA is the same as HALB CA scheme but without load balancing among cliques; and then we compare the two last schemes (4) HALBTH+ is the proposed CA heuristic based on handoff aware with load balancing as shown in Algorithm 10; and (5) W.O.HALBTH+ is the distributed CA dynamic scheme without good initial solution.

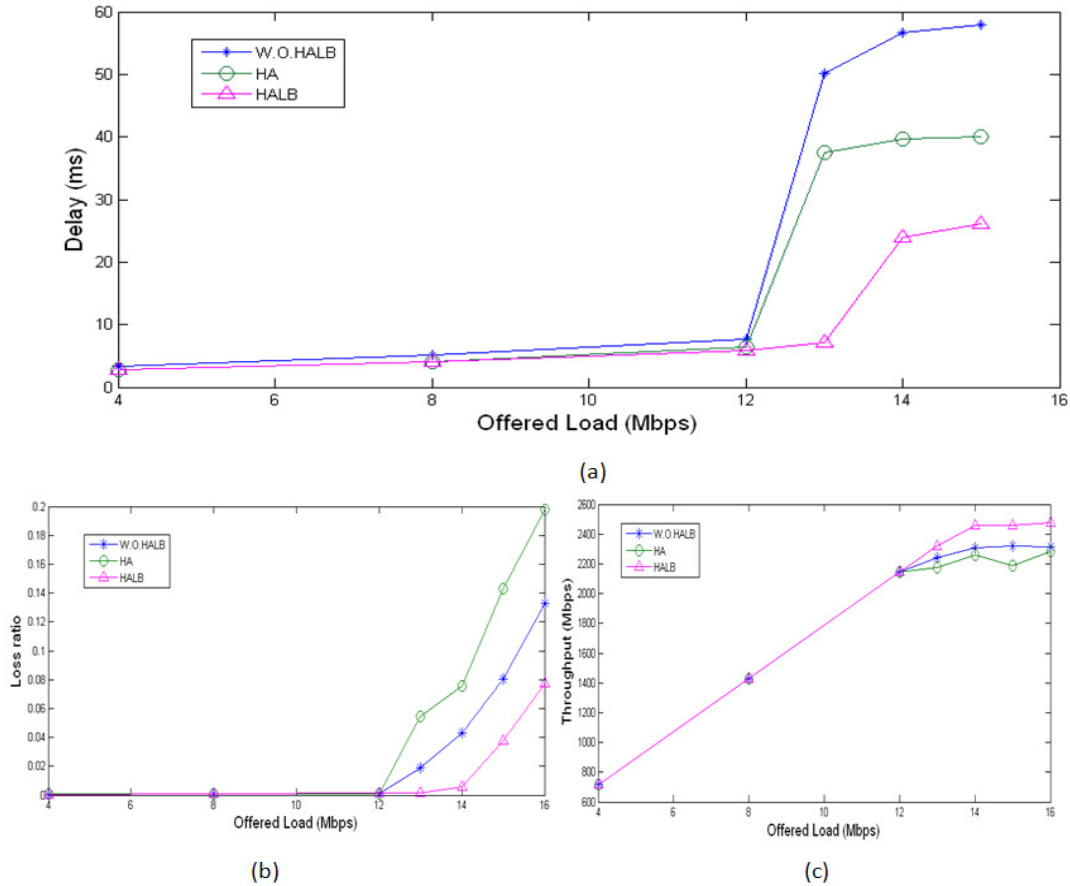


Fig. 56. Delay (a); packet loss ratio (b); throughput (c)

In Fig. 56, we evaluate the performance of the MR-WMN under our proposed CA schemes (HALB and HA) and we compare them to the existing CA scheme (W.O.HALB) that does not take into account handoff and load balancing.

Fig. 56-(a) shows that HALB provides the best result in terms of delay followed by HA; the worst results are generated by W.O.HALB. This can be explained by the fact that our scheme re-assigns carefully channels to MR interfaces so that most of the handoff traffic is re-routed by only changing first-hop MRs of original paths; this will definitely reduce delay compared with arbitrary re-routing (the case of W.O.HALB). HALB outperforms HA because of the considerable delay increase in overloaded cliques (caused by HA); this, in turn, causes the increase of the average delay (shown in Fig. 56-(a)). When no load balancing is considered (HA), the exponential increase of

the delay in overloaded cliques results in a higher average delay over all cliques compared to the average delay of load balanced cliques.

Fig. 56-(b) shows that HALB outperforms considerably HA and W.O.HALB; for example, at load 16, HA (resp. W.O.HALB) generates 60% (resp. 41%) more losses than HALB. One can observe, in Fig. 56-(b) that W.O.HALB produces better results than HA. This is due to higher contentions and collisions in a larger clique extended by one or more logical links (because of MRs switching cliques due to the operation of HALB) in order to confine its handoff in the same clique.

