

**UNIVERSITÉ DE PIERRE ET MARIE CURIE – PARIS 6  
U.F.R. DE SCIENCES**

***THÈSE***

pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITÉ DE PIERRE ET MARIE CURIE**

***Discipline* : SYSTÈMES INFORMATIQUES**

présentée et soutenue publiquement

par

**M. Sidi-Mohammed SENOUCI**

le 09 Octobre 2003

Titre :

**APPLICATION DE TECHNIQUES D'APPRENTISSAGE DANS LES  
RESEAUX MOBILES**

**Directeur de thèse :  
M. Guy PUJOLLE**

**JURY**

Président	Mlle. Dominique GAITI	Professeur à l'UTT
Rapporteur	M. André-Luc BEYLOT	Professeur à l'ENSEEIH
Rapporteur	M. Younes BENNANI	Professeur à l'université de Paris 13
Examineur	M. Sami TABBANE	Enseignant-chercheur à l'ESPTT
Examineur	M. Otto SPANIOL	Professeur à l'université d'Aachen
Examineur	M. Laurent REYNAUD	Chercheur chez France Télécom R&D
Examineur	M. Samir TOHME	Professeur à l'ENST
Examineur	M. Guy PUJOLLE	Professeur à l'université de Paris 6



# Remerciements

Je tiens à remercier M. Guy Pujolle, Professeur à l'université de Pierre et Marie Curie, de m'avoir accueilli au sein de l'équipe Réseau et Performance (RP) du Laboratoire Informatique de Paris 6 (LIP6) et de m'avoir permis de travailler dans d'aussi bonnes conditions. Ce travail n'aurait pu voir le jour sans sa présence attentive et ses conseils toujours pertinents. Mes plus vifs remerciements lui sont adressés.

J'adresse toute ma gratitude à M. André-Luc Beylot, Professeur à l'ENSEEIH, pour ses judicieux conseils qui ont permis d'améliorer la qualité de ce rapport, et pour avoir contribué à la réalisation de certains des travaux présentés dans cette thèse. Je le remercie également M. Younes Bennani, Professeur à l'Université de Paris 13, pour avoir immédiatement accepté le rôle de rapporteur et de l'avoir assumé avec beaucoup de patience et de pertinence.

Je souhaite aussi exprimer mes vives remerciements à Mlle. Dominique Gaïti, Professeur à l'UTT, M. Otto Spaniol, Professeur à l'université technologique d'Aachen en Allemagne, M. Sami Tabbane, enseignant-chercheur à l'ESPTT en Tunisie, M. Samir Tohmé, Professeur à l'ENST, et M. Laurent Reynaud, chercheur chez France Télécom R&D, d'avoir accepté sans hésitation de prendre part à ce jury de thèse.

Je remercie tous les membres du groupe RP pour le climat sympathique dans lequel ils m'ont permis de travailler. Les nombreuses discussions que j'ai pu avoir avec chacun m'ont beaucoup appris.

Une mention toute spéciale est réservée dans mon cœur pour tous ceux qui ont relu soigneusement toutes les étapes de ce mémoire ou qui m'ont grandement aidé à avancer tout au long de cette thèse : Selma, Khaled, Laetitia et en particulier Yacine.

Je tiens à redire toute mon amitié à la joyeuse bande des "jeunes" : Nadjib, Prométhée, Lila, Anelise, Mauro, Marcelo, Louis, Hakima, Nazim, Youcef, Kamel, Samir, Maha, Davor, Julien, Benoit, Konstantin, Dominique, Leila, Arthur, Erwan, avec qui j'ai partagé ces quelques années de thèse.

Enfin, j'adresse toute mon affection à ma famille, et en particulier mes sœurs et mes frères, à mon père dont j'aimerais dire à quel point sa présence m'est précieuse et à ma "beautiful sœur". Leur intelligence, leur confiance, leur tendresse, leur amour me portent et me guident tous les jours.

*À ma maman.*



## Résumé

Les demandes en transmissions sans fil, fournissant des communications fiables de voix et de données "n'importe où et n'importe quand", ont augmenté massivement ces dernières années. L'explosion de ce marché, sa croissance soutenue et l'apparition de nouveaux services amènent les réseaux mobiles actuels à leur limite. Contrairement aux réseaux filaires, plusieurs problèmes tels que le contrôle d'admission d'appels (CAC – Call Admission Control), l'allocation des ressources, la gestion de la localisation et le routage sont plus difficiles à résoudre dans les réseaux mobiles, et doivent leur complexité aux imperfections du support sans fil. Les solutions existantes ignorent souvent l'expérience et la connaissance qui pourraient être acquises pendant l'exécution du système. Nous démontrons, dans cette thèse, l'utilité des techniques d'apprentissage pour résoudre un certain nombre de problèmes rencontrés dans ces réseaux mobiles.

Nos premières contributions consistent à améliorer, grâce aux techniques d'apprentissage par renforcement, le contrôle d'admission (CAC) et l'allocation dynamique des ressources dans les réseaux cellulaires multiservices. Les solutions proposées sont robustes, améliorent considérablement la qualité de service (QoS), et réduisent les probabilités de coupure et ceci malgré les variations dans les conditions de trafic.

Constatant le grand succès de ces réseaux cellulaires, de nouvelles solutions de communication sans fil de plus en plus performantes sont en train de paraître. Parmi ces solutions, nous nous sommes particulièrement intéressé aux réseaux ad hoc ; ce sont des réseaux sans fil ne nécessitant aucune infrastructure fixe pour communiquer. Nos contributions se concentrent sur le problème du routage ad hoc, et plus spécialement à l'aspect minimisation de la consommation de l'énergie, et par conséquent l'amélioration de la durée de vie du réseau (survivabilité). L'une des quatre solutions de routage ad hoc proposées utilise l'algorithme d'apprentissage par renforcement (RL).

**Mots-clés :** Réseaux cellulaires, Réseaux ad hoc, Contrôle d'admission CAC, Allocation dynamique de ressources DCA, routage ad hoc, Apprentissage par renforcement, Q-learning.

## **Abstract**

The increasing demand and rapid growth of mobile communications that provide reliable voice and data communications “anytime and anywhere” has massively grown. However, compared to the wired networks, several problems such as call admission control (CAC), channel allocation, location management and routing are more difficult to solve and owe their complexity to the shortcomings of the wireless medium. The existing solutions completely ignore the experience or knowledge that can be acquired during system operation. We argue that the use of learning techniques can lead to very good solutions in reasonable times. In this thesis, we show the benefits of learning techniques to solve some problems encountered in mobile networks.

Our first contribution, consists on improving call admission control and dynamic channel allocation in multi-service cellular networks, using reinforcement learning techniques. The suggested solutions are robust, improve the Quality of Service (QoS), and reduce call-blocking probabilities for handoff calls in spite of variations in the traffic conditions.

Given the great success of cellular networks, new powerful wireless communication solutions have been developed. Among these solutions, we were particularly interested in ad hoc networks, which are wireless networks that do not need any fixed infrastructure. Our work, has focused on routing problem, and more specifically on mechanisms minimizing energy consumption, and extending the network lifetime (survivability). One of the four suggested solutions of ad hoc routing uses the reinforcement learning algorithm (RL).

**Keywords :** Cellular networks, ad hoc networks, Call Admission Control CAC, Dynamic channel allocation DCA, ad hoc routing, Reinforcement learning, Q-learning.

# Table des matières

<b>Introduction.....</b>	<b>7</b>
1. Motivations.....	7
2. Contributions.....	9
<b>Réseaux cellulaires : Vue d'ensemble.....</b>	<b>13</b>
1. Introduction.....	13
2. Notions de base.....	14
2.1 La technique FDMA.....	14
2.2 La technique TDMA.....	15
2.3 La technique CDMA.....	15
3. Les réseaux cellulaires.....	16
3.1 Fondements.....	17
3.2 La cellule.....	17
3.3 La réutilisation des fréquences.....	17
3.4 Le transfert inter-cellulaire (ou Handoff).....	17
4. Schémas d'allocation de canaux.....	17
4.1 Fixed Channel Allocation.....	17
4.2 Dynamic Channel Allocation.....	17
4.3 Comparaison.....	17
4.4 Hybrid Channel Allocation.....	17
5. Diminution des blocages.....	17
5.1 Réserve des canaux.....	17
5.2 Mise en attente de nouveaux appels.....	17
5.3 Mise en attente des handoffs.....	17
5.4 Contrôle d'admission.....	17
5.5 Recouvrement des cellules.....	17
6. Conclusion.....	17
<b>L'Apprentissage.....</b>	<b>17</b>
1. Introduction.....	17
2. Apprentissage numérique.....	17
2.1 Apprentissage supervisé.....	17
2.2 Apprentissage non-supervisé.....	17
2.3 Apprentissage par renforcement.....	17
3. Applications.....	17
3.1 Applications classiques.....	17
3.2 Routage.....	17

3.3	Allocation de ressources.....	17
3.4	Contrôle d'admission CAC .....	17
4.	Conclusion.....	17
<b>Contrôle adaptatif d'admission des appels.....</b>		<b>17</b>
1.	Motivations.....	17
2.	Cas d'une seule cellule FCA .....	17
2.1	Formulation du problème sous forme de SMDP .....	17
2.2	Implantation des algorithmes .....	17
2.3	Étude Analytique du mécanisme de réservation de canaux .....	17
2.4	Résultats expérimentaux.....	17
3.	Système FCA .....	17
3.1	Implantation de l'algorithme QCAC-FCA .....	17
3.2	Résultats expérimentaux.....	17
4.	Conclusion.....	17
<b>Allocation dynamique et adaptative des ressources .....</b>		<b>17</b>
1.	Motivations.....	17
2.	Formulation du problème sous forme de SMDP .....	17
3.	Implantation de l'algorithme Q-DCA.....	17
3.1	Représentation des Q-valeurs .....	17
3.2	Implantation .....	17
3.3	Exploration.....	17
4.	Résultats expérimentaux .....	17
4.1	Répartition uniforme du trafic .....	17
4.2	Répartition non uniforme du trafic .....	17
4.3	Charge de trafic variable dans le temps .....	17
4.4	Panne d'équipement dans un système DCA.....	17
5.	Conclusion.....	17
<b>Réseaux ad hoc : Vue d'ensemble .....</b>		<b>17</b>
1.	Introduction .....	17
2.	Présentation des réseaux Ad hoc .....	17
2.1	Applications cibles .....	17
2.2	Avantages.....	17
2.3	Inconvénients .....	17
3.	Principaux algorithmes de routage ad hoc .....	17
3.1	Protocoles proactifs .....	17
3.2	Protocoles réactifs .....	17
3.3	Protocoles hybrides .....	17
4.	Conclusion.....	17
<b>Consommation de l'énergie dans les réseaux ad hoc.....</b>		<b>17</b>
1.	Motivations.....	17
2.	Protocoles minimisant la consommation des batteries .....	17
3.	Impact du routage sur la consommation de l'énergie .....	17
3.1	Caractéristiques énergétiques des protocoles de routage .....	17
3.2	Étude comparative des protocoles de routage.....	17
4.	Conclusion.....	17
<b>Protocoles de routage minimisant la consommation de l'énergie.....</b>		<b>17</b>

1. Motivations.....	17
2. Le protocole LEAR-AODV .....	17
2.1 Découverte de route .....	17
2.2 Maintien de la route.....	17
3. Le protocole PAR-AODV .....	17
3.1 Découverte de route .....	17
3.2 Maintien de la route.....	17
4. Le protocole LPR-AODV .....	17
4.1 Découverte de route .....	17
4.2 Maintien de la route.....	17
5. Le protocole Q-AOMDV .....	17
5.1 Découverte de route .....	17
5.2 Maintien de la route.....	17
6. Résultats expérimentaux .....	17
6.1 Nœuds fixes .....	17
6.2 Nœuds mobiles.....	17
7. Conclusion.....	17
<b>Conclusion.....</b>	<b>17</b>
1. Perspectives.....	17
<b>Liste des publications.....</b>	<b>17</b>
<b>Bibliographie .....</b>	<b>17</b>



# Chapitre 1

## Introduction

### 1. Motivations

Avec l'évolution de la société actuelle, les personnes se déplacent de plus en plus, tout en ayant besoin de communiquer pendant leurs déplacements. Ce phénomène a provoqué une demande accrue et orienté les études vers le développement de systèmes très sophistiqués afin de répondre aux nouveaux besoins des utilisateurs. Ces besoins ont effectivement changé ; si la voix était à l'origine le seul besoin, les demandes en transmissions sans fils fournissant des communications fiables de son à haute définition, d'images, voire de vidéos de haute qualité est devenu de plus en plus prisé par un nombre croissant d'utilisateurs. Ces derniers, souhaitent que la mobilité soit complètement transparente afin de bénéficier de performances comparables à celles des réseaux filaires, malgré la gourmandise en bande passante de ces nouveaux services.

Les systèmes cellulaires sont sans aucun doute, ceux qui ont connu la plus grande évolution ces dernières années. Un des grands challenges pour ce type de réseaux est la mobilité des usagers durant une même communication (changement de cellules). En effet, la disponibilité des ressources radios durant toute la durée de la communication n'est pas nécessairement garantie, et ces utilisateurs peuvent subir une dégradation ou même une rupture de la communication lors d'un passage d'une cellule à une autre. Cet événement, appelé transfert inter-cellulaire ou "handoff", doit être transparent à l'utilisateur. L'un des soucis majeurs lors de la conception des réseaux cellulaires, consiste à éviter aux usagers le désagrément causé par des coupures de communication. En effet, du point de vue du client, cette coupure de communication est beaucoup plus désagréable qu'un échec de connexion.

Étant donné que la ressource radio est une ressource rare, il est impératif de l'utiliser au maximum et particulièrement dans le cas d'un réseau cellulaire multiservices. Pour un opérateur il est parfois préférable de bloquer un appel d'une classe de service moins prioritaire (données par exemple) et d'en accepter un autre appel d'une classe plus prioritaire (voix par exemple). Ainsi, une bonne politique de contrôle d'admission d'appel (CAC) est certainement nécessaire afin de permettre de maximiser l'utilité de l'ensemble de ces ressources radio. Pour aboutir à cet objectif, il est également indispensable de trouver une bonne méthode d'allocation de la totalité de la bande passante disponible à l'ensemble des cellules. Ces nouveaux mécanismes (CAC, allocation dynamiques des ressources) doivent également faire face aux changements fréquents des conditions de trafic dans les réseaux cellulaires.