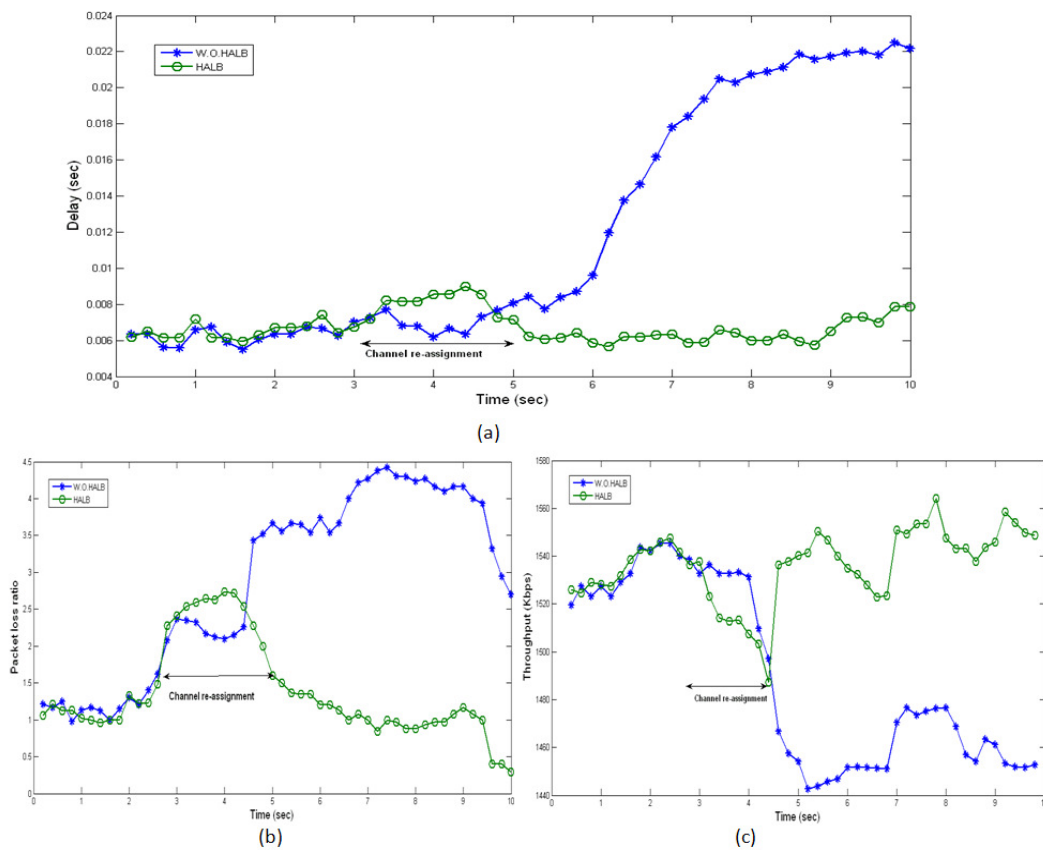


Fig. 57. Delay (a); packet loss ratio (b); throughput (c)

Since the throughput is tightly related to packet losses, we observe in Fig. 56-(c) a slight decrease in the throughput provided by HA over W.O.HALB. Fig. 56-(c) shows that HALB provides, at load 16, 10% and 6.2% more throughput than HA respond W.O.HALB respectively.

At saturation, the throughput reaches its maximal value constrained by the capacity of the MR-WMN; when operating at full capacity, new flows are rejected if admission control is implemented; otherwise, packet loss rates increase dramatically and delay reaches its maximum.

Fig. 57-(a) shows the delay variation in both HALB and W.O.HALB. This figure aims to show the effectiveness of intra-clique handoffs in decreasing the delay, especially when we take into account traffic load balancing. It shows also that the delay with W.O.HALB increases significantly after the channel re-assignment process; it exceeds 0.023 sec at time 6.8 s. Generally speaking, the delay with W.O.HALB after the channel re-assignment process is about three times the delay with HALB. HALB performs, whenever adequate, in-clique handoff and produces a lower delay; we conclude that intra-clique handoff has better delay performance than inter-clique handoff.

Fig. 57-(b) illustrates the variation of packet loss ratio; the channel re-assignment starts at time 3.2 sec and finishes at 4.5 sec. Note that before this period, W.O.HALB and HALB provide similar loss ratios; however, during the channel re-assignment process, when using HALB, the loss rate reaches its maximum value but never exceeds the loss ratio incurred by W.O.HALB. After the channel re-assignment process, HALB outperforms handily W.O.HALB.

Fig. 57-(c) shows that HALB has better throughput than W.O.HALB; its outperformance is even more marked after the channel re-assignment process. This can be explained by the fact that good load balancing capability leads to efficient usage of network resources and results in higher throughput. More specifically, when using W.O.HALB, the throughput decreases compared with HALB starting from 4.8 s; this is expected because of the increase in packet losses using W.O.HALB (see Fig. 57-(b)).

During the channel re-assignment process, HALB faces slight degradation in terms of delay, loss and throughput. In fact, the latency of transmitting signaling messages that perform the re-routing and the context transfer of the session from one path to another cause a path outage, and therefore a QoS degradation, during this short

transient period of channel re-assignment. This is a price worth paying to provide high throughput and low losses after this period while satisfying QoS requirements of new traffic flows.

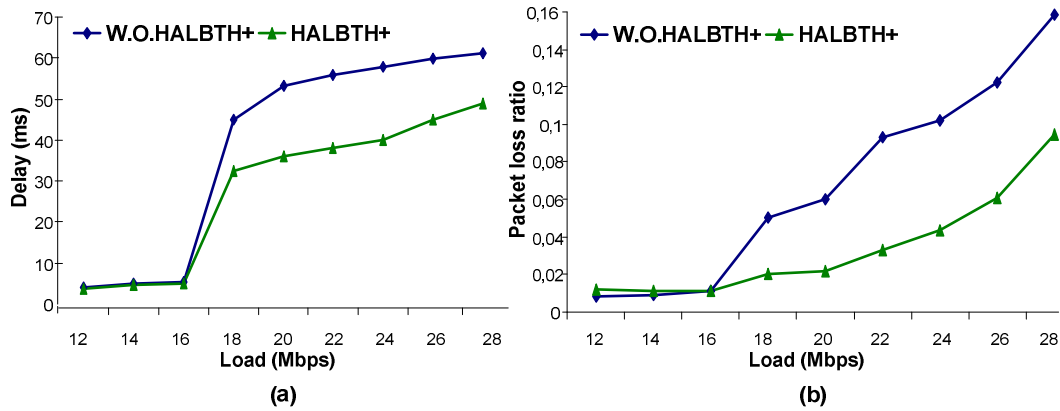


Fig. 58. Delay (a) and packet loss rate (b) with and without HALBTH+

Fig. 58 shows the performance, in terms of delay and packet loss, of the MR-WMN under a channel assignment using the proposed heuristic HALBTH+ (Algorithm 10). In Fig. 58-(a), we observe that HALBTH+ is able to reduce the delay by 26% compared to W.O.HALBTH+. Fig. 58-(b) shows that HALBTH+ reduces efficiently loss rate compared to W.O.HALBTH+. In fact, the mean loss rate (over all loads) of HALBTH+ decreases by 47% compared to W.O.HALBTH+. This improved performance provided by HALBTH+ (HALB combined with HALBTH) compared to the distributed CA dynamic scheme without good initial solution (HALB), shows the importance of a centralized CA heuristic in providing a good efficient assignment that serves as an initial solution to be dynamically improved by the distributed CA. However, unlike HALB, the centralized heuristic HALBTH cannot be performed dynamically in an online fashion, due to its time and message complexity.

5.6.3 Case III: Dynamic CA based heuristics MHALB+/MTABU+

Fig. 59, Fig. 60 and Fig. 61 show the performance of the proposed dynamic CA based heuristics (MHALB+/MTABU+) in terms of delay, throughput and packet loss.

Table 13 shows the CPU time of MHALB+/MTABU+ compared to the optimal solution computed by CPLEX with network of 19 nodes.

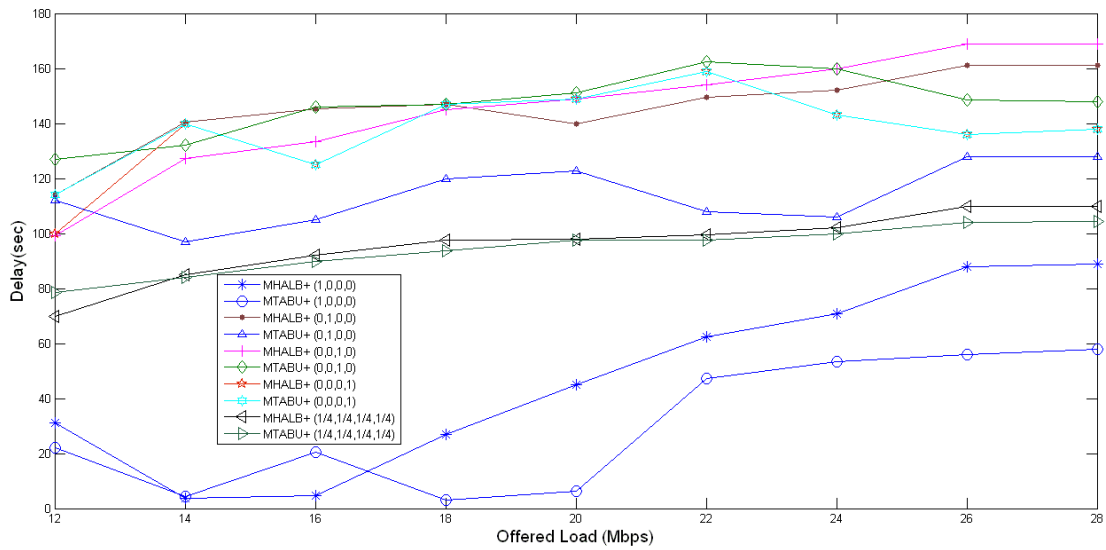


Fig. 59. Delay with MHALB+ and MTABU+

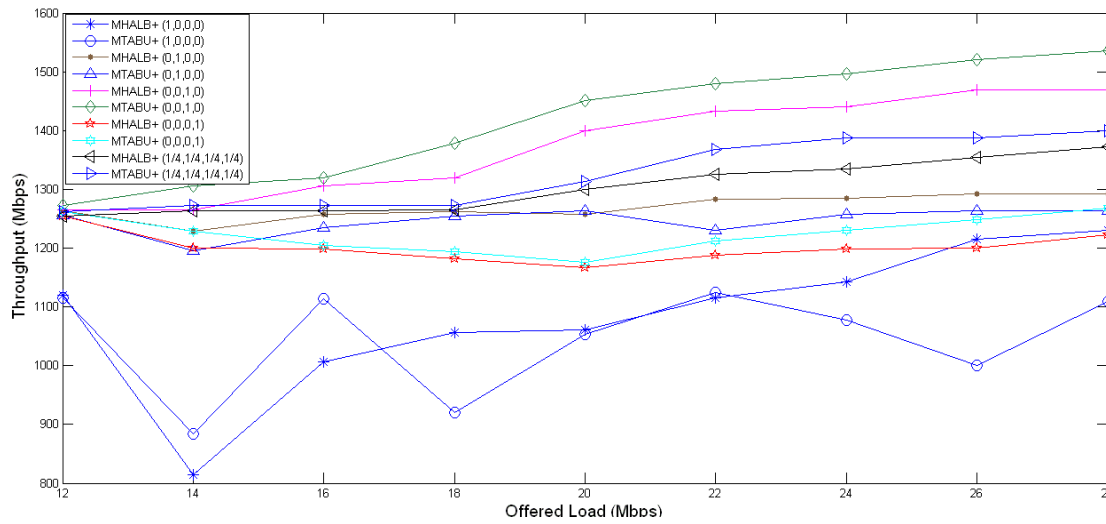


Fig. 60. Throughput with MHALB+ and MTABU+

Fig. 59 shows that the schemes, MHALB+/MTABU+, taking into account only load balancing (1,0,0,0) present the lowest delay (MHALB+ : 46ms, MTABU+: 30ms). This can be explained by the fact that intra-clique handoff has better delay performance than inter-clique handoff. With a multi-objective giving exclusive preference to load balancing, i.e. using (1,0,0,0), our approach outperforms the other weighted approaches

in terms of delay but at the expense of throughput degradation and higher losses (see Fig. 60 and Fig. 61).

However, when we consider a tradeoff between the four objective functions $(1/4, 1/4, 1/4, 1/4)$, the two meta-heuristics have an average delay for all loads equals to 94 ms which is good enough for multimedia traffic, such as voice which requires a maximal delay of 100 ms that does not exceed 150 ms required by ITU-T G.114 [Si99].

Fig. 60 shows that, considering only the throughput maximization $(0,0,1,0)$ for the two meta-heuristics, we obtain the highest throughput (1374 Kbps for MHALB+ and 1417 for MTABU+ in average). However, combining all objectives $(1/4, 1/4, 1/4, 1/4)$ has also a good throughput (5% for MHALB+ and 7% for MTABU+ less compared to $(0,0,1,0)$) which can be explained by the tight relationship between all of these performance metrics. In fact, increasing the throughput either directly or indirectly through load balancing will automatically decrease packet losses as shown in Fig. 61. The price to pay in order to provide high throughput, lower losses, and MCs fairness is an increase in terms of delay.

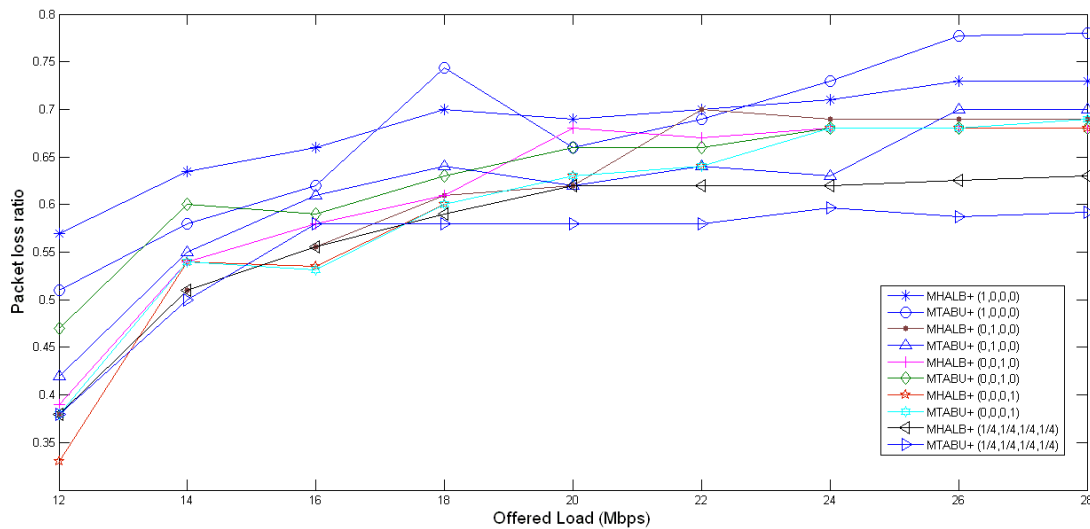


Fig. 61. Loss rate with MHALB+ and MTABU+

In summary, it is observed, through simulations, that several tradeoffs exist between load balancing, handoff overhead, throughput and fairness to determine a good solution that satisfies adequately all the criteria. In our simulations, we found that the

combination $(1/4, 1/4, 1/4, 1/4)$ provides the best results of our test runs; other combinations may provide better results since it is unfeasible to run all possible combinations. A sensitivity analysis is performed on each objective function in order to determine/study the impact and the usefulness of each objective function/criterion alone without taking into account the other criteria (e.g., $(1,0,0,0)$).