L'engouement suscité par ces réseaux cellulaires laisse à penser que l'utilisateur va devenir de plus en plus mobile. Ces systèmes connaissent une très forte expansion à l'heure actuelle, cependant ils requièrent une importante infrastructure logistique et matérielle fixe. L'ambition vouée des réseaux ad hoc est d'étendre les notions de mobilité pour permettre l'accès à l'information et à la communication "n'importe où et n'importe quand" (Anywhere and Anytime). Les réseaux ad hoc éliminent le besoin d'infrastructure fixe pour communiquer. L'activité du groupe MANET [1] de l'IETF<sup>1</sup> montre que le développement de ces réseaux sans fil et sans infrastructure est en plein essor. Les industriels imaginent déjà toutes sortes d'applications. Militaires bien sûr pour la création de réseaux tactiques mobiles, mais aussi civiles pour les interventions d'urgence, les communications avec les automobiles, la reconfiguration de réseaux sans câblage dans les entreprises ou bien la création de réseaux temporaires autour d'événement. Sans conteste, les atouts majeurs de cette nouvelle génération de réseaux mobiles sont la flexibilité et leur faible coût.

Un des grands challenges pour ce type de réseaux réside dans l'autonomie restreinte des stations mobiles le constituant. Effectivement, cette autonomie est fournie par de simple batteries, et donc représente une ressource finie et rare. Chaque paquet envoyé ou reçu, de même que chaque utilisation du terminal mobile profite de cette ressource. Et comme l'amélioration du confort, et de l'ensemble des fonctionnalités offertes aux utilisateurs est de plus en plus appréciable, réduire la consommation de l'énergie au minimum est un défi important dans les réseaux mobiles. Cet objectif devient davantage considérable pour les

---

1. IETF : Internet Engineering Task Force.

réseaux ad hoc, où les stations ont de surcroît la fonction de routage. En effet, relayer des paquets au nom d'autres nœuds, consomme l'énergie propre au nœud.

## 2. Contributions

Constatant qu'il est intéressant d'exploiter l'expérience et la connaissance qui pourraient être acquises en cours de fonctionnement d'un réseau cellulaire, nous proposons, dans le cadre de cette thèse, de nouveaux mécanismes de contrôle d'admission d'appels dans ces réseaux. Ces mécanismes considèrent les deux schémas d'allocation de canaux : fixe (FCA) et dynamique (DCA), et sont obtenus en utilisant l'algorithme d'apprentissage par renforcement Q-learning.

Nous proposons, également, pour les systèmes d'allocation dynamique, où toutes les ressources radios sont groupées dans un *pool* (ou groupe) commun, un nouveau mécanisme basé sur le Q-learning et permettant de choisir le meilleur canal parmi tous les canaux disponibles dans le pool commun. Ce mécanisme couplé avec le mécanisme de contrôle d'admission a pour objectif de maximiser le taux d'utilisation de chaque canal.

Ces propositions, que ce soit pour des systèmes d'allocation fixe ou dynamique, permettent de favoriser les appels plus prioritaires, tout en minimisant les probabilités de coupure des appels en cours de communication.

En ce qui concerne les réseaux ad hoc, nous avons constaté que la totalité des protocoles de routage actuels, proposés au sein du groupe MANET de l'IETF, utilisent la même métrique (nombre de sauts ou temps de transmission minimum). Constatant, également, que le choix du protocole de routage influe réellement sur la consommation de l'énergie dans les réseaux ad hoc, nous proposons de nouveaux protocoles de routage utilisant des métriques basées sur l'énergie. Ces métriques vont permettre de rallonger la durée de vie des batteries, et par conséquent la durée de vie du réseau ad hoc (survivabilité du réseau).

Ces nouveaux protocoles permettent ainsi, d'assurer que la connectivité du réseau soit maintenue aussi longtemps que possible, tout en évitant de le partitionner en sous-réseaux disjoints. Les protocoles que nous avons développé (LEAR-AODV, PAR-AODV, LPR-AODV et Q-AOMDV) sont des protocoles réactifs et sont basés sur l'un des plus importants protocoles de routage actuels : AODV<sup>2</sup>. L'idée fondamentale du protocole LEAR-AODV est de mettre en

---

2. Ad hoc On demand Distance Vector.

cause la "bonne volonté" de chaque nœud mobile à participer au processus de sélection de route et de relayer des paquets de données au nom d'autres nœuds. Ainsi, un nœud dont la batterie est trop utilisée, peut refuser de relayer ces paquets. Le routage PAR-AODV recherche la meilleure route minimisant une fonction coût prenant en compte l'énergie résiduelle de chaque nœud dans la route. Le protocole LPR-AODV favorise la route dont la durée de vie est maximum, et par conséquent la route qui ne contient pas de nœuds dont la durée de vie est faible. Le protocole Q-AOMDV, quant à lui, recherche plusieurs routes allant de la source jusqu'à la destination, et essaie ensuite d'équilibrer le trafic sur cet ensemble de routes. Le choix de la meilleure route se fait de proche en proche en utilisant un algorithme de routage basé sur l'apprentissage par renforcement.

L'ordre quasi-linéaire de ce document correspond à une volonté d'amener le lecteur à mieux comprendre les différentes caractéristiques de chacune des solutions proposées. Ce document est organisé comme suit :

Le prochain chapitre de cette thèse présente les principaux éléments nécessaires à la compréhension du rapport, en particulier les notions de canal, cellule et de transfert inter-cellulaire (handoff). Nous exposons aussi les principaux schémas d'allocation de ressources implantés dans les réseaux cellulaires. Nous mettons l'accent sur les différentes techniques permettant la diminution des ruptures de communication.

Dans le troisième chapitre, nous exposons brièvement la notion d'apprentissage, en mettant l'accent sur l'apprentissage par renforcement.

Le chapitre quatre présente une nouvelle méthode de contrôle d'admission des appels (CAC) dans les réseaux cellulaires supportant plusieurs classes de trafic. La méthode proposée utilise l'apprentissage par renforcement, et a été réalisée pour le cas d'une seule cellule de type FCA, et le cas d'un système cellulaire complet de type FCA. Différentes implantations de l'algorithme d'apprentissage par renforcement ont été réalisées. Cette étude a été validée par un ensemble de simulations et comparée aux solutions existantes.

Le cinquième chapitre expose une nouvelle politique d'allocation dynamique des ressources radios, en considérant également le contrôle d'admission d'appels dans des systèmes cellulaires supportant plusieurs classes de trafic. La méthode proposée utilise, également,

l'apprentissage par renforcement et a été testée pour différents scénarios de simulations, et comparée aux solutions traditionnelles.

Le sixième chapitre présente une vision d'ensemble sur les réseaux ad hoc. Nous y détaillons les principaux algorithmes de routage ad hoc existants. Ces notions sont nécessaires pour la compréhension des deux derniers chapitres du document.

La consommation de l'énergie dans ce type de réseau est étudiée dans le chapitre 7. Nous arborons quelques propositions parmi les plus intéressantes, permettant de réduire cette consommation pour chacune des couches du modèle de référence. Nous mettons l'accent sur la couche réseau (routage) et comparons, par un ensemble de simulations, l'impact du routage sur la consommation de l'énergie.

Dans le dernier chapitre, nous présentons de nouveaux protocoles réactifs (LEAR-AODV, PAR-AODV, LPR-AODV et Q-AOMDV) permettant d'assurer la survivabilité du réseau, en réduisant autant que possible, la consommation des batteries dans les nœuds. Les procédures de découverte et de maintien de la route y sont décrites pour chaque protocole et sont comparés à celles utilisées dans AODV. Un ensemble de simulation a été réalisé pour comparer leurs performances.

Enfin, une synthèse globale reprenant l'ensemble du travail effectué pendant cette thèse, conclut ce document.



# Chapitre 2

## Réseaux cellulaires : Vue d'ensemble

### 1. Introduction

En 1876, *Graham Bell* invente le téléphone fixe, premier moyen de communication moderne. Onze ans plus tard, le physicien allemand *Heinrich Hertz* découvre les ondes radio. La première transmission radio fut réalisée en 1896 par le physicien italien *Guglielmo Marconi*, et le premier service de radiotéléphone a vu le jour aux Etats-Unis au début des années 50. Malgré une percée assez timide dans les années 60, les réseaux mobiles connaissent actuellement un énorme succès. L'avantage de tels systèmes est la possibilité de communiquer de n'importe où, même en se déplaçant. Cependant, l'utilisation de la voie hertzienne pour le transport de l'information a donné naissance à des architectures de réseau assez différentes de celles des réseaux fixes. L'une des raisons de cela est que la communication, dans les réseaux mobiles, doit continuer sans interruption, même en cas de déplacement de l'émetteur ou du récepteur. L'autre raison, est l'apparition de difficultés qui n'existaient pas lors des transmissions câblées, telles que la limitation naturelle de la bande passante, l'instabilité de la qualité du lien radio ou encore la variation des points d'accès au réseau.

Les systèmes cellulaires, sont sans aucun doute, ceux qui ont connu la plus grande évolution ces dernières années. Ce chapitre se propose de présenter les principaux éléments des réseaux cellulaires utiles à la compréhension de cette thèse tels que les notions de canal, de cellule, transfert inter-cellulaire, et d'allocation de ressources.

## 2. Notions de base

La fréquence radio est une ressource naturellement limitée, elle doit être alors judicieusement utilisée afin de permettre d'écouler la maximum de trafic. Ainsi, plusieurs méthodes d'accès multiple ont été définies, permettant le partage de la bande passante en plusieurs canaux. Ceux-ci seront alloués aux usagers afin de pouvoir communiquer avec le réseau. La définition d'un canal de communication dépend de la méthode d'accès choisie, il en existe principalement trois :

- Accès Multiple par Répartition dans les Fréquences (AMRF) ou Frequency-Division Multiple Access (FDMA) ;
- Accès Multiple par Répartition dans le Temps (AMRT) ou Time-Division Multiple Access (TDMA) ;
- Accès Multiple par Répartition par Code (AMRC) ou Code-Division Multiple Access (CDMA).

La combinaison des trois techniques d'accès (FDMA, TDMA et CDMA) est également envisagée.

### 2.1 La technique FDMA

La méthode d'accès FDMA [2][3][4][5][6], repose sur un multiplexage en fréquences et est essentiellement utilisée dans les réseaux analogiques. Le multiplexage fréquentiel divise la bande de fréquences en plusieurs sous-bandes. Chacune est placée sur une fréquence dite porteuse, ou carrier, qui est la fréquence spécifique du canal. Chaque porteuse ne peut transporter que le signal d'un seul utilisateur. La Figure 2.1 illustre un multiplexage FDMA de trois porteuses acceptant trois utilisateurs sur le même support. Cette méthode nécessite une séparation entre les porteuses pour éviter les interférences. Ce mode de partage est simple à mettre en œuvre et il ne nécessite pas de synchronisation entre l'émetteur et le récepteur. L'un des grands inconvénients de FDMA est la sous utilisation de la bande passante.

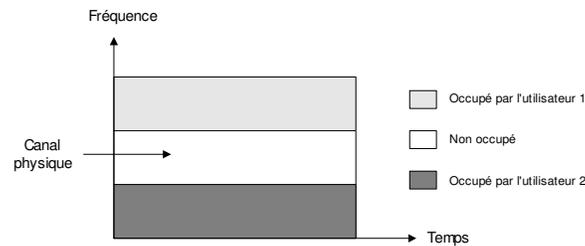


Figure 2.1 Le FDMA (Frequency-Division Multiple Access).

## 2.2 La technique TDMA

La méthode TDMA [2][3][4][5][6] offre la totalité de la bande de fréquences à chaque utilisateur pendant une fraction de temps donnée, appelée *time slot* (intervalle). TDMA permet de transmettre un débit plus important que la solution FDMA. Pour communiquer, un abonné utilise un slot (cf. Figure 2.2). Ainsi, plusieurs terminaux peuvent se partager la même porteuse. Pour permettre ce partage, les mobiles doivent être fortement synchronisés. Le TDMA s'applique principalement à la transmission de signaux numériques. Les accès combinés FDMA/TDMA montrent une efficacité spectrale importante.

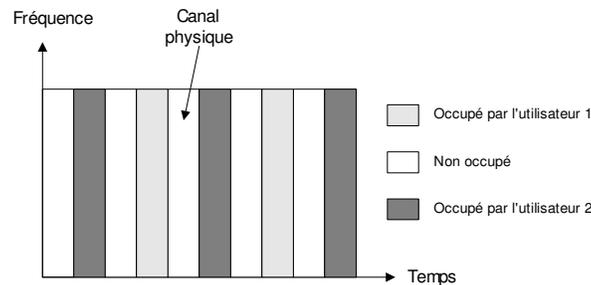


Figure 2.2 Le TDMA (Time-Division Multiple Access).

## 2.3 La technique CDMA

La troisième méthode, le CDMA autorise l'allocation de la totalité de la bande de fréquences, de manière simultanée, à tous les utilisateurs d'une même cellule. Pour ce faire, un code binaire spécifique est octroyé à chaque utilisateur. L'utilisateur se sert de son code pour transmettre l'information qu'il désire communiquer en format binaire d'une manière orthogonale, c'est-à-dire sans interférence entre les signaux, ou autres communications. Le nombre de canaux avec CDMA [2][3][4][5][6] est théoriquement illimitée ; toute la difficulté, cependant, est de trouver des codes suffisamment différents pour éviter les interférences et

permettre ainsi à la station de base de récupérer des émissions en parallèle. La combinaison FDMA/CDMA offre un bon niveau de résistance à l'interférence co-canal [5].

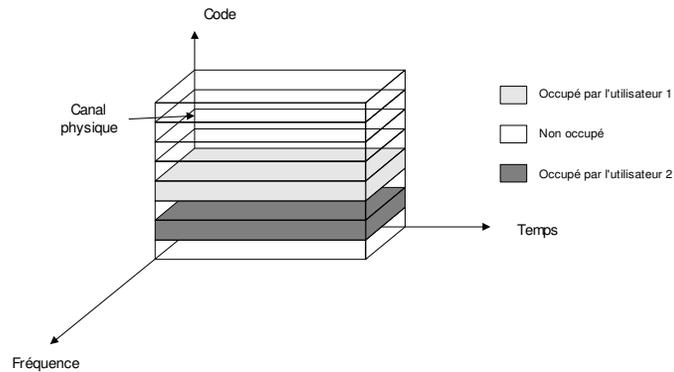


Figure 2.3 Le CDMA (Code-Division Multiple Access).