Table 12 and Table 13 show that for a large network size CPLEX cannot return a solution in a reasonable time. Therefore, in this case meta-heuristics based on VNS or Tabu search provide a good alternative to get near-optimal solutions. As we observe from Fig. 59, Fig. 60, Fig. 61 and Table 13, although Tabu search has a higher CPU time than VNS, it provides a CA solution that gives a slightly better performance compared to VNS.

Table 13. CPU time (sec) with MHALB+, MATBU+ and CPLEX: 19-nodes network

Network size: 19 nodes			
Load (Mbps)	MHALB⁺ (1/4, 1/4, 1/4, 1/4)	MTABU⁺ (1/4, 1/4, 1/4, 1/4)	CPLEX
12	639	1476	∞
14	766	1663	∞
16	829	1484	∞
18	836	1501	∞
20	902	1422	∞
22	877	1387	∞
24	957	1500	∞
26	990	1499	∞
28	990	1600	∞

5.7 Conclusions and Future Work

In this paper, we propose a new unified/generalized model for the CA problem in MR-WMNs with handoff support. The uniqueness of our scheme lies in the fact that rather than treating just non uniform traffic load during the CA process, as has been done so far in the literature, we propose a novel CA scheme that optimizes several objectives. We developed a novel multi-objective optimization model that integrates important network characteristics, such a handoff and load balancing, by exploiting the intimate relationships among them. In this paper, we used this model to determine a CA that minimizes inter-clique handoffs and traffic load variation, besides maximizing the throughput and ensuring MCs fairness while satisfying all technical constraints.

Moreover, we propose centralized static CA heuristics (variable neighborhood search: MHALB/Tabu search: MTABU) to search for a near-optimal initial CA solution optimizing the four performance metrics regardless of traffic variation. Then, we propose online distributed dynamic CA schemes (MHALB+/MTABU+) that continuously improve the near optimal solution produced by the static heuristics to the dynamic variation of traffic load and handoff rates. Particularly, our proposed online CA schemes try to continually minimize the handoffs involving path re-routing beyond the first hop MR and to ensure a better balanced load in the network. Using simulations, we show that our proposed approaches, integrating the centralized CA heuristics with the distributed online CA schemes, achieve good performance in terms of delay, loss rate and overall throughput.

Currently, we plan to use pure multi-objective methods to solve this problem and then to produce non-dominated solutions.

Chapitre 6: Conclusion et recherches futurs

Suite à l'explosion vertigineuse de la quantité d'informations échangée via Internet et à la multiplication croissante du nombre d'utilisateurs, il est devenu nécessaire voir même indispensable de recourir à des techniques et approches plus efficaces pour satisfaire une clientèle de plus en plus exigeante, ayant des besoins de plus en plus pointus, surtout en termes de la qualité des services offerts. Ces besoins se font sentir notamment au niveau des délais de transmission et de la perte de paquets qu'on souhaiterait aussi petits que possible. Les techniques de gestion adaptative de ressources ont pour objectif de satisfaire cette nécessité, d'où l'énorme intérêt qu'elles ne cessent de susciter. Il y a eu de ce fait plusieurs tentatives d'amélioration de ces techniques de gestion, mais le gain en délai et en débit reste toujours en deçà des exigences des utilisateurs. C'est pourquoi il faut introduire de nouvelles approches (théoriques et/ou pratiques) pour améliorer la performance de ces techniques.

C'est dans ce cadre de réflexion que nous avons proposé dans cette thèse de nouveaux modèles et algorithmes pour la QoS des WMNs. Notre démarche a consisté à agir au niveau des différentes composantes qui influencent le bon fonctionnement du réseau WMN.

Nous nous sommes principalement intéressés à trois problèmes majeurs : le contrôle d'admission du trafic, la différenciation du trafic et l'affectation adaptative des canaux en présence du trafic en relève. Nous avons vu que, malgré le nombre important de contributions dans le domaine de QoS de ce type de réseaux, la majorité des études est restée limitée à l'augmentation de capacité ou à la proposition de métriques de performance.

Le principe de contrôle d'admission consiste à n'accepter de nouvelles demandes que si l'on peut assurer la QoS de toutes les entités dont la demande a déjà été acceptée en même temps que celle de cette nouvelle demande. L'entité de demande en question

correspond à des unités de services très différentes : un appel téléphonique, un transfert de clip vidéo, un fichier, une page Web, etc. La fonction de contrôle d'admission à appliquer dépend de la nature des entités de demande, de leurs caractéristiques de trafic et de leurs exigences de QoS.

Offrir à des hôtes mobiles/trafic en relève un accès ubiquitaire à l'Internet devient de plus en plus important en raison de l'émergence de nouvelles applications, telles que l'accès mobile à l'information et les communications multimédia temps-réel. Un réseau local sans fil comme IEEE 802.11s est une solution intéressante pour le déploiement du réseau d'accès WMN dans des zones denses, en raison de son déploiement simple et non coûteux. La différenciation de services est impérative pour garantir une certaine QoS aux applications multimédia temps-réel. Cette tendance est née suite à l'hétérogénéité des équipements mobiles vendus sur le marché.

De même, la QoS observée dans un réseau WMN dépend de la gestion des canaux sans fil utilisés. En particulier, la gestion efficace des canaux dans un WMN à multiples-radios (MR-WMN) en présence du trafic en relève évitera des communications interrompues aux utilisateurs mobiles et leur fournira ainsi une QoS grandement supérieure.

Dans ce qui suit, nous résumons les contributions majeures de cette thèse et nous concluons avec quelques perspectives pour des recherches futures.

6.1 Résumé des contributions

Dans le chapitre 3, nous avons élaboré un nouveau mécanisme de contrôle d'admission appelé RCAC pour le protocole IEEE 802.11 qui utilise la méthode d'accès DCF. Ce mécanisme prend en compte les interférences et les ressources locales et dans un voisinage avant de prendre des décisions relatives à l'acceptation ou au rejet d'un trafic entrant. Les expériences que nous avons menées montrent que RCAC est capable de fournir la QoS exigée par les utilisateurs, particulièrement en termes de délai et de taux de perte des paquets, qui n'excèdent pas des seuils prédéfinis. De même, nous avons démontré par des simulations que la méthode d'accès CSMA/CA couplée

avec RCAC donne de meilleurs résultats que ceux de la méthode d'accès MDA en termes de débit, de surcoût, de délai et de taux de blocage.

Dans le chapitre 4, nous avons démontré que la méthode d'accès MDA renforcée par des mécanismes de contrôle d'admission et de gestion des tranches de temps ("time-slots"), selon la classe de service en présence du trafic en relève, est plus efficace que les méthodes d'accès DCF et EDCA en termes de longueur du délai, du taux de perte de paquets et surtout du taux de suppression du trafic en relève .

Finalement, dans le chapitre 5, nous avons développé en premier lieu, un modèle d'optimisation qui, en plus de maximiser le débit, améliore l'équité entre utilisateurs et minimise le surcoût dû à la relève des appels. Ce modèle a été résolu par le logiciel CPLEX pour un nombre limité de nœuds. En second lieu, nous avons élaboré des heuristiques/méta-heuristiques centralisées dans ce contexte pour permettre la résolution d'instances de tailles semblables à celle des réseaux déployés en pratique. En dernier lieu, nous avons développé un algorithme pour réaffecter en temps-réel et de façon prudente les canaux aux interfaces. Cet algorithme a pour objectif de minimiser le surcoût et le délai de re-routage, spécialement du trafic dynamique, généré par les appels en relève. Ce mécanisme est amélioré en tenant compte de l'équilibrage de la charge entre différents groupes de nœuds.

L'objectif du projet financé par Bell et le CRSNG était d'explorer de nouveaux concepts pour les environnements de réseaux maillés sans fil, de développer des architectures, des procédés et des outils utiles pour le développement futur de produits, garantissant la QoS des applications multimédia.

Nous avons participé à ce projet en élaborant une plateforme pour les réseaux WMNs dans laquelle on dispose d'algorithmes et d'heuristiques pour satisfaire les applications avec des contraintes temporelles et/ou les clients mobiles. Notre participation s'est concrétisée par la conception d'un modèle qui supporte la QoS en termes de court délai et de faible taux de pertes de paquets ou de blocage pour les applications temps-réel.

6.2 Directions pour les travaux futurs

Dans la continuité du travail présenté, nous pourrions approfondir notre étude afin d'améliorer les résultats obtenus. Dans le premier volet de cette thèse, nous avons proposé un mécanisme de contrôle d'admission qui considère une affectation statique de canaux. Cependant, l'impact d'une affectation dynamique des canaux lors de l'admission des flux est à étudier pour les MR-WMNs.

Dans le deuxième volet, nous nous sommes limités au contexte des WMNs à deux-sauts. Il serait pertinent d'intégrer des travaux qui améliorent la méthode d'accès MDA, en dehors du contexte deux-sauts, en termes d'interférences tels que présentés dans [CL08] et dans [GM09] pour essayer de généraliser notre modèle de différenciation de services et de minimiser l'interférence en dehors de l'environnement deux-sauts.

Dans le troisième volet, nous avons proposé une solution basée sur une agrégation de plusieurs fonctions objectif, or une solution purement multi-objectifs mériterait d'être étudiée. Dans ce contexte, nous constatons que la détermination de l'ensemble optimal de variables de décision pour un seul objectif, tel que minimiser le surcoût du trafic en relève, se traduit souvent par un ensemble non-optimal par respect à d'autres objectifs, comme celle de maximiser l'équité et le débit. L'ensemble optimal pour plusieurs fonctions objectif est connu sous le nom "front de Pareto" [FF95], [FI99] et [ZT99]. Ce front représente l'ensemble des solutions qui ne peuvent pas être améliorées dans toutes les dimensions. Les solutions sur le "front de Pareto" sont non-dominées et coexistent en raison de compromis entre les divers objectifs. Nous projetons développer des méthodes permettant de caractériser les solutions Pareto-optimales dans notre cas, ce qui nous donnera un ensemble de solutions optimales au lieu d'une seule solution pour des poids prédéfinis. Nous pensons que ceci peut s'avérer très utile pour un opérateur d'entreprise de télécommunications.

Dans le prolongement des travaux de cette thèse, nous avons déjà amorcé des recherches portant sur l'intégration des WMNs dans les réseaux métropolitains (MANs) [RB09], [RB10], [BR10a], [BR10b]. L'idée sous-jacente à ces recherches était

d'assurer une convergence entre les WMNs et les réseaux OBS (Optical Burst Switching) afin de bénéficier de leurs avantages respectifs. Nous souhaitons proposer une architecture MAN qui combine une couverture large et à faible coût des WMNs et une disponibilité de la bande passante des OBS. En effet, une architecture MAN qui relie l'utilisateur final à l'Internet tout en satisfaisant la demande de la bande passante et les spécifications des réseaux de la prochaine génération (NGNs) [CZ08] est nécessaire aujourd'hui. Nous réservons l'approfondissement de ce sujet de recherche pour des études futures.

Bibliographie

- [AB06] M. Alicherry, R. Bhatia, and L. Li, "Joint Channel Assignment and Routing for Throughput Optimization in Multi-radio WMNs", in *IEEE Journal Selected Areas in Communications*, Vol. 24, No.11, pp. 1960-1971, 2006.
- [AC02] G. Ahn, A. T. Campbell, A. Veres, and L. Sun, "SWAN: Service differentiation in stateless wireless ad hoc networks", in *Proc. of IEEE Infocom*, pp. 457-466, 2002.
- [AD06] Y. Amir, C. Danilov, M. Hilsdale, R. Musaloiu-Elefteri and N. Rivera, "Fast handoff for seamless wireless mesh networks", in *Proc. of ACM MobiSys*, pp.83-95, 2006.
- [AW05] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey", in *Journal of Computer Networks*, Vol. 47, No.4, pp.445-487, 2005.
- [AW07] I. Akyildiz and X. Wang, "Wireless Mesh Networks", *Advanced texts in Communications and Networking*, John Wiley & Sons, 2007.
- [AW08] I. Akyildiz and X. Wang, "Cross-Layer design in Wireless mesh networks", in *IEEE Journal Transactions on Vehicular Technology*, Vol. 57, No.2, pp.1061-1076, 2008.
- [BA06] N. Bisnik and A. A. Abouzeid, "Delay and throughput in random access wireless mesh networks", in *Proc. of IEEE ICC*, pp. 403-408, 2006.
- [BC04] P. Bahl, R. Chandra, and J. Dunagan, "SSCH: Slotted seeded channel hopping for capacity improvement in IEEE 802.11 ad-hoc wireless networks", in *Proc. of ACM MobiCom*, pp. 216-230, 2004.
- [BH07] A. Beljadid, A. Hafid, M. Gendreau, "Optimal Design of Broadband Wireless Mesh Networks", in *Proc. of IEEE Globecom*, pp.4840-4845 2007.
- [BH08] D. Benyamina, A. Hafid and M. Gendreau, "WMN Planning, A Multi-objective Optimization Approach", in *Proc. of IEEE Broadnets*, pp.602-609, 2008.
- [BR10a] A. Belbekkouche J. Rezgui and A. Hafid., "QoS provisioning for Wireless Mesh and Optical Burst Switching Convergence", in *Proc. of IEEE WCNC*, pp. 1-6, 2010.