### 3. Les réseaux cellulaires

Les systèmes cellulaires (cf. Figure 2.4) offrent la possibilité d'acheminer automatiquement une communication jusqu'au destinataire quel que soit l'endroit où il se trouve, que le demandeur appartienne au réseau de télécommunications commuté ou au système cellulaire. Ils assurent le maintien de cette communication au cours du déplacement. Deux générations de ces systèmes coexistent, la première dite analogique tend à disparaître au profit de la seconde qui repose, elle, sur une technologie numérique.

- **Première génération** : apparus dans les années 70, les réseaux cellulaires analogiques, fournissant un service téléphonique à des usagers pouvant se déplacer sur de grandes distances. La taille des cellules était importante ce qui ne permettait d'écouler que peu de trafic. Cette génération utilisait la technique d'accès FDMA, et la transmission de la voix se faisait grâce à une modulation de fréquence ;
- **Deuxième génération** : profitant du développement des techniques de codage numérique de la parole, les réseaux cellulaires de cette génération autorisent l'utilisation de TDMA et CDMA comme une alternative à FDMA. Ainsi, le spectre radio est mieux rentabilisé et le nombre d'usagers possibles augmente considérablement.

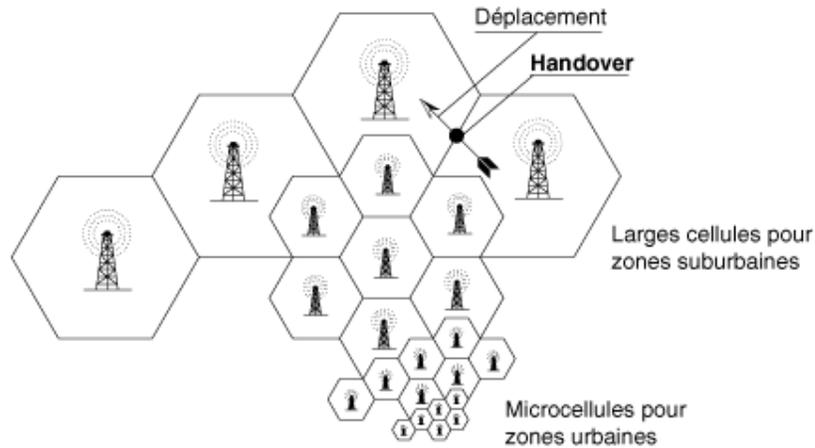


Figure 2.4 Exemple de réseau cellulaire.

De nouvelles générations de réseaux mobiles sont apparues [4] :

- **Troisième génération** : La troisième génération de mobiles, telle que l'UMTS (Universal Mobile Telecommunications System), propose de définir un système de communication sans fil visant à réaliser la convergence fixe-mobile. Les mobiles de la troisième génération sont des terminaux aux débits supérieurs à ceux des mobiles actuels. Ils sont capables, ainsi, d'offrir une importante gamme de services multimédias ;
- **Quatrième génération** : La quatrième génération de mobiles et sans fil se donne pour ambition non seulement d'améliorer le débit mais de mettre en commun la grande variété de solutions mobiles, souvent complémentaires entre elles, et de les proposer sous forme unifiée, dans un équipement terminal unique. La cinquième génération souhaite aller encore plus loin dans la même direction par une unification des interfaces radio, des techniques d'accès et des services.

### 3.1 Fondements

Le concept cellulaire repose sur une propriété essentielle des ondes radio qui est de s'atténuer en fonction de la distance. Une bande de fréquences utilisée sur un site peut, grâce à cette propriété, être utilisée sur un autre site à condition que ce dernier soit suffisamment éloigné du premier. Ce concept, combiné le plus souvent avec le mécanisme de réutilisation des fréquences, permet à des systèmes de communication radio de couvrir des zones très vastes et desservir des densités de trafic élevées [5].

Les réseaux cellulaires mobiles reposent sur deux principes fondamentaux à savoir :

- La cellule ;
- La réutilisation de fréquences.

Ils gèrent principalement deux fonctions de base :

- Le trafic intercellulaire (ou Handoff) ;
- L'itinérance (ou Roaming).

### 3.2 La cellule

La zone géographique desservie par le réseau mobile est divisée en petites surfaces appelées cellules. Chacune d'elles est couverte par un émetteur nommé "Station de base".

#### 3.2.1 Modélisation d'une cellule

La cellule est souvent schématisée par un hexagone ; cependant, cette forme ne peut exister dans la réalité. La Figure 2.5 montre ce qui peut ressembler à des cellules réelles. Contrairement à la forme circulaire idéale, l'hexagone régulier assure une couverture jointive du plan sans recouvrement. Il permet aussi un contour de niveau du signal constant depuis une antenne omnidirectionnelle. C'est pourquoi, les ingénieurs de planification utilisent cette forme lors de la conception des réseaux.

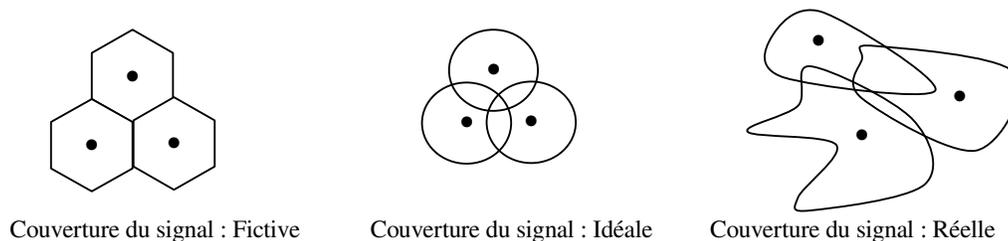


Figure 2.5 Formes de cellules.

#### 3.2.2 Taille de la cellule

La taille des cellules est très variable, elle peut varier de quelques mètres (pico-cellules) à quelques kilomètres (macro-cellules). Dans les zones denses, on trouve principalement des cellules de petites tailles (pico- ou micro-cellules), alors que dans les zones peu peuplées, on

utilise des macro-cellules. La tendance actuelle, dans les systèmes cellulaires, est l'utilisation de cellules de plus en plus petites pour écouler un maximum de trafic.

### 3.3 La réutilisation des fréquences

Les fréquences utilisées par une cellule peuvent être utilisées par une autre cellule suffisamment éloignée. Cela peut se faire grâce à la propriété d'atténuation des ondes radioélectriques. Ce principe permet d'augmenter significativement l'utilisation de la bande passante.

Ainsi, au lieu d'utiliser une seule antenne (1 cellule) pour tout le réseau, la zone de couverture est divisée en petites cellules regroupées en cluster (ou motifs cellulaires) dont la taille dépend du système. À chaque groupe ainsi formé est attribué l'ensemble des canaux du système. Les cellules utilisant les mêmes fréquences appelées co-cellules, doivent être situées à une distance suffisamment grande pour ne pas causer d'interférences dites "co-canal". La distance minimale requise est appelée distance moyenne de réutilisation. Dans la Figure 2.6, si par exemple, le système possède  $N$  canaux. Dans un système sans réutilisation, il pourrait y avoir au maximum  $N$  appels. En revanche, dans un système possédant  $K$  motifs de réutilisation,  $K \times N$  communications peuvent être acheminées.

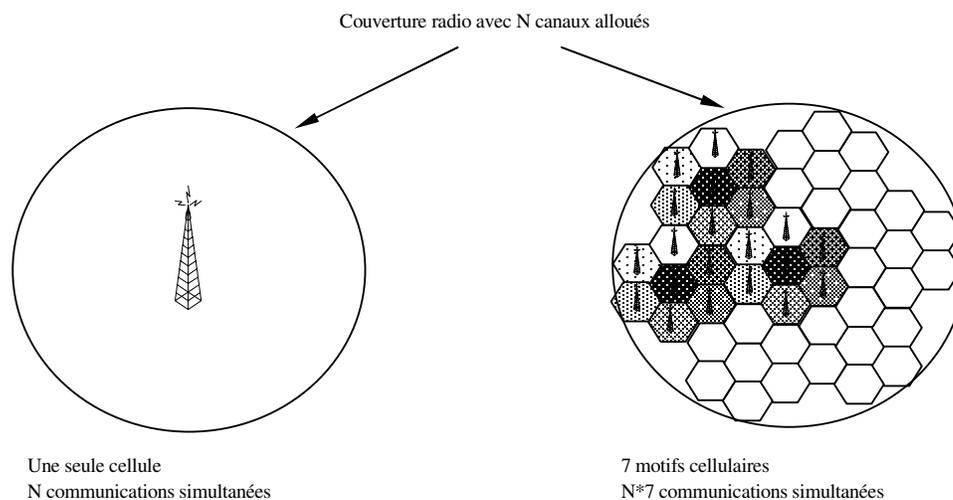


Figure 2.6 Gain de capacité réalisé par la réutilisation des fréquences.

### 3.3.1 Paramètres de réutilisation

#### 3.3.1.1 Définitions

- $C/I$  [5] : Indicateur de performance utilisé pour mesurer la qualité du signal reçu. Il dépend du signal utile  $C$  et du niveau d'interférence co-canal  $I$ . Ce rapport est affecté par des phénomènes aléatoires tels que la localisation du mobile, l'évanouissement de Rayleigh, les caractéristiques des antennes et la localisation des émetteurs/récepteurs ;
- $R$  : Rayon de la cellule. Il est fonction de la puissance d'émission de la station de base ;
- $D$  : Distance entre deux co-cellules.

Afin d'écouler le maximum de charge, il est nécessaire de réduire la distance de réutilisation au maximum. Cela ne peut se faire sans risque d'interférence. Lors de la conception du réseau et de la mise en place des stations de base, il est nécessaire de bien définir la distance de réutilisation ainsi que la taille optimale du motif. Ces deux nombres sont fonctions des rapports  $C/I$  et  $D/R$ .

Le nombre  $N$  de cellules par motif est donné par :  $N = D^2/(3R^2)$ .

Une grande valeur de  $N$  donne une distance de réutilisation importante, mais réduit le risque d'interférences co-canal. L'objectif dans un réseau cellulaire, est de trouver la plus petite valeur de  $N$  possible donnant un seuil d'interférence minimal requis par le système ; ceci permet d'obtenir une utilisation maximale de la bande passante [7].

En considérant que les cellules ont des rayons identiques, on obtient la relation suivante entre  $C/I$  et  $D/R$  [7] :

$$\frac{C}{I} = \frac{1}{\sum_{k=0}^{K-1} \left(\frac{D_k}{R}\right)^{-\gamma}} \quad (2.1)$$

Où  $K$  représente le nombre de cellules co-canaux de la première couronne et  $\gamma$  un paramètre caractérisant l'environnement.

Cette relation permet de constater que plus le rapport  $D/R$  est grand, plus le niveau d'interférence diminue. La valeur de  $C/I$  requise dépend du système utilisé ; c'est ainsi que pour FDMA, par exemple, le  $C/I$  minimal est de 18 dB. Par ailleurs, la connaissance de cette valeur

permet de déterminer la taille du motif. Dans le cas de FDMA, il est possible d'avoir un motif de 7 à 9 cellules [5].

### **3.4 Le transfert inter-cellulaire (ou Handoff)**

Une des procédures les plus complexes dans un réseau cellulaire est le transfert inter-cellulaire (ou Handoff). Il permet à un abonné de poursuivre sa communication sans coupure lorsqu'il change de cellule (Handoff inter-cellulaire – cf. Figure 2.4). Le mobile peut également effectuer un handoff à l'intérieur de la cellule si le signal devient faible par exemple (on parle alors de handoff intra-cellulaire).

Pour l'étude des réseaux cellulaires, plusieurs paramètres sont évalués. Parmi les plus importants nous pouvons citer : la probabilité de rejet des nouveaux appels (demandes de connexion rejetées) et la probabilité d'échec de transfert inter-cellulaire (se traduisant par une coupure de communication). Ainsi, un des soucis majeurs lors de la conception des réseaux est la réduction de la probabilité de coupure puisque, du point de vue de l'utilisateur, elle est beaucoup plus désagréable qu'un échec de connexion. Ceci est d'autant plus important qu'afin de répondre à la croissance des réseaux mobiles, la taille des cellules est de plus en plus réduite, ce qui augmente considérablement le nombre de transferts inter-cellulaires.

#### **3.4.1 Déroutement d'un handoff**

Le mécanisme de handoff peut être divisé en trois phases [8] : (i) mesure de la qualité du lien radio, (ii) initiation de la procédure, et (iii) exécution.

##### *3.4.1.1 Mesure de la qualité du lien radio*

Durant la communication, la station de base et le mobile effectuent des mesures sur le lien radio (la puissance du signal et le taux d'erreur bit). Les intervalles de mesure doivent être suffisamment petits pour permettre au réseau de réagir rapidement à la détérioration de la qualité du lien. Pendant cette phase, une liste de cellules candidates, susceptibles de prendre en charge la communication, est maintenue à jour. Une cellule est dans la liste si le signal qu'elle émet est reçu par le mobile avec une puissance au dessus d'un niveau prédéterminé.

##### *3.4.1.2 Initiation de la procédure de handoff*

Le processus de handoff est enclenché dès lors que la qualité du signal est au-dessous d'un seuil prédéterminé. L'initiation de ce mécanisme peut être faite par le réseau, comme le

cas d'un système analogique, ou bien par le mobile lui-même. Les systèmes cellulaires numériques utilisent les deux techniques.

#### 3.4.1.3 Exécution du handoff

Une fois le handoff décidé, la connexion est transférée vers un nouveau canal. Il existe trois mécanismes permettant de basculer d'un canal à l'autre [8] :

**Handoff doux (Soft handoff)** : À l'approche de la périphérie d'une cellule, les couvertures de deux stations de base voisines se chevauchent, et l'utilisateur distingue deux signaux forts pour sa transmission. S'il se connecte aux deux stations d'accès à la fois, on dit qu'il exécute un handoff doux. Ce procédé consomme deux fois plus de ressources, mais le passage d'une cellule à une autre est confortable pour l'utilisateur. Ce type de handoff a été introduit dans le système CDMA ;

**Handoff dur (Hard handoff)** : Se produit lorsque le changement de cellules s'accompagne d'une modification de la fréquence porteuse du signal. Le handoff n'est pas forcément synonyme de changement de fréquence, et il est possible de faire un handoff sur la même fréquence en changeant uniquement le slot. La communication et le routage des données s'effectuent simultanément sur le nouveau lien ; cela induit une coupure durant la transmission. Ce type de handoff est complètement géré par le réseau, il s'adapte bien à la technique FCA (décrite ci-dessous) et il est utilisé dans le réseau GSM. L'avantage de cette procédure est que le mobile ne monopolise qu'un canal à la fois ;

**Seamless Handoff** : Le nouveau lien est établi en parallèle à l'ancien. Le mobile transmet sur l'ancien et le nouveau lien en utilisant un partage temporel. Quand le réseau effectue la communication l'ancien lien est libéré. Ce handoff permet de minimiser la probabilité de coupure pendant le transfert, il s'adapte bien à la technique DCA (décrite ci-dessous) ;

**Handoff analogique** : Dans certains pays, tels que les Etats-Unis, les systèmes analogiques de première génération sont toujours opérationnels et cohabitent avec ceux entièrement numériques. Un abonné de la deuxième génération peut donc traverser une cellule dotée d'une transmission numérique pour se rendre dans une zone couverte par un système analogique. Si les deux systèmes l'acceptent, et que son terminal le permette, l'abonné exécute un handoff analogique pour assurer la poursuite de la connexion.

### 3.4.2 L'itinérance ou le Roaming

L'itinérance englobe toutes les fonctions du réseau mobile permettant de localiser géographiquement un abonné. Cette procédure joue un rôle très important puisqu'elle permet d'acheminer les communications de/vers un usager du réseau. En effet, contrairement aux réseaux fixes, où un numéro de téléphone correspond à une adresse physique, un numéro de terminal mobile possède une adresse logique à laquelle il faut faire correspondre une adresse physique qui varie lors des déplacements [5].

La méthode de localisation la plus répandue actuellement divise l'ensemble des cellules en zones de localisation. Le réseau doit connaître en permanence la zone dans laquelle se trouve le mobile. Ainsi, dès la mise sous tension du terminal, le réseau enregistre sa localisation actuelle dans une base de données. Pour mener à bien cette gestion, le réseau maintient généralement deux types de bases de données [5] :

- Une base de données principale nommée HLR (*Home Location Register*), qui stocke les informations de tous les abonnés du réseau (noms, droits d'accès, etc.). La recherche de l'abonné commence toujours par cette base.
- Plusieurs bases de données visiteurs nommées VLR (*Visitor Location Register*). Il existe une base de données de ce type par zone de localisation. Chaque VLR recopie du HLR les informations concernant les abonnés se trouvant dans sa zone de localisation.

## 4. Schémas d'allocation de canaux

Dans les réseaux cellulaires, la bande passante est divisée en un ensemble disjoint de canaux radio qui peuvent être utilisés simultanément, à condition de maintenir une qualité du signal radio acceptable. Cette division peut se faire grâce à différentes techniques telle que FDMA, TDMA, CDMA ou toutes combinaisons de ces méthodes. Il existe principalement trois méthodes :

- Fixed Channel Allocation (FCA) ;
- Dynamic Channel Allocation (DCA) ;
- Hybrid Channel Allocation (HCA).

Ces méthodes tentent de maintenir la ressource radio entre les différentes cellules afin d'obtenir une utilisation efficace du point de vue qualitatif et quantitatif. Cette partie décrit brièvement ces mécanismes d'allocation de canaux, leurs performances respectives (Figure 2.7), ainsi que quelques méthodes permettant d'améliorer la QoS des utilisateurs.

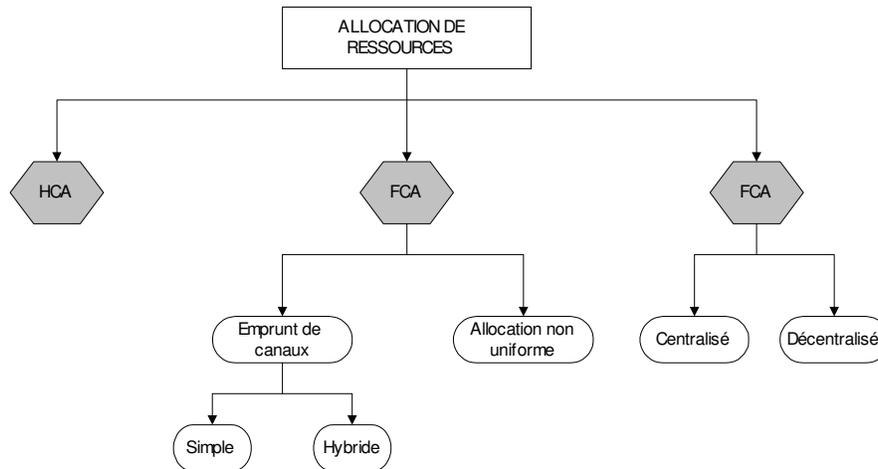


Figure 2.7 Techniques d'allocation de canaux.

## 4.1 Fixed Channel Allocation

FCA est une méthode adaptée aux systèmes cellulaires utilisant FDMA et TDMA. Elle reste la technique la plus utilisée dans les réseaux mobiles actuels. Elle consiste à allouer d'une manière quasi-permanente des canaux à des cellules, selon un plan de fréquence déterminé lors de la planification du réseau. Toute modification de cette allocation entraîne une re planification des fréquences. L'un des principaux avantages de cette allocation est qu'elle affecte les canaux aux cellules en respectant une distance de réutilisation optimale.

La technique FCA de base attribue le même nombre de canaux aux différentes cellules, ce qui se révèle efficace à condition que le trafic soit uniforme. Cependant, de grandes pertes peuvent être causées s'il y a des différences de charge entre les stations de base. En effet, la demande de trafic peut excéder la capacité de certaines cellules alors qu'au même moment d'autres sites du système ont leurs canaux sous-utilisés. Pour pallier à cela, des techniques d'allocation non-uniformes ou des techniques d'emprunt de canaux sont proposées.

### 4.1.1 Allocation non-uniforme

L'allocation non-uniforme [9][10] consiste à attribuer aux cellules des canaux en fonction du trafic qu'elles écoulent. C'est ainsi qu'un plus grand nombre de canaux est alloué aux

cellules plus chargées. Les auteurs de [9], proposent des algorithmes d'allocation non-uniformes permettant de diminuer les blocages du réseau et d'augmenter le trafic écoulé.

#### 4.1.2 Emprunt de canaux

Une cellule ayant épuisé toute sa capacité de trafic demande à ses voisines des canaux pour pouvoir les attribuer à ses propres appels. Cet emprunt ne peut se faire que s'il ne cause aucune interférence. À la fin de l'utilisation, les canaux sont rendus à la cellule initiale. De nombreux mécanismes d'emprunt existent ; ils diffèrent les uns des autres par le choix des canaux à emprunter. Deux familles d'algorithmes [11][12] peuvent être distinguées :

##### 4.1.2.1 Algorithmes d'emprunt de canaux simples

Si tous les canaux d'une cellule sont occupés, un canal est emprunté aux cellules voisines s'il ne cause pas d'interférences. Les canaux ainsi empruntés sont verrouillés pour toutes les cellules où ils pourraient causer des interférences. Plusieurs stratégies d'emprunt existent [11], comme par exemple :

- Choisir un canal de la cellule la plus riche en canaux libres ;
- Choisir un canal de la cellule la plus riche en canaux libres non verrouillés ;
- Transférer, dès que possible, la communication d'un canal emprunté à un canal propre.

##### 4.1.2.2 Algorithmes d'emprunt de canaux hybrides

Les canaux d'une cellule sont divisés en deux groupes, les premiers ne pouvant être utilisés que par la cellule elle-même et les seconds pouvant être prêtés aux autres sites du réseau. Le *ratio* entre les deux groupes est déterminé en fonction des paramètres de trafic [12][13]. Ici encore, différentes stratégies existent [11], on peut envisager entre autre de :

- faire varier le *ratio* entre les deux groupes en fonction des conditions de trafic ;
- donner des priorités aux canaux pouvant être empruntés ;
- déterminer, pour chaque cellule, les canaux pouvant être empruntés par chacune de ses voisines [13].

### 4.1.2.3 Performances des algorithmes d'emprunt

Les algorithmes d'emprunt reposent, pour la plupart, sur le principe de verrouillage des canaux ce qui présente de nombreux inconvénients. Un des plus importants, est que le nombre de canaux pouvant être empruntés est limité à ceux qui ne sont pas verrouillés et qui ne sont pas utilisés dans les cellules se trouvant au-delà de la distance de réutilisation minimale. Il est en outre difficile de maintenir la distance de réutilisation optimale dans un tel système [13].

Pour des charges faibles et moyennes, les algorithmes d'emprunt présentent de meilleures performances que FCA. Cependant, pour des charges importantes, FCA est meilleur. En effet, dans ce cas, le nombre d'emprunts augmente, et de nombreux canaux libres ne peuvent plus être utilisés pour respecter la distance de réutilisation minimale requise. Cette raison fait aussi que les algorithmes hybrides sont meilleurs que les algorithmes simples en forte charge.

## 4.2 Dynamic Channel Allocation

L'allocation dynamique [11][14] est une technique actuellement utilisée dans les systèmes sans cordon tel que DECT [15]. Dans DCA, tous les canaux, ou un ensemble de canaux, disponibles dans le système sont mis dans un pool commun utilisable par toutes les stations de base. Lors d'une demande de communication, une cellule choisit un canal du pool commun qui sera restitué à la fin de l'appel. Le choix du canal n'est pas aléatoire, mais repose sur le calcul d'interférences. Dans le cas de disponibilité de plusieurs canaux pour la cellule, plusieurs stratégies de sélection peuvent être appliquées [16][17][18][19][20][21]. Nous proposons, dans le Chapitre 5, une méthode permettant de choisir le meilleur canal disponible.

Les algorithmes DCA peuvent être divisés en deux catégories [11] : (i) algorithmes centralisés et (ii) algorithmes distribués.

### 4.2.1 Algorithmes centralisés

Dans les algorithmes d'allocation centralisés, les canaux sont alloués aux cellules par un contrôleur centralisé. Pour effectuer cette opération, ce dernier utilise toutes les mesures effectuées sur les fréquences par les mobiles et les stations de base. Ce type d'algorithme permet une allocation optimale avec un risque de surcharge du réseau. La différence entre les différentes propositions existantes est la fonction coût utilisée pour le choix du meilleur canal. La stratégie appelé 'First Available – FA' [18], par exemple, alloue le premier canal disponible satisfaisant la distance de réutilisation. Dans la stratégie LODA (Locally Optimized Dynamic Assignment) [17] la fonction coût choisie est basée sur la future probabilité de blocage du

voisinage de la cellule en laquelle un appel a été initié. Dans la stratégie RING [18], le canal choisi est celui le plus utilisé dans l'ensemble des cellules co-canaux.

Il existe dans la littérature plusieurs autres propositions d'algorithmes d'allocation centralisés. Des études de performances ont été faites afin de les comparer [18][11].

#### 4.2.2 Algorithmes distribués

Plusieurs résultats de simulation [22][23][24][25] ont démontré que les stratégies d'allocation centralisés peuvent produire une allocation de canaux proche de l'optimum, mais aux dépens d'une surcharge, due à la centralisation, très élevée. Dans les algorithmes d'allocation distribués, la cellule choisit le canal qu'elle peut prendre en fonction des informations qu'elle possède. Ces stratégies sont donc plus attrayantes, en raison de leur simplicité d'implémentation dans chaque station de base. Les stratégies d'allocation dynamique proposés utilisent généralement soit : (i) des informations locales sur les canaux disponibles au voisinage de la cellule (cell-based) [26][27], soit (ii) des mesures sur la puissance du signal [28][29][30].

##### 4.2.2.1 Stratégies basées sur des informations sur cellules (cell-based)

Dans les stratégies cell-based, un canal est alloué à un appel par la station de base dans laquelle l'appel à été initié. La différence avec l'approche centralisée est que chaque station de base maintient des informations sur les canaux disponibles dans le voisinage. L'information sur le statut des canaux est mise à jour grâce un échange régulier d'informations entre les stations de base. La stratégie cell-based fournit une allocation de canaux proche de l'optimum aux dépens d'un échange excessif d'informations de statut entre les stations de base.

##### 4.2.2.2 Stratégies basées sur la puissance du signal

Ce sont des politiques DCA qui adaptent leurs allocations selon les mesures du rapport signal/interférence  $C/I$ . Un exemple de système basé sur ce principe est le standard européen de téléphonie sans fil DECT (Digital Enhanced Cordless Telecommunications). La station de base n'utilise que des informations locales, sans avoir à communiquer avec n'importe quelle autre station de base dans le réseau. Ainsi, les stations de base et les mobiles estiment le rapport signal/interférence  $C/I$ , et allouent un canal au nouvel appel tant que ce rapport est au-dessus d'un certain seuil. Ces stratégies permettent un traitement en temps réel assez rapide mais aux dépens d'une forte probabilité d'interférence co-canal avec les appels des cellules adjacentes. En fait, il est possible que l'allocation d'un canal au nouvel appel, provoque une détérioration du rapport signal/interférence des appels déjà établis.

### 4.3 Comparaison

La demande de communication est le résultat d'un processus aléatoire, l'attribution des canaux dans le cas dynamique est aussi aléatoire. Dans un tel cas, les cellules utilisant la même fréquence sont souvent plus éloignées que la distance minimale requise. En conséquence, le réseau n'est pas toujours capable d'attribuer, en même temps, un canal le maximum de fois possible. Dans une allocation fixe, les canaux sont alloués au départ aux cellules de manière optimale. Ces raisons font qu'en cas de fortes charges, FCA est meilleure, puisqu'elle permet d'écouler le maximum de trafic. Cependant, pour les faibles et moyennes charges, DCA utilise les canaux d'une manière plus efficace, surtout s'il y a des disparités dans le trafic des cellules, alors que FCA peut engendrer des blocages même s'il existe des canaux libres dans le système.

De plus, le mécanisme FCA engendre un nombre plus important de handoffs. En effet, lors d'un transfert inter-cellulaire, le mobile -dans le cas de DCA- ne change pas le canal si les conditions d'interférence le permettent, alors que pour FCA un changement de cellule entraîne un changement de canal. En conséquence, les pertes de handoff, dans le cas d'une allocation fixe, sont beaucoup plus importantes, surtout dans un contexte micro-cellulaire où la procédure de changement de cellule est très sollicitée.