- [BR10b] A. Belbekkouche J. Rezgui and A. Hafid, "Wireless Mesh and Optical Burst Switching Convergence for a Novel Metropolitan Area Network Architecture", *To appear in Journal of Computer Networks*, 2010.
- [BX10] N. Bayer, B. Xu, V. Rakocevic, and J. Habermann, "Application-aware scheduling for VoIP in Wireless Mesh Networks", *in Journal of Computer Networks*, Vol. 54, No.2, pp. 257-277, 2010.
- [CC04] C. Casetti, Chiasserini, "Improving fairness and throughput for voice traffic in 802.11e EDCA", *in Proc. of IEEE Pimrc*, pp. 525-530, 2004.
- [Ch06] L. Chen, "Wireless Mesh Networks (WMNs)", *Technical Report UIUCDCS-R-2006-2874*, Dept. of Computer Science, UIUC, 2006.
- [CJ03] T. Clausen and P. Jacquet, "Optimized Link State Routing Protocol (OLSR)", *IETF RFC 3626*, 2003.
- [CL08] C. Cicconetti, L. Lenzini, and E. Mingozzi, "Scheduling and dynamic relocation for IEEE 802.11s mesh deterministic access", *in Proc. of IEEE Secon*, pp. 19-27, 2008.
- [CN98] Crawley, E., R. Nair, B. Rajagopalan and H. Sandick, "A Framework for QoS-based Routing in the Internet", *RFC Editor*, 1998.
- [CP] "CPLEX 10.1.1", www.ILOG.com.
- [CS03] C. Savin, "Réseaux et Télécoms", livre, DUNDO, ISBN 2100079867, 2003.
- [CZ08] Y. Chi, L. Zhengbin, X. Anshi, "Dual-Fiber-Link OBS for Metropolitan Area Networks: Modelling, Analysis and Performance Evaluation", *in Proc. of IEEE Globecom*, pp.1-6, 2008.
- [DA05] A. K. Das, H. Alazemi, R. Vijaykumar, S. Roy, "Optimization models for fixed channel assignment in Wireless Mesh Networks with Multiple Radios", *in Proc. of IEEE Secon*, pp. 463-47, 4 2005.
- [DK88] E. Dahlhaus and M. Karpinski, "A fast parallel algorithm for computing all maximal cliques in a graph and the related problems", *in Proc. of SWAT*, pp.314-144, 1988.
- [DV06] F. De Greve, W. Vandenberghe, F. De Turck, I. Moerman, and P. De meester, "Towards Ethernet-based wireless mesh networks for fast moving users", *in Proc. of SEAA*, pp.387-397, 2006.
- [DY02] Y. Dong, T. Yang, and D. Makrakis, "SRL-enabled QoS model for mobile ad hoc networks", *in Proc. of ICCAS*, pp.414-418, 2002.

- [EG08] M. Ergin, M. Gruteser, L. Luo, D. Raychaudhuri, and H. Liu, "Available bandwidth estimation and admission control for QoS routing in wireless mesh networks", *Computer Communications*, Vol. 31, pp.1301-1317, 2008.
- [FF95] C. M. Fonseca and P. J. Fleming, "An overview of evolutionary algorithms in multiobjective optimization", in *Evolutionary Computation Journal*, Vol.3, No.1, pp.1–16, 1995.
- [F199] P. J. Fleming, "Designing control systems with multiple objectives", in *IEEE Colloquium Advances in Control Technology*, pp. 4/1–4/4, 1999.
- [GB07] M. Ghaderi, R. Boutaba, and G. W. Kenward, "Joint call and packet QoS in cellular packet networks", in *Journal of Computer Networks*, Vol. 51, No.4, pp.1060-1071, 2007.
- [Ge02] M. Gendreau, "Recent Advances in Tabu Search", in *Essays and Surveys in Metaheuristics*, C.C. Ribeiro and P. Hansen (eds.), Kluwer Academic Publishers, pp.369-377, 2002.
- [GM09] J. Gallardo, D. Makrakis, and H. Mouftah, "MARE: an Efficient Reservation-Based MAC Protocol for IEEE 802.11 s Mesh Networks", in *Proc. of IEEE MESH*, pp.97-102, 2009.
- [GZ03] D. Gu and J. Zhang, "A new measurement-based admission control method for IEEE 802.11 wireless local area networks ", in *Proc. of IEEE PIMRC*, pp.2009-2013, 2003.
- [HM03] P. Hansen and N. Mladenovic, "Variable neighbourhood search", in F.W. Glover and G.A. Kochenberger, editors, *Handbook of Metaheuristics*, Dordrecht Kluwer, 2003.
- [HM07] G. R. Hiertz, S. Max, Y. Zang, T. Junge, and D. Denteneer, "IEEE 802.11s MAC fundamentals", in *Proc. of IEEE MeshTech*, pp.1-8, 2007.
- [HM08] G. R. Hiertz, S. Max, T. Junge, D. Denteneert, and L. Berlemann, "IEEE 802.11s-Mesh deterministic access", in *Proc. of EW*, pp.1–8, 2008.
- [HP02] Z. Haas, M. Pearlman, and P. Samar, "The Zone Routing Protocol (ZRP) for Ad Hoc Networks", in *draft-ietf-manet-zone-zrp-02.txt*, 2002.
- [HY04] Z. Hongqiang, K. Younggoo, and F. Yuguang, "Performance analysis of IEEE 802.11 MAC protocols in wireless LANs", in *Wireless Communications. Mobile Computer*, Vol.4, No.8, pp.917–931, 2004.
- [JC84] R. Jain, D.M. Chiu and W. Hawe, "A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer

Systems", *DEC Research Report TR-301*, 1984.