À la lumière de cette comparaison entre allocation fixe et dynamique, il est naturel de penser qu'une méthode combinant les deux donnerait de très bons résultats. De nombreux auteurs ont ainsi suggéré des mécanismes hybrides permettant d'avoir les avantages des deux.

### 4.4 Hybrid Channel Allocation

Lors d'une allocation hybride, les canaux sont divisés en deux catégories : fixe et dynamique. À chaque cellule est attribué un ensemble de canaux permanents et le reste est mis dans un pool commun utilisable par toutes les stations de base. Quand une cellule épuise ses canaux propres, elle utilise les canaux se trouvant dans le pool commun. Un des paramètres importants de ce mécanisme est le ratio entre le nombre de canaux fixes et le nombre de canaux dynamiques. Généralement, l'optimum est fonction de la charge.

Les performances de HCA sont proches de celles de DCA. Pour les mêmes raisons que celles citées dans la section 4.3, HCA présente de bons résultats dans un système moyennement chargé, alors que pour les fortes charges l'allocation fixe est meilleure.

## 5. Diminution des blocages

Les techniques d'attribution des canaux présentées ne prennent pas en compte l'impact des handoffs sur les performances. Comme il a été décrit précédemment, le handoff est défini comme le changement de canal de transmission du mobile. Généralement, ce changement est dû à la dégradation du signal radio ou à une réorganisation des canaux pour éviter la congestion. Le nouveau canal appartient soit à la station de base courante (dans le cas d'un handoff intra-cellulaire), soit, à une autre station de base (handoff inter-cellulaire).

Dans cette partie, notre attention porte uniquement sur les transferts inter-cellulaires. Lorsqu'un mobile change de site et qu'il ne trouve pas de canal disponible, la communication est coupée et le handoff échoue. Dans un système composé de cellules de petite taille, la procédure de handoff a un impact très important sur les performances du réseau. Il est donc important de limiter les échecs de transfert inter-cellulaire, d'autant que, du point de vue de l'utilisateur, il est préjudiciable de perdre la communication en cours que de ne pas être accepté au départ.

Ainsi, dans un réseau mobile, implanter des mécanismes permettant de réduire la probabilité de coupure de communication devient très important. De nombreuses méthodes sont supposées ; elles permettent de favoriser les handoffs par rapport aux nouveaux appels. On obtient ainsi de bonnes performances, mais, en contrepartie, le trafic écoulé est généralement réduit et la probabilité d'échec des nouveaux appels est augmentée. Ces mécanismes utilisent globalement 5 principes :

- réservation de canaux à l'usage des handoff ;
- réservation de canaux pour les handoffs avec mise en attente des nouveaux appels ;
- mise en attente de handoffs ;
- contrôle d'admission ;
- recouvrement des cellules.

Les sections suivantes décrivent brièvement quelques unes des méthodes les plus citées dans la littérature.

## 5.1 Réserveation des canaux

Introduites dans les années 80 [12], les techniques de réserveation de canaux (souvent nommées Guard Channel Schemes) permettent de réduire la probabilité de coupure de communication en réservant simplement, dans chaque cellule, des canaux à l'usage exclusif des handoffs. Le reste des ressources peut être utilisé pour tous les types d'appels. L'inconvénient de cette approche est évidemment la perte de trafic offert, qui peut être réduite grâce à un bon dimensionnement des canaux de garde et/ou à une mise en attente des nouveaux appels [31].

Les techniques Guard channel peuvent être utilisées avec FCA ou DCA. Dans ce dernier schéma, les canaux réservés ne sont pas alloués à une cellule particulière mais à toutes les cellules utilisant le même pool commun [11].

### 5.1.1 Dimensionnement des cellules

Le facteur le plus important dans les mécanismes de réserveation est le choix du nombre de canaux à allouer pour les handoffs, de manière à obtenir un bon niveau de QoS. En effet, une sur-estimation de ce nombre peut causer des pertes importantes de bande passante, principalement dans le cas de FCA. Certains auteurs [31][32] suggèrent une réserveation du même nombre de canaux par cellule. Ce pendant, cette uniformité ne tient pas compte de la dissymétrie du trafic qui peut exister dans un réseau cellulaire et peut ainsi engendrer des pertes. Par ailleurs, d'autres travaux se sont intéressés à des mécanismes de réserveation de canaux en fonction du trafic.

Ainsi, les auteurs de [9] montrent que le blocage des nouveaux appels dans une cellule, croît avec une réserveation et décroît à l'ajout d'un canal. Suite à cette constatation, ils développent des algorithmes qui permettent d'attribuer à chaque cellule des canaux propres (utilisables par tout type de trafic) et des canaux pour les handoffs, de manière à satisfaire un niveau de QoS prédéfini. La méthode est appliquée à un ensemble de cellules, ce qui permet d'avoir un FCA non uniforme avec les réserveations. Cependant, les auteurs améliorent la QoS de chaque cellule indépendamment les unes des autres, alors que dans un système réel, l'amélioration d'une cellule peut entraîner la détérioration de ses voisines.

Le travail cité dans [33] propose quant à lui, une méthode permettant de déterminer instantanément le nombre de canaux à réserver, en fonction du nombre d'appels courant dans les cellules. Ce mécanisme repose sur la notion d'intervalle d'observation qui représente une tranche de temps où il ne peut y avoir qu'un seul changement dans une cellule (augmentation ou

diminution du nombre de mobiles). Ainsi, cette méthode suggère qu'après une période d'observation, chaque cellule prévient ses voisines des modifications éventuelles qui se sont produites sur son territoire. Les informations ainsi collectées permettent à chaque site, d'estimer le trafic entrant au cours de la tranche de temps suivante, et ainsi de calculer la réservation adéquate pour maintenir la QoS à un niveau prédéterminé. L'inconvénient majeur de cette proposition est qu'elle risque de générer beaucoup de signalisation, en particulier lors de fortes charges puisque les périodes d'observation doivent être petites pour couvrir tous les changements.

## **5.2 Mise en attente de nouveaux appels**

La mise en attente des nouveaux appels est plus simple à mettre en œuvre que celles des handoffs [36]. En effet, ces derniers sont plus sensibles aux délais que les nouvelles communications. Dans [17], les auteurs proposent de réserver des canaux pour le handoff et de mettre en attente les nouvelles communications. Les canaux de garde permettent de réduire le blocage alors que l'attente augmente le trafic offert. Les auteurs considèrent qu'il n'y a pas de délai maximal d'attente dans une file d'attente, parce qu'un client s'élimine automatiquement grâce à un 'time-out' interne. Cette hypothèse est prise en compte dans les calculs ce qui améliore les performances. En effet, les auteurs ne déterminent pas le blocage des nouveaux appels mais le délai d'attente. Les résultats obtenus montrent une augmentation sensible du trafic écoulé avec une réduction des différents blocages.

## **5.3 Mise en attente des handoffs**

La mise en attente des handoffs [32][34][35], avec ou sans réservation, permet de réduire la probabilité des transferts inter-cellulaires. Un nouvel appel n'est alors admis que si la file d'attente est vide.

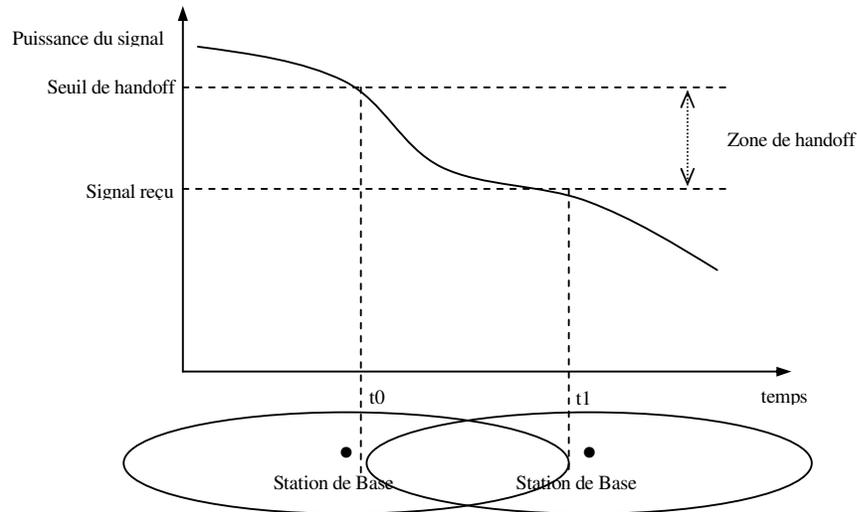


Figure 2.8 Zone de handoff.

Un tel mécanisme est possible grâce au recouvrement des cellules qui permet d'avoir des zones dans lesquelles les mobiles peuvent physiquement communiquer avec deux stations de base (Figure 2.8). La puissance du signal, dans cet intervalle, est entre le seuil d'initialisation des handoffs et le seuil de réception de la station de base. Lors du séjour d'un appel dans cet espace, si aucune ressource n'est disponible dans la cellule destination, une demande de communication est mise dans une file d'attente. Un client quitte cette file lorsqu'un canal devient disponible ou lors de sa sortie de la zone de handoff.

La mise en œuvre de l'attente des handoffs est délicate [36]. Pour pouvoir l'étudier il est important de bien estimer les paramètres de trafic tels que le temps de communication, le temps d'attente dans la file, la vitesse de déplacement, etc. Ainsi deux travaux [34][35] se basant sur des hypothèses différentes ont abouti à des résultats contradictoires. En effet, les auteurs comparent deux types de files : une file FIFO et une file prioritaire non préemptive. Dans cette dernière, la priorité est donnée aux clients dont la qualité du signal est mauvaise. Les résultats de [35] montrent que le mécanisme prioritaire est plus performant que FIFO alors que [34] montrent qu'il y a peu de différence entre ces deux politiques de gestion.

Selon [36], il est évident que si les handoffs peuvent attendre longtemps, une file d'attente prioritaire permet d'avoir de meilleurs résultats qu'une file FIFO. Cependant, si le signal des mobiles se dégrade rapidement, ce qui revient à dire que les handoffs ne peuvent pas attendre longtemps, les deux politiques sont équivalentes. En effet, dans ce dernier cas, le premier d'une

file FIFO va certainement être le plus prioritaire. La question primordiale de ce mécanisme est alors combien peut-on réellement faire attendre les handoffs dans un cas pratique ?

#### 5.4 Contrôle d'admission

Le principe du contrôle d'admission [37][38] repose sur la notion de *Cell-Cluster* qui représente un groupe de cellules adjacentes reliées à un même contrôleur d'appel. Un nouveau mobile est accepté par le contrôleur si le nombre de communications dans le cluster n'a pas atteint un niveau prédéfini. Cette décision est prise en fonction d'informations globales. L'objectif est une généralisation de la réservation pour un groupe de cellules. L'objectif du contrôle d'admission est de "garantir" aux communications admises dans un cluster une faible probabilité de coupure. Dans ces travaux, les auteurs se sont également intéressés à des trafics ayant des besoins de QoS différentes.

#### 5.5 Recouvrement des cellules

La représentation hexagonale disjointe des cellules est uniquement utilisée dans le cadre d'études théoriques. Dans la réalité, les cellules se recouvrent les unes et les autres. Ce recouvrement peut permettre à une partie des mobiles se trouvant sur un site, de pouvoir transmettre sur les sites voisins. Les auteurs de [39][40][41][43] exploitent cette caractéristique et proposent des mécanismes qui réduisent les blocages sans trop perdre en terme de trafic offert. Les méthodes les plus communs sont : *Directed Retry* (DR) [39][41][43] et *Directed Handoff* (DH) [43].

- **Directed Retry** : un mobile qui ne trouve pas de canal dans sa cellule, en cherche un autre dans une des cellules voisines qu lui fournirait une qualité de signal suffisante ;
- **Directed handoff** : une cellule qui a presque atteint sa capacité de trafic maximale force une partie de ses clients à effectuer un handoff vers les cellules voisines. Cette opération ne peut se faire que si la qualité du signal est suffisante.

Les méthodes DR et DH reposent sur la technique FCA et permettent théoriquement de diminuer les pertes. Cependant, il est difficile d'estimer réellement le nombre de mobiles pouvant être servi par deux ou plusieurs stations de base. Les performances réelles pourraient être moins optimistes que prévues [36].

## 6. Conclusion

Dans ce chapitre, nous avons donné un aperçu sur les différentes solutions qui ont été proposées dans la littérature, permettent de réduire les probabilités de coupure de communication.

Dans toutes ces propositions, la politique qui doit être suivie par le gestionnaire de ressources, de chaque cellule, est celle où un certain nombre de canaux radios sont exclusivement réservés pour les handoffs. Cette politique est optimale lorsqu'on considère une seule classe de trafic (i.e. appel téléphonique).

Toutefois, dans un contexte multiservices, cette politique n'est pas optimale. Nous proposons dans la suite de cette thèse la recherche de la politique optimale en utilisant l'apprentissage par renforcement.

# Chapitre 3

## L'Apprentissage

### 1. Introduction

Donner une définition précise de l'apprentissage reste un problème délicat et difficile, comme le note [44] : "L'apprentissage je sais ce que c'est tant qu'on ne me demande pas de le définir". Mais, une des définitions qui nous paraît assez complète vient du domaine de l'Intelligence Artificielle [45] : le terme apprentissage désigne, la capacité à organiser, à construire et à généraliser des connaissances pour une utilisation ultérieure. C'est donc la capacité à tirer profit de l'expérience pour améliorer la résolution d'un problème.

L'apprentissage automatique peut être de différents types. On distingue, entre autres, l'apprentissage numérique et l'apprentissage symbolique. Le premier met en œuvre des données quantitatives et peut ainsi permettre de résoudre des problèmes numériques complexes, mais ne permet pas de gérer aisément des données multi-relationnelles, ni de produire des connaissances explicites. L'apprentissage symbolique présente des avantages et inconvénients sensiblement opposés. Nous nous intéressons dans cette thèse à l'apprentissage numérique. Pour voir les différentes techniques utilisées dans les deux approches, l'ouvrage de *T. Mitchell* [45] nous paraît être une des meilleures références. Dans ce chapitre, nous allons, tout d'abord, aborder l'apprentissage numérique dans sa généralité, ainsi que les grandes catégories d'approches d'apprentissage numérique existantes. Nous énumérerons, ensuite, quelques applications de cet apprentissage à différents domaines et, en particulier, au domaine des réseaux de télécommunications. Enfin, nous détaillons l'approche apprentissage par renforcement.

## 2. Apprentissage numérique

Selon le type d'informations disponibles dans l'apprentissage numérique, deux grandes catégories d'approches peuvent être distinguées. La première qualifiée d'apprentissage *non-supervisé* vise à regrouper des objets en classes, en se basant sur des ressemblances entre eux. La deuxième approche est l'apprentissage *supervisé*, qui se base quant à elle sur un ensemble d'apprentissage constitué d'objets dont la classe est connue à priori.

### 2.1 Apprentissage supervisé

Dans ce type d'apprentissage, un maître (ou superviseur, d'où le nom d'apprentissage supervisé) fournit soit l'action qui devrait être exécutée, soit un gradient sur l'erreur commise (cf. Figure 3.9). Dans les deux cas, le maître fournit au contrôleur une indication sur l'action qu'il devrait générer afin d'améliorer ses performances. L'utilisation d'une telle approche présuppose l'existence d'un expert capable de fournir un ensemble d'exemples, appelé base d'apprentissage, formés de situations et d'actions correctes associées. Ces exemples doivent être représentatifs de la tâche à accomplir.

L'une des variantes de l'apprentissage supervisé, dans lequel une "critique" de la réponse calculée est fournie au réseau, est l'apprentissage par renforcement (Reinforcement Learning – RL). C'est cette variante d'algorithmes qui nous a paru la plus adaptée pour résoudre les problèmes traités dans cette thèse. Cette méthode est détaillée plus bas.

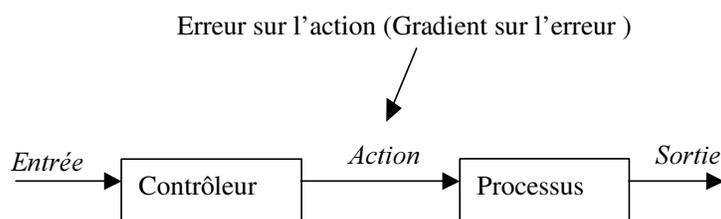


Figure 3.9 Apprentissage supervisé.

### 2.2 Apprentissage non-supervisé

L'apprentissage non-supervisé, appelé aussi apprentissage à partir d'observations, consiste à déterminer une classification à partir d'un ensemble d'objets ou de situations données. On dispose d'une masse de données indifférenciées, et l'on désire savoir si elles possèdent une quelconque structure de groupes. Il s'agit d'identifier une éventuelle tendance

des données à être regroupées en classes. Ce type d'apprentissage appelé 'Clustering', est retrouvé dans la classification automatique et dans la taxinomie numérique. Il recherche des régularités parmi un ensemble d'exemples, sans être nécessairement guidé par l'utilisation qui sera faite des connaissances apprises. Il regroupe l'ensemble des exemples de manière à ce que les exemples au sein d'un même groupe se ressemblent suffisamment, et que les exemples de groupes différents soient suffisamment différents.

### 2.3 Apprentissage par renforcement

L'apprentissage par renforcement (dit aussi apprentissage semi-supervisé) est une variante de l'apprentissage supervisé [47]. Par opposition à l'approche supervisée, le maître dans l'apprentissage par renforcement a un rôle d'évaluateur et non pas d'instructeur. Il est en général appelé critique. Le rôle du critique est de fournir une mesure indiquant si l'action générée est appropriée ou non. Il s'agit de programmer un agent au moyen d'une évaluation par pénalité/récompense sans avoir besoin de spécifier comment la tâche doit être remplie. Dans ce cadre, on doit indiquer au système quel est le but à atteindre, et celui-ci doit apprendre par une succession d'essais/erreurs (en interaction avec l'environnement) comment atteindre le but fixé.

Les composants de l'apprentissage par renforcement sont l'"apprenti" agent, l'environnement où il agit ainsi que la tâche qu'il doit réaliser (cf. Figure 3.10). L'interaction entre l'agent et l'environnement est continue. D'une part le processus de décision de l'agent choisit des actions selon les situations perçues de l'environnement, et d'autre part ces situations évoluent sous l'influence de ces actions. Chaque fois que l'agent effectue une action, il reçoit une récompense. Celle-ci est une valeur scalaire indiquant à l'agent la conséquence de cette action.

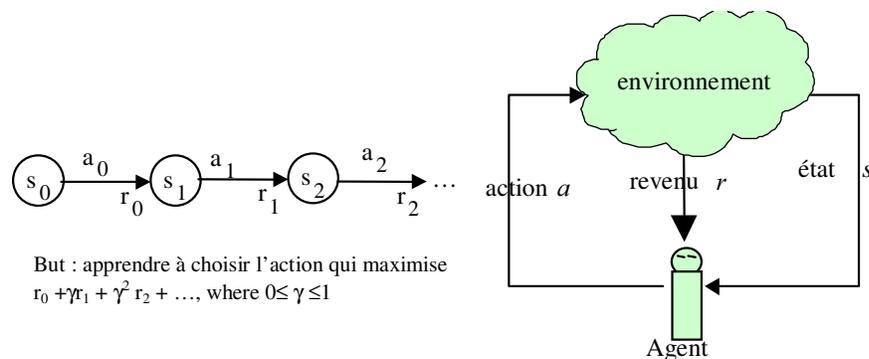


Figure 3.10 Interaction Agent-Environnement.

Pour être plus formel, dénotons  $s$  ( $\in S$ , un ensemble fini), une représentation de l'état actuel de l'environnement,  $a$  ( $\in A$ , un ensemble fini) l'action choisie et  $r$  ( $\in R$ , un ensemble fini) la récompense reçue. L'interaction entre l'agent et l'environnement consiste, à chaque instant, en les séquences suivantes :

- (i) l'agent observe l'état actuel de l'environnement  $s_t \in S$  ;
- (ii) en se basant sur l'état  $s_t$ , l'agent prend une décision en exécutant une action  $a_t \in A$  ;
- (iii) l'environnement fait alors une transition vers un nouvel état  $s_{t+1} = s' \in S$  suivant la probabilité  $P_{ss',(a_t)}$  ;
- (iv) l'agent reçoit instantanément un certain revenu  $r_t = r(s_t, a_t)$  indiquant la conséquence de cette décision.

Le processus de décision de l'agent s'appelle *politique* et c'est une fonction de l'ensemble des états vers l'ensemble des actions ( $\pi: S \rightarrow A$ ). L'agent doit apprendre une politique,  $\pi$ , permettant de choisir la prochaine action  $a_t = \pi(s_t)$  à effectuer, et ceci en fonction de l'état actuel  $s_t$ . L'interaction entre l'agent et l'environnement est continue et l'apprenti agent modifie sa politique selon son expérience et selon le but consistant à maximiser le cumul des récompenses dans le temps. Ce cumul  $V^\pi(s_t)$ , réalisée en suivant une politique arbitraire  $\pi$ , à partir d'un état initial  $s_t$ , est définie comme suit :

$$V^\pi(s) = E \left\{ \sum_{t=0}^{\infty} \gamma^t r(s_t, \pi(s_t)) \mid s_0 = s \right\} \quad (3.1)$$

où  $E$  désigne l'opérateur espérance et le facteur  $0 \leq \gamma \leq 1$  représente la constante de propagation temporelle.

L'objectif, pour l'agent est donc de maximiser cette somme des renforcements reçus, et son apprentissage s'effectue par de nombreuses expériences. L'agent est guidé en cela par divers algorithmes cités dans le paragraphe 2.3.1 suivant.

### 2.3.1 Méthodes de résolution

Il existe trois classes fondamentales de méthodes permettant de résoudre un problème d'apprentissage par renforcement : la programmation dynamique PD, les méthodes de Monte Carlo MC, et l'apprentissage par différences temporelles TD [48]. Chaque classe possède des avantages et des inconvénients. La programmation dynamique possède des fondements mathématiques bien connus/étudiés mais nécessite un modèle complet et précis de l'environnement. Les méthodes MC ne nécessitent pas de modèle et sont conceptuellement simples, mais sont inadaptées à un calcul pas à pas incrémentiel. Enfin, l'approche TD combine les deux premières méthodes, et récupère ainsi la meilleure part de chacune. Cette approche ne nécessite pas de modèle et est incrémentale. Ces méthodes se distinguent aussi en matière d'efficacité et de rapidité de convergence. Nous décrivons, dans ce qui suit, une des méthodes les plus prisée des algorithmes d'apprentissage par différences temporelles : le Q-learning.

#### L'algorithme Q-Learning :

Le Q-learning, développé en 1989 par Watkins [49][50], fait partie des méthodes de résolution sans modèle puisqu'il s'agit d'apprendre par l'expérience, les actions à effectuer en fonction de l'état actuel. Le Q-learning a été développé afin de déterminer une politique optimale  $\pi^*$  sans disposer de toutes les données pour faire les calculs par programmation dynamique ; il s'agit alors d'essayer de lier une estimation des actions et des espérances  $V(s)$ .

L'équation (3.1) peut être réécrite comme suit [50] :

$$V^{\pi}(s) = R(s, \pi(s)) + \gamma \sum_{s' \in \pi(s)} P_{ss'}(\pi(s)) V^{\pi}(s') \quad \text{où } R(s, \pi(s)) = E\{r(s, \pi(s))\}.$$

L'utilité d'une politique optimale  $\pi^*$  peut s'écrire de manière récursive sous la forme de l'équation de Bellman :

$$V^*(s) = V^{\pi^*}(s) = \max_{a \in A} \left[ R(s, a) + \gamma \sum_{s' \in S} P_{ss'}(a) V^{\pi^*}(s') \right] \quad (3.2)$$

La tâche du Q-learning consiste à déterminer cette politique  $\pi^*$  sans connaître  $R(s, a)$  et  $P_{ss'}(a)$ . Ceci est réalisé en reformulant (3.2) : À chaque politique  $\pi$  est alors associée une valeur

$Q^\pi(s, a)$ , appelée Q-valeur. Cette dernière fournit la "qualité" de l'action à effectuer dans l'état  $s$ , et elle est définie comme suit :

$$Q^\pi(s, a) = R(s, a) + \gamma \sum_{s'} P_{ss'}(a) V^\pi(s')$$

$Q^\pi(s, a)$  représente la moyenne des gains prévue si l'action  $a$  est exécutée lorsque, l'état courant du système est  $s$  et que  $\pi$  est adoptée, ensuite, comme politique de décision.

$$\text{Soit : } Q^*(s, a) = Q^{\pi^*}(s, a) = R(s, a) + \gamma \sum_{s'} P_{ss'}(a) V^{\pi^*}(s')$$

On a alors :

$$V^*(s) = \max_{a \in A} [Q^*(s, a)]$$

La valeur optimale de la fonction utilité  $V^*$  peut, ainsi, être obtenue en utilisant  $Q^*(s, a)$ .

$$V^*(s) = V^{\pi^*}(s) = \max_{a \in A} [Q^*(s, a)]$$

En revanche, nous obtenons une définition récurrente pour  $Q^*$  :

$$Q^*(s, a) = R(s, a) + \gamma \sum_{s'} \left\{ P_{ss'}(a) \left[ \max_{b \in A} Q^*(s', b) \right] \right\}$$

L'objectif de l'algorithme du Q-learning est de trouver une approximation de  $Q^*$ , en connaissant seulement le quadruplet  $(s_t, a_t, s'_t, r_t)$ . Cette information comprend l'état à l'instant  $t$  ( $s_t$ ), l'état à l'instant  $t+1$  ( $s'_t = s_{t+1}$ ), l'action prise quand le système est à l'état  $s_t$  ( $a_t$ ) et le revenu reçu à l'instant  $t$  ( $r_t$ ) suite à l'exécution de cette action. Les Q-valeurs sont mises à jours, de façon récursive à chaque transition comme suit :

$$Q_{t+1}(s, a) = \begin{cases} Q_t(s, a) + \alpha_t \Delta Q_t(s, a), & \text{si } s = s_t \text{ et } a = a_t \\ Q_t(s, a), & \text{sinon} \end{cases} \quad (3.3)$$

$$\text{où } \Delta Q_t(s, a) = \left\{ r_t + \gamma \max_b [Q_t(s'_t, b)] \right\} - Q_t(s, a) \quad (3.4)$$

D'où l'algorithme :

<b>Algorithme Q-learning</b>	<p><i>Initialiser <math>Q_0(s, a)</math> à des valeurs aléatoires</i></p> <p><i>Choisir un point de départ <math>s_0</math></i></p> <p><i>tant que la politique n'est pas suffisamment bonne</i></p> <p><i>choisir <math>a_t</math> en fonction des valeurs <math>Q_t(s_t, \cdot)</math></i></p> <p style="text-align: center;"><math>a_t = f(Q_t(s_t, \cdot))</math></p> <p><i>obtenir en retour : <math>s_{t+1}</math> et <math>r_t</math></i></p> <p><i>mettre à jour <math>Q_{t+1}(s_t, a_t)</math> en utilisant la formule (3.3)</i></p> <hr/> <p><i>Fin tant que</i></p>
----------------------------------	--

Il est alors nécessaire de régler le coefficient  $\alpha$  afin de fixer progressivement la politique apprise.  $\gamma$  permet pour sa part de moduler l'importance des récompenses escomptées à venir. Dans [49], les auteurs démontrent que si chaque paire  $(s, a)$  est infiniment visitée, et que le taux d'apprentissage  $\alpha$  tend vers zéro,  $Q_t(x, a)$  converge, lorsque  $t \rightarrow \infty$ , vers  $Q^*(x, a)$  avec une probabilité de 1. La politique optimale sera, alors, celle avec la plus grande Q-valeur :

$$\pi^*(s) = \arg \max_{a \in A(s)} Q^*(s, a).$$

### **Exploration :**

Dans l'algorithme du Q-learning, le choix de l'action (la fonction  $f$ ) n'est pas décrit. Il est possible d'imaginer différentes stratégies de sélection, par exemple l'action qui a été la moins utilisée, ou celle qui renvoie la plus forte Q-valeur. Dans [51] l'auteur identifie deux approches dans l'exploration : (i) une exploration *dirigée*, reposant sur les actions ou sur les Q-valeurs correspondants, et (ii) une exploration *non-dirigée* dite aussi 'aléatoire'. Le choix de l'action, pour la première approche, peut se faire par exemple soit sur la base du nombre d'utilisation de l'action, soit parce que l'action a eu, par le passé, un fort pouvoir prédictif. Quant au choix de l'action pour la seconde approche, il se fait à l'aide d'un tirage aléatoire. L'auteur dans [51] étudie l'impact de l'exploration sur le contrôle de l'apprentissage, et compare ces deux approches

en terme de temps moyen nécessaire pour la convergence. Les méthodes non dirigées ont besoin de beaucoup plus de temps (entre 5 et 10 fois plus). Cependant, le processus mis en œuvre dans cette approche est sans mémoire, puisqu'une fois l'action choisie, l'agent n'a plus accès aux informations relatives à son point de départ. L'intérêt d'une telle approche est de ne pas nécessiter une mémoire dépendante de la taille de l'environnement.