- [JL05] L. Jiang and S. Liew, "Proportional fairness in wireless LANs and ad hoc networks", in *Proc. of IEEE WCNC*, pp.1551-1556, 2005.
- [JM04] B. Johnson David, A. Maltz David, and H. Yih-Chun, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR)", in *Internet Draft : <draft-ietf-manetsdr- 10.txt>*, 2004.
- [JP03] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance", in *Proc. of ACM Mobicom*, pp.66-80, 2003.
- [KG02] Kwon, T. J. and M. Gerla, "Efficient flooding with Passive Clustering (PC) in ad hoc networks", in *ACM of Sigomm Computer Communication Review*, Vol.32, No.1, pp.44-56, 2002.
- [KG07] A. Kashyap, S. Ganguly, S. R. Das, and S. Banerjee, "VoIP on wireless meshes: Models, algorithms and evaluation", in *Proc. of IEEE Infocom*, pp.2036-2044, 2007.
- [KL75] L. Kleinrock, "Queueing Systems", Vol. 1: Theory, New York, USA, Wiley Interscience. 448 pages, 1975.
- [KL03] Y. L. Kuo, C. Lu, E. Wu, and G. Chen, "An admission control strategy for differentiated service in IEEE 802.11", in *Proc. of IEEE Globecom*, pp.707-712, 2003.
- [KM04] Kumar, V. S. A., M. V. Marathe, et al, "End-to-end packet-scheduling in wireless ad-hoc networks", in *Proc. of ACM-SIAM symposium on Discrete algorithms*, pp.1021-1030, 2004.
- [KN05] M. Kodialam and T. Nandagopal, "Characterizing the Capacity Region in Multi-radio Multi-channel Wireless Mesh network", in *Proc. of ACM Mobicom*, pp.73-87, 2005.
- [Kr04] L. Krishnamurthy, "Making radios more like human ears: alternative MAC techniques and innovative platforms to enable large-scale meshes", in *Proc. of Microsoft Mesh Networking Summit*, 2004.
- [KV05] P. Kyasanur and N. Vaidya, "Routing and Interface Assignment in Multi-Channel Multi-Interface Wireless Networks", in *Proc. of IEEE WCNC*, pp.2051-2056, 2005.
- [KV06] P. Kyasanur and N. Vaidya, "Routing and Link-layer Protocols for Multi-Channel Multi-Interface Ad Hoc Wireless Networks", in *Mobile Computing and Communications Review*, Vol.10, No.1, pp.31-43, 2006.

- [LC00] L. Chung-Ching, "An On-Demand QoS Routing Protocol For Mobile Ad Hoc Networks", in *Proc. of IEEE Globecom*, pp.1783-1787, 2000.
- [LG06] L. Lou, M. Gruteser, H. Liu, D. Raychaudhuri, K. Huang, and S. Chen, "A QoS routing and admission control scheme for 802.11 ad hoc networks", in *Proc. of ACM DIWANS*, pp.19-28, 2006.
- [LW08] Y. Lin and V.W. S.Wong, "An admission control algorithm for multi-hop 802.11e based WLANs ", in *Computer Communications journal*, Vol.31, No.14, pp.3510-3520, 2008.
- [MD05] M. Marina and S. R. Das, "A Topology Control Approach for Utilizing Multiple Channels in Multi-Radio Wireless Mesh Networks", in *Proc. of Broadnets*, pp.381-390, 2005.
- [Ne08] T.R. Newman, "Multiple Objective Fitness Functions for Cognitive Radio Adaptation", Ph.D. Dissertation, University of Kansas, Lawrence, KS, 2008.
- [NS] NS-2 simulator. [Online]. Available: <http://www.isi.edu/nsnam/ns/>
- [Ne04] M. Nehdi, "Évaluation de protocole EDCA", Tunis, Tunisie, 2004. [Online]. Available: http://www.etudionet.com/v0/communaute/xuser/etudionet/docs/PFE_NEHDI_Mourad.pdf.
- [NG06] D. Niculescu, S. Ganguly, K. Kim, and R. Izmailov, "Performance of VoIP in a 802.11 wireless mesh network," in *Proc. of IEEE Infocom*, pp.1–11, 2006.
- [OL09] [Online]. Available: http://fr.wikipedia.org/wiki/Algorithme_de_Ford-Fulkerson, visited Jan. 15, 2009.
- [OJ03] D. O'Neill, D. Julian, and S. Boyd, "Adaptive management of network resources", in *Proc. of IEEE VTC*, pp.1929-1933, 2003.
- [PB94] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers", in *ACM Sigcomm Computer Communications Review archive*, Vol.24, No.4, pp.234–244, 1994.
- [PB03] C. E. Perkins, E. M. Belding-Royer, and S. Das, "Ad Hoc on demand distance vector (AODV) routing", *IETF RFC 3561*, 2003.
- [PB96] Peterson, L. L. and B. S. Davie, "Computer networks: a systems approach", *Morgan Kaufmann Publishers Inc*, 1996.

- [RB09]** J. Rezgui, A. Belbekkouche and A. Hafid, "Novel Analytical Delay Model and Burst Assembly Scheme for Wireless Mesh and Optical Burst Switching Convergence", in *Proc of ICST QShine*, pp.104-119, 2009.
- [RB10]** J. Rezgui, A. Belbekkouche and A. Hafid, "On Delay performance and Burst Assembly for Wireless Mesh and Optical Burst Switching Converged Metro Area Network", in *ACM/Springer Mobile Networks and Applications*, DOI: 10.1007/s11036-010-0249-3, 2010.
- [RC04]** A. Raniwala and T. Chiueh, "Evaluation of a Wireless Enterprise Backbone Network Architecture", in *Proc. of Hot-Interconnects*, pp.25-27, 2004.
- [RC05]** A. Raniwala and T. Chiueh, "Architecture and Algorithms for an IEEE 802.11-Based Multi-Channel Wireless Mesh Network", in *Proc. of IEEE Infocom*, pp.2223-2234, 2005.
- [RG04]** A. Raniwala, K. Gopalan and T. Chiueh., "Centralized Channel Assignment and Routing Algorithms for Multichannel Wireless Mesh Networks", in *ACM of Mobile Computing and Communications Review*, Vol.8, pp.50-65, 2004.
- [RH08]** J. Rezgui, A. Hafid, and M. Gendreau, "A distributed admission control scheme for Wireless Mesh Networks", in *Proc. of Broadnets*, pp.594-601 2008.
- [RH09]** J. Rezgui, A. Hafid, M. Gendreau and B.Rong, "Congestion-aware Clique-based Handoff in Wireless Mesh Networks", in *Proc. of IEEE MSN*, pp.388-395, 2009.
- [RH10b]** J. Rezgui, A. Hafid, R. Ben Ali and M. Gendreau, "Handoff-Aware Channel Assignment for Multi-Radio Wireless Mesh Networks", in *Proc. of IEEE ICC*, pp.1-6, 2010.
- [RH10c]** J. Rezgui, A. Hafid, R. Ben Ali and M. Gendreau, "Meta-heuristics for (Re-) Channel Assignment problem in Multi-Radio Wireless Mesh Networks", in *Proc. of INFORMS Telecom*, 2010.
- [RH10d]** J. Rezgui, A. Hafid, and M. Gendreau, "Distributed Admission Control in Wireless Mesh Networks: Models, Algorithms, and Evaluation", in *IEEE Journal Transactions on Vehicular Technology*, Vol.59, No.3, pp.1459-1473, 2010.
- [RH10e]** J. Rezgui, A. Hafid, M.Gendreau, "Admission Control and QoS Provisioning for IEEE 802.11s Mesh Deterministic Access ", *Submitted in Journal of Computer Networks*, 2010.