Dans nos travaux, nous avons testé un ensemble de méthodes d'exploration appartenant aux deux approches, telles que  $\epsilon$ -gloutonne (ou  $\epsilon$ -directed, Boltzmann et même des stratégies aléatoires que nous détaillerons dans les chapitres suivants.

### 3. Applications

Les travaux utilisant les techniques d'apprentissage, comme solution, pour résoudre les problèmes rencontrés dans les réseaux de télécommunications ne sont pas aussi nombreux. Nous allons énumérer, ci-dessous, quelques uns de ces travaux allant d'applications classiques (robotique, jeux, etc.) à des applications réseaux. (routage, CAC, etc.).

#### 3.1 Applications classiques

De nombreux travaux ont été effectués dans le cadre de l'apprentissage de jeu à deux joueurs (jeu de dames, jeu du Backgammon, etc.). L'autre grand domaine d'applications de l'apprentissage par renforcement est celui de la robotique "autonome", un domaine très peu abordé avec l'apprentissage supervisé du fait de l'impossibilité de modéliser le monde réel (c.a.d. de connaître la sortie désirée) avec suffisamment de précision pour tenir compte de l'hétérogénéité des capteurs, du bruit ambiant et la dynamique robot-monde extérieur. *Schaal et Atkeson* (1994) ont construit un robot à deux bras, devant jongler avec un stick. *Mahadevan et Connell* (1992) ont utilisé cette approche pour apprendre à un robot mobile à déplacer des objets par poussées. Le Q-Learning a été utilisé par *Crites et Barto* (1996) pour gérer 10 ascenseurs d'un bâtiment avec pour objectif la minimisation du temps d'attente par les utilisateurs.

Dans [52], les auteurs proposent un système multi-agents (SMA) capable de prendre des décisions économiques telles que fixer les prix dans un contexte de marché compétitif. Ils utilisent un modèle où chacun des vendeurs fixe son prix à tour de rôle et reçoit aussitôt un certain revenu. Ils considèrent deux cas : (i) un seul agent a la capacité d'apprendre ; (ii) tous les agents essaient d'apprendre en même temps une politique optimale. Ils démontrent, qu'avec le

Q-learning l'agent permet de trouver la politique de prix optimale pour le premier cas, et de diminuer le phénomène de 'guerre des prix' entre les différents vendeurs pour le deuxième cas.

### 3.2 Routage

Les travaux [53][54][55] sont des propositions de routage dans les réseaux de télécommunications, en utilisant des techniques d'apprentissage comme solution. Nous nous intéressons, plus particulièrement, au travail proposé par *Boyan et Littman* et qu'ils ont nommé Q-routing [53]. Ce travail fait l'objet d'une extension traitée dans le Chapitre 8 de ce document. Les auteurs proposent une extension de l'algorithme à vecteur de distance de *Bellman-Ford* (DBF) basé sur l'apprentissage par renforcement. Le module d'apprentissage par renforcement est intégré dans chaque nœud d'un réseau de commutation. La politique de routage tente de trouver le meilleur nœud adjacent permettant d'atteindre la destination avec un "temps de transmission" minimum. Chaque nœud,  $x$ , estime le temps de transmission pour envoyer un paquet vers une destination  $d$  par l'intermédiaire de l'un de ses voisins,  $y$ . Cette estimation est améliorée après une rétroaction directe du temps de transmission mesuré quand le paquet atteint sa destination. En fait, en utilisant l'apprentissage par renforcement, cette estimation peut être mise à jour localement avant que le paquet n'atteigne sa destination. Effectivement, et comme le décrit [53], lors de l'envoi d'un paquet à  $y$ , le nœud  $x$  récupère immédiatement l'estimation du temps de transmission de  $y$  pour la partie restante de la route.

Dans l'algorithme Q-routing, chaque Q-valeur  $Q_x(y,d)$  représente une estimation, pour chaque nœud  $x$ , du temps de transmission de bout en bout au cas où un paquet est routé vers le voisin  $y$  pour arriver à destination. Ce temps de transmission inclut, également, le temps de mise en file d'attente dans  $x$ .

Quand un nœud  $x$  doit prendre une décision de routage il choisit simplement le voisin  $y$  pour lequel  $Q_x(y,d)$  est minimum. À l'envoi d'un paquet au nœud  $y$ ,  $x$  récupère immédiatement l'estimation de  $y$  concernant le temps de transmission sur la route restante, à savoir :

$$t = \min_{z \in N(y)} Q_y(z, d)$$

où  $N(w)$  représente l'ensemble des voisins du nœud  $w$ .

Si le paquet met  $q$  unités de temps dans la file du nœud  $x$  et  $s$  unités de temps pour arriver au nœud  $y$ , alors le nœud  $x$  peut mettre à jour son estimation en se basant sur cette rétroaction comme suit :

$$\Delta Q_x(y, d) = \eta ( \overbrace{q + s + t}^{\text{nouvelle estimation}} - \overbrace{Q_x(y, d)}^{\text{ancienne estimation}} )$$

où  $\eta$  représente le taux d'apprentissage.

En d'autres termes, les informations sur la route restante sont employées pour mettre à jour la Q-valeur du nœud émetteur. Les auteurs démontrent par un ensemble de simulations que le Q-routing est plus performant qu'un simple routage basé sur le plus court chemin (*Bellman-Ford*).

### 3.3 Allocation de ressources

Une proposition particulièrement intéressante traitant de l'allocation de ressources, est celle proposée par *Nie* et *Haykin* [56]. Nous nous intéressons plus spécialement à cette proposition puisqu'elle fait l'objet d'une extension abordée dans le Chapitre 5. Les auteurs proposent de résoudre le problème d'allocation dynamique de ressources dans un réseau cellulaire en utilisant le Q-learning. Ils considèrent un système à allocation dynamique de ressources (DCA – Dynamic Channel Allocation) supportant une seule classe de trafic. Dans un système DCA, et comme on l'a vu dans le chapitre précédent, chaque canal est maintenu dans un pool commun, et il est temporairement alloué à une des cellule à condition que la contrainte sur la distance de réutilisation soit satisfaite. Le réseau doit assigner les canaux disponibles aux nouveaux appels dans le but de maximiser l'ensemble des gains à long terme. Les auteurs donnent une formulation sous forme d'un programme dynamique dont l'ensemble des états, des actions et des revenus sont :

- l'état du système  $s_t = (i, A(i))_t$ , où  $i$  est l'index de la cellule où se produit un événement (arrivée/départ d'un appel).  $A(i)$  représente le nombre de canaux disponibles dans la cellule  $i$  ;
- l'action  $a$  définit le fait d'allouer un des canaux disponibles dans  $A(i)$  au nouvel appel ;

- le revenu  $r(s, a)$  représente le gain instantané obtenu après l'exécution d'une action  $a$  lorsque le système est dans un état  $s$ . Ce revenu prend en compte les conditions d'utilisation dans les cellules co-canal<sup>3</sup> liées à la cellule  $i$  :

$$r(s, a) = n_1(a)r_1 + n_2(a)r_2 + n_3(a)r_3 \quad (3.5)$$

Dans la formule (3.5),  $n_1(k)$  représente le nombre de cellules compactes liées à la cellule  $i$  dans lesquelles le canal  $k$  est utilisé. Les cellules compactes sont les cellules avec une distance moyenne minimum entre les cellules co-canaux [17]. Dans un modèle hexagonal (comme celui présenté dans la Figure 4.19), les cellules compactes sont situées sur la troisième rangée avec trois cellules les séparant;  $n_2(k)$  est le nombre de cellules co-canaux qui sont situées sur la troisième rangée et qui ne sont pas compactes dans lesquelles le canal  $k$  est utilisé;  $n_3(k)$  est le nombre des autres cellules co-canaux utilisant le canal  $k$ .  $r_1$ ,  $r_2$  et  $r_3$  sont des constantes associées respectivement à  $n_1(k)$ ,  $n_2(k)$  et  $n_3(k)$ .

L'agent détermine la politique permettant d'allouer le meilleur canal qui maximise le cumul des gains reçus et ceci en connaissant uniquement l'état actuel du réseau. Ils démontrent par un ensemble de simulations que la politique d'allocation basée sur le Q-learning (DCA-Nie) est plus performante que d'autres heuristiques en terme de probabilité de blocage des appels.

### 3.4 Contrôle d'admission CAC

Les problèmes de contrôle d'admission CAC dans les réseaux fixes et utilisant l'apprentissage par renforcement ont été traités dans quelques travaux [54][55][57][58][59]. Les auteurs de [54] proposent, par exemple, de résoudre ce problème dans un réseau à intégration de services tel que ATM. Les auteurs se mettent dans le contexte où un fournisseur de services désire vendre ses ressources réseaux dans le but de maximiser l'ensemble de ses revenus. Ils considèrent différentes classes de services et le réseau doit accepter ou rejeter les nouveaux appels (Call Admission Control). Si un appel est accepté, le réseau doit trouver la meilleure route au sein du réseau (Routage). Ces problèmes de CAC et de routage sont formulés sous forme de problèmes de programmation dynamique, et l'apprentissage par renforcement est utilisé pour les résoudre. Ils démontrent par un ensemble de simulations que les politiques de contrôle d'admission et de routage obtenues sont plus performantes que OSPF (Open-Shortest-Path-First).

---

3. Cellules utilisant le même canal sans causer d'interférences.

## 4. Conclusion

Nous venons de voir que l'apprentissage numérique permet de résoudre des problèmes numériques assez complexes. L'objectif est toujours le même, et il consiste à établir des relations numériques entre les entrées (données des exemples à apprendre) et les sorties (comportement observé des exemples). Nous avons évoqué différentes approches de l'apprentissage numérique et nous nous sommes, particulièrement, intéressé à l'approche "apprentissage par renforcement". Ce type d'apprentissage est une alternative possible à la définition de la base d'apprentissage. La forme des exemples d'apprentissage est du type (entrée, sortie, utilité), où la dernière composante encode l'utilité de produire cette "sortie" pour cette "entrée". Les exemples d'apprentissage sont générés automatiquement au cours d'une phase dite d'exploration. Le succès de l'application dépendra de la qualité de la fonction spécifiant l'utilité d'une paire. Echanger la construction "manuelle" de la base d'apprentissage contre une simple mesure de la performance est certainement économique pour l'opérateur. Cependant, le seul inconvénient est une convergence de l'apprentissage un peu plus lente. Il est important de noter qu'en réalité l'apprentissage supervisé et l'apprentissage par renforcement ne se concurrencent pas. Chacun correspond à une niche d'applications particulière, selon que l'on dispose d'une base d'exemples représentative ou d'une mesure de la performance du comportement recherché.

Pour résoudre les quelques problèmes liés aux réseaux mobiles étudiés dans cette thèse, l'apprentissage par renforcement nous a paru la seule solution possible vu que l'on ne dispose d'aucune base d'apprentissage représentative au préalable.

# Chapitre 4

## Contrôle adaptatif d'admission des appels

### 1. Motivations

Nous avons vu dans le Chapitre 2 que la zone géographique desservie par un réseau mobile est divisée en cellules se partageant l'ensemble de la bande de fréquences suivant deux méthodes : FCA et DCA. Dans cette partie, nous nous intéressons uniquement à FCA, où à chacune des cellules est allouée un ensemble fixe de canaux. Un canal est alloué à chaque usager. Quand un usager passe d'une cellule à une autre (handoff - transfert inter-cellulaire), il doit demander un nouveau canal dans la cellule destination. Si la cellule destination n'a aucun canal disponible, l'appel est coupé. Pour un client, cette coupure est beaucoup plus désagréable qu'un échec de connexion. L'un des soucis majeurs lors de la conception des réseaux mobiles est la réduction de la probabilité de coupure. Ceci est d'autant plus important, qu'afin de répondre à la croissance des réseaux mobiles, la taille des cellules est de plus en plus réduite, ce qui augmente considérablement le nombre de transferts inter-cellulaires.

Nous avons également, vu que les techniques de réservation de canaux (ou Guard Channel) permettent de réduire la probabilité de coupure de communication en réservant, dans chaque cellule, des canaux à l'usage exclusif des handoffs. Si ces techniques sont simples à dimensionner lorsqu'on considère une seule classe de trafic (i.e. appel téléphonique), elles deviennent beaucoup plus compliquées à optimiser, et moins optimales dans un contexte multi-classes de trafic.

En effet, dans un contexte multi-classes de trafic, il est parfois préférable de bloquer un appel d'une classe moins prioritaire, et d'accepter un autre appel appartenant à une classe de priorité supérieure. Le contrôle d'admission d'appels proposé dans ce chapitre, permet ce genre de mécanismes. Il est aussi important d'exploiter l'expérience et la connaissance qui pourraient être acquises en cours de fonctionnement du système.

Nous proposons une nouvelle approche pour résoudre le problème du contrôle d'admission des appels dans les réseaux cellulaires supportant plusieurs classes de trafic. Cette politique (ou stratégie) optimale de CAC est obtenue en utilisant l'algorithme d'apprentissage par renforcement : Q-learning [49] présenté dans le Chapitre 3. Le double objectif de ce type d'apprentissage est de conduire de manière optimale un système au cours du temps, et d'apprendre cette conduite optimale au travers d'un ensemble d'expériences.