- [RH10f] J. Rezgui, A. Hafid, R. Ben Ali and M. Gendreau, "A Variable Neighborhood Search Method for Multi-Objective Channel Assignment Problem in Multi-Radio WMNs", *to appear in Proc. of IEEE LCN*, 2010.
- [RH10g] J. Rezgui, A. Hafid, R. Ben Ali and M. Gendreau, "Optimization Model for Handoff-Aware Channel Assignment Problem for Multi-Radio Wireless Mesh Networks", Submitted in *IEEE Journal Transactions on Vehicular Technology*, 2010.
- [RH10h] J. Rezgui, A. Hafid, R. Ben Ali and M. Gendreau, "Admission Control and QoS Provisioning in Multi-Service MDA for IEEE 802.11s-based Wireless Mesh Networks ", Accepted in *Proc. of ICST Qshine*, 2010.
- [RV06] K. N. Ramachandran, E. M. Belding, K. C. Almeroth and M. M. Buddhikot, "Interference Aware Channel Assignment in Multi-Radio Wireless Mesh Networks", in *Proc. of IEEE Infocom*, pp.1-12, 2006.
- [RW07] A. Rad and V. Wong, "Joint channel allocation, interface assignment and MAC design for multi-channel wireless mesh networks", in *Proc. of IEEE Infocom*, pp.1469-1477, 2007.
- [SC06] S. Speicher, and C. Cap, "Fast Layer 3 Handoffs in AODV-based IEEE 802.11 Wireless Mesh Networks", in *Proc. of ISWCS*, pp. 233-237, 2006.
- [SG07] H. Skalli, S. Ghosh, S. K. Das and L. Lenzini, "Channel Assignment Strategies for Multiradio Wireless Mesh Networks: Issues and Solutions", in *IEEE Communications Magazine*, Vol.45, No.11, pp.86-95, 2007.
- [SG08] A. Subramanian, H. Gupta, S. Das, and J. Cao, "Minimum interference channel assignment in multiradio wireless mesh networks", in *IEEE Transactions Journal on Mobile Computing*, Vol.7, No.12, pp.1459-1473, 2008.
- [SH05] N. Salem and J. Hubaux, "A fair scheduling for wireless mesh networks", in *Proc. of IEEE WiMesh*, 2005.
- [Si99] E. D. Siegel, "Designing Quality of Service Solutions for the Enterprise", *Book, John Wiley & Sons, Inc*, ISBN: 0471333131, 1999.
- [SS02] C. C. Shen, C. Srisathapornphat, and C. Jaikaeo, "The Guirrella Management Architecture for Ad hoc Networks", in *Proc. of MILCOM*, pp.467-472, 2002.
- [SS03] C. C. Shen, C. Srisathapornphat, and C. Jaikaeo, "An adaptive management architecture for ad hoc networks", in *IEEE Communications*

Magazine, Vol.41, No.2, pp.108–115, 2003.

- [So07] M. Soleri, IEEE 802.11s mesh deterministic access: Design and analysis, Pisa, Italy, 2007. [Online]. Available: <http://etd.adm.unipi.it/ETD-db/NDLTD-OAI/oai.pl>
- [SV04] J. So and N. Vaidya, "Multi-channel mac for ad hoc networks: Handling multi-channel hidden terminals using a single transceiver", in *Proc. of ACM MobiHoc*, pp.222-233, 2004.
- [SX03] C. Shim, L. Xie, B. Zhang and C.J. Sloane, "How delay and packet loss impact voice quality in Voip", *Qovia, Inc*, <http://www.qovia.com>, 2003.
- [SX08] A.M. Srivatsa and J. Xie, "A Performance Study of Mobile Handoff Delay in IEEE 802.11-Based Wireless Mesh Networks", in *Proc. of IEEE ICC*, pp.2485-2489, 2008.
- [Qo99] QoS Forum. QoS protocols and architectures. White paper of QoS Forum, <http://www.qosforum.com>, 1999.
- [WK06] H. Wei, K. Kim, A. Kashyap, and S. Ganguly, "On admission of VoIP calls over wireless mesh network", in *Proc. of IEEE ICC*, pp.1990-1995, 2006.
- [WL04] Y.Weii, C.Lin, F.Ren, R.Raad, E.Dutkiewicz, "Dynamic handoff scheme in differentiated QoS wireless multimedia networks", in *Computer and Communications Journal*, Vol.27, No.1, pp.1001–1011, 2004.
- [WN99] C. Wenli, J. Nitin, and S. Singh, "ANMP: ad hoc network management protocol", in *IEEE journal on Selected Areas in Communications*, Vol.17, No.8, pp.1506-1531, 1999.
- [WT07] H. Wu, K. Tan, Y. Zhang and Q. Zhang, "Proactive Scan: Fast Handoff with Smart Triggers for 802.11 Wireless LAN", in *Proc. of IEEE Infocom*, pp.749-757, 2007.
- [WW06] W. Wang, Y. Wang, X. Li, W. Song, and O. Frieder, "Efficient interference-aware TDMA link scheduling for static wireless networks," in *Proc. of ACM MobiCom*, pp.262-273, 2006.
- [XL04] Y. Xiao and H. Li, "Local data control and admission control for QoS support in wireless ad hoc networks ", in *IEEE journal Transactions on Vehicular Technology*, Vol.52, No.5, pp.1558-1572, 2004.
- [XS00] H. Xiao, K. G. Seah, A. Lo, and K. C. Chua, "A flexible quality of service model for mobile ad-hoc networks", in *Proc. of IEEE VTC*, pp.445-449, 2000.

- [XX06] C. XiaoLin., L. Xiaoyang, et al, "A control-theoretic approach to improving fairness in DCF based WLANs. Performance", in *Proc. of IEEE IPCCC*, pp.86-93, 2006.
- [YJ08] Y. Yaling and W. Jun, "Design guidelines for routing metrics in multihop wireless networks", in *Proc. of IEEE Infocom*, pp.1615-1623, 2008.
- [YK05] Y. Yaling and R. Kravets, "Contention-aware admission control for ad hoc networks ", in *IEEE Journal Transactions On Mobile Computers*, vol.4, No.4, pp.363-377, 2005.
- [ZD01] Y. Zhang and D. Liu, "An adaptive algorithm for call admission control in wireless networks", in *Proc. of IEEE Globecom*, pp.3628-3632, 2001.
- [ZT99] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach", in *IEEE Journal Transactions on Evolutionary Computation*, vol.3, No.4, pp.257-271, 1999.