Un revenu, représentant le coût du service fourni aux utilisateurs, est associé à chaque classe de trafic. Ces revenus dépendent également du fait que c'est un nouvel appel ou bien un appel de handoff. Les handoffs sont plus prioritaires et donc les revenus associés sont plus grands. L'objectif est donc, d'accepter des clients et d'en rejeter d'autres dans le but de maximiser le cumul des revenus reçus au cours du temps. Ce problème peut être formulé comme un processus de décision semi-markovien (SMDP) dont l'apprentissage est une solution idéale.

Nous proposons d'étudier, en premier lieu, un cas simple où on considère une seule cellule de type FCA modélisée de façon adéquate. Nous généralisons ensuite, pour un système cellulaire complet de type FCA, avec plusieurs cellules et plusieurs canaux distribués suivant un motif cellulaire.

## 2. Cas d'une seule cellule FCA

Après une formulation du problème du contrôle d'admission d'appels dans une simple cellule de type FCA (cf. Figure 4.1) en tant que SMDP, nous détaillerons deux implantations de l'algorithme du Q-Learning (TRL-CAC et NRL-CAC) capables de résoudre ce SMDP ainsi que l'évaluation des performances et les résultats de simulation.

## 2.1 Formulation du problème sous forme de SMDP

Concentrons-nous sur une simple cellule FCA avec  $N$  canaux disponibles, et deux<sup>4</sup> classes de trafic  $C_1$  et  $C_2$ . Les appels de la première classe demandent un seul canal, alors que les appels de la seconde demandent deux canaux. De nouveaux appels, ainsi que des handoffs venant des cellules voisines, peuvent surgir dans la cellule (cf. Figure 4.1).

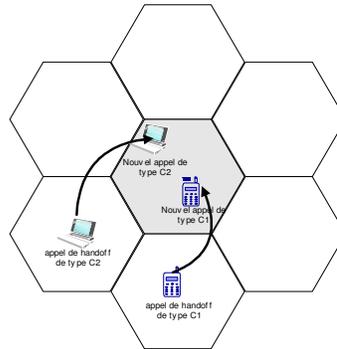


Figure 4.1 Modèle d'une cellule FCA avec deux classes de trafic.

Ce système cellulaire peut être considéré comme un système à événement discret. Les principaux événements pouvant se produire dans une cellule sont les arrivées et les départs d'appels. Ces événements sont modélisés par des variables stochastiques avec des distributions appropriées. En particulier, les arrivées des nouveaux appels et des handoffs sont Poissonniennes. Le temps de séjour de chaque appel est exponentiellement distribué (les paramètres expérimentaux sont récapitulés dans le Tableau 4.3). Les appels arrivent dans la cellule puis partent, soit en effectuant un handoff vers une autre cellule, soit en se terminant normalement. Le réseau devra alors, choisir d'accepter ou de rejeter ces demandes de connexion. En retour, il collecte l'ensemble des revenus reçus de la part des clients acceptés (les gains), ainsi que l'ensemble des revenus reçus de la part des clients rejetés (les pertes). L'objectif de l'opérateur du réseau est de trouver une politique de CAC qui maximise l'ensemble des gains à long terme et qui réduise les probabilités de blocage des handoffs ainsi que l'ensemble des pertes.

L'ensemble des états " $s$ ", des actions " $a$ " et des revenus " $r$ " ont été définis comme suit :

4. L'idée peut facilement être étendue à plusieurs classes de trafic.

**Etats :** À l'instant  $t$ , le système est dans une configuration particulière  $x$  définie par le nombre d'appels en cours pour chaque type de trafic. À un instant aléatoire, un événement  $e$  peut se produire indiquant ou bien l'arrivée d'un appel (nouveau ou handoff) ou bien le départ d'un appel. L'événement de départ est dû à une terminaison normale de l'appel, ou à un transfert inter-cellulaire de l'appel vers une cellule voisine. La configuration  $x$  et l'événement  $e$  déterminent ensemble l'état du système  $s=(x, e)$  :

–  $x=(x_1, x_2)$ , où  $x_1$  et  $x_2$  représentent le nombre d'appels de chacune des classes de trafic ( $C1$  et  $C2$  respectivement) dans la cellule.

–  $e=\{1,2,3,4\}$  où

$$e = \begin{cases} 1 & \text{arrivée d'un nouvel appel de type C1} \\ 2 & \text{arrivée d'un nouvel appel de type C2} \\ 3 & \text{arrivée d'un appel de handoff de type C1} \\ 4 & \text{arrivée d'un appel de handoff de type C2} \end{cases} \quad (4.1)$$

Les états associés au départ d'un appel ne sont pas pris en compte. La raison de cette simplification est que le départ d'un appel n'est pas une décision en soit dans un contrôle d'admission, et donc aucune mesure ne doit être prise.

**Actions :** Appliquer une action consiste à accepter ou à rejeter un appel. Ainsi, l'ensemble des actions possibles est  $A=\{1, 0\}$  où :

$$a_i = \begin{cases} 1 & \text{accepter} \\ 0 & \text{rejeter} \end{cases}$$

**Revenus :** Le revenu  $r(s,a)$  représente le gain instantané, obtenu après l'exécution d'une action  $a$  lorsque le système est dans un état  $s$ . Quand un appel de classe  $i$  arrive ( $e = e_i \in \{1,2,3,4\}$ ), le revenu associé sera nul si l'appel est rejeté ( $a=0$ ), et égal à  $\eta_i$  si l'appel est accepté ( $a=1$ ) :

$$r(s,a) = \begin{cases} \eta_i & \text{si } a=1 \text{ et } e = e_i \\ 0 & \text{sinon} \end{cases}$$

Les valeurs de  $\eta_i$  sont données par le Tableau 4.1. Étant donné que l'objectif principal de l'opérateur du réseau est de diminuer les probabilités de blocage des handoffs, des valeurs de

revenus assez grandes leurs ont été affectées. Les valeurs de revenus pour les appels de la classe C1 sont plus importantes que celles des appels de la classe C2, étant donné que C1 est supposé prioritaire par rapport à C2.

$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$
5	1	50	10

Tableau 4.1 Revenus Instantanés.

En résumé, nous avons choisi la description de l'ensemble des états comme étant  $s = ((x_1, x_2), e)$ , où  $x_i$  est le nombre d'appels en cours de la classe  $C_i$ , et  $e \in \{1,2,3,4\}$  représente l'arrivée d'un nouvel appel ou d'un appel de handoff dans la cellule. Quand un événement se produit, l'agent doit choisir une des actions possibles  $A(s) = \{0=rejeter, 1=accepter\}$ . À la fin d'un appel, aucune mesure ne doit être prise. L'agent doit donc déterminer une politique d'acceptation des appels permettant de maximiser le cumul des gains reçus et ceci en connaissant uniquement l'état du réseau  $s$ . Ce système constitue un SMDP avec comme espace d'états un ensemble fini  $S = \{(x, e)\}$  et comme espace d'actions possibles l'ensemble fini  $A = \{0,1\}$ .

## 2.2 Implantation des algorithmes

Après avoir formulé le problème CAC sous forme d'un SMDP, nous allons décrire deux implantations de l'algorithme Q-learning capables de résoudre ce problème. Nous les avons nommées TRL-CAC (Table Reinforcement Learning CAC) et NRL-CAC (Neural network Reinforcement Learning CAC). TRL-CAC utilise une simple table pour représenter la fonction  $Q$  (l'ensemble des Q-valeurs). En revanche, NRL-CAC utilise un réseau de neurones multicouches [45].

<b>Représentation des Q-valeurs dans le Q-learning</b>	La majorité des preuves de convergence du Q-learning ont été faites, en supposant le cas où l'espace des états est assez petit, et donc une simple table serait utilisée pour représenter la Q-fonction. En outre, la principale difficulté avec les problèmes SMDP provient du passage à l'échelle ou scalabilité (explosion exponentielle de l'espace des états avec la dimension du problème). Ainsi, lorsque l'ensemble des paires état-action (s,a) est grand ou lorsque les variables d'entrée (x,e) constituant l'état s sont des variables continues, l'utilisation d'une simple table devient rédhibitoire à cause des immenses besoins de stockage. Dans ce cas, quelques fonctions d'approximation, telles que l'agrégation des états, les réseaux de neurones [45][61] ou même les arbres de régression [60] peuvent être utilisés de manière efficace.
--	---

Les différences entre ces deux algorithmes (TRL-CAC et NRL-CAC) sont exprimées en terme de complexité de calcul et de taille mémoire. Ces différences sont résumées dans le Tableau 4.2.

	Nombre d'opérations de calcul nécessaires	Unités mémoire nécessaires
Table	0	$676 \times 5 = 3380$
Réseau de neurones	$4 \times 10 + 10 \times 4 + 10 + 10 = 100$	$4 \times 10 + 10 = 50$

Tableau 4.2 Table vs. Réseaux de neurones.

### 2.2.1 Représentation des Q-valeurs sous forme d'une Table

L'approche utilisant une table est la méthode la plus simple et la plus efficace. Comme le montre la Figure 4.2 (a), les premières colonnes (de 1 à 3) représentent l'état  $s = ((x_1, x_2), e)$ . Les colonnes 4 et 5 représentent, respectivement, la valeur d'acceptation et la valeur de rejet. Cette approche conduit à un calcul exact et elle est complètement conforme aux hypothèses de structures faites afin de prouver la convergence de l'algorithme du Q-learning [49].

S'il y a seulement deux classes de trafic, une seule opération de comparaison entre les deux Q-valeurs d'acceptation et de rejet est nécessaire (cf. Tableau 4.2). En revanche, pour obtenir ces deux valeurs, l'indexation associée peut être fastidieuse. En effet, la taille de la table utilisée pour stocker les Q-valeurs est très grande ( $676 \times 5 = 3380$  unités mémoires sont nécessaires dans notre cas)<sup>5</sup>.

### 2.2.2 Représentation des Q-valeurs sous forme d'un réseau de neurones

Les réseaux de neurones sont des modèles mathématiques et informatiques, dont l'inspiration originelle était le système nerveux humain. Les réseaux de neurones possèdent plusieurs propriétés qui les rendent intéressants d'un point de vue théorique, et fort utiles en pratique.

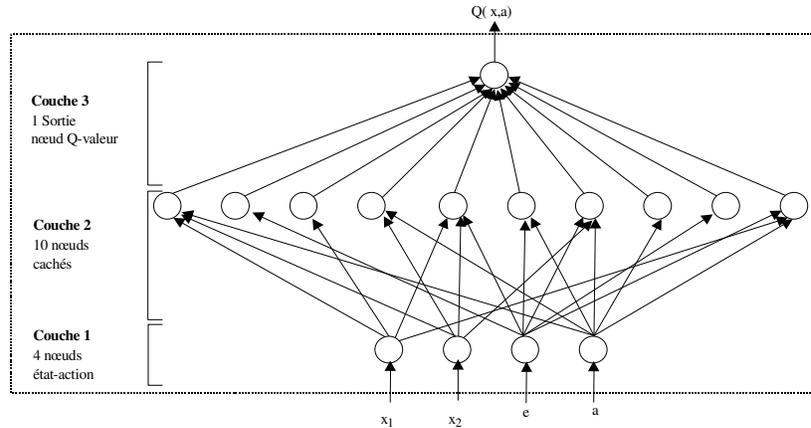
5. La taille réelle de l'espace des états a été réduite de 2304 ( $12 \times 24 \times 4 \times 2$ ) états à 676 états, en éliminant les combinaisons impossibles.

<p><b>Les réseaux de neurones ?</b></p>	<p>Formellement, un réseau de neurones [61] est un graphe dont les nœuds sont des unités de calcul appelés <i>neurones formels</i>, et les arêtes orientées et pondérées, s'appellent <i>liens synaptiques</i>. Chaque neurone formel est un automate indépendant dont l'état, une valeur scalaire, se nomme activation. Une fonction d'activation anime le neurone en déterminant son activation, influencée par les neurones en amont (Ces neurones influencent en fonction de leur propre activation, et de la pondération du lien appelé <i>poids synaptique</i> les reliant au neurone courant). Une fois déterminée, l'activation du neurone courant est propagée à son tour le long de liens synaptiques vers les neurones en aval, et de proche en proche, à travers le réseau. Cette propagation d'activation constitue le mode de calcul principal du réseau de neurones.</p> <p>Malgré la simplicité des neurones individuels, le comportement collectif (ou émergent) du réseau de neurones peut être fort complexe, et les calculs qu'il effectue, très puissants. Parce qu'il est collectif, le comportement du réseau est en général réglé par sa connectivité, et plus particulièrement, par les poids synaptiques des liens. Ainsi, un même réseau est capable de calculs très différents, selon les valeurs exactes de ses poids synaptiques. Les architectures en réseaux de neurones apprennent, à partir d'exemples, à acquérir une certaine carte d'entrée/sortie par l'adaptation de ses poids synaptiques.</p>
---	--

Le réseau de neurones que nous choisissons pour représenter les Q-valeurs, est un réseau multi-couche classique (Figure 4.2 (b)). Il est constitué de 4 entrées, 10 unités cachées et une unité de sortie. La complexité de calcul dépend de la taille du réseau de neurones. Dans notre cas, 100 opérations (multiplications et additions) sont nécessaires pour calculer les deux Q-valeurs d'acceptation et de rejet (cf. Tableau 4.2). En revanche, en terme de besoins de stockage, l'approche nécessite moins de mémoire de stockage (50 unités seulement) pour sauvegarder les poids synaptiques du réseau de neurones.

	1	2	3	4	5
	$x_1$	$x_2$	$e$	$Q(x,a=1)$	$Q(x,a=0)$
1					
:					
:					
676					

(a)



(b)

Figure 4.2 Représentation des Q-valeurs (a) Table, (b) Réseau de Neurones.

### 2.2.3 Implantation

L'organigramme de la Figure 4.4 récapitule les différentes phases des algorithmes TRL-CAC et NRL-CAC. Nous avons initialisé les Q-valeurs à des valeurs aléatoires. Nous rappelons que les états les plus importants sont ceux associés aux arrivées des appels et, par conséquent, les états associés aux départs des appels ne sont pas pris en compte. Ceci réduit considérablement les capacités de calcul et de stockage des Q-valeurs.

Lorsqu'un appel arrive (nouvel appel ou handoff), l'algorithme détermine si la qualité de service n'est pas violée en acceptant cet appel (en vérifiant tout simplement s'il y a assez de canaux disponibles dans la cellule). Si cette qualité de service est violée, l'appel est rejeté ; sinon l'action est choisie selon la formule :

$$a = \arg \max_{a \in A(s)} Q^*(s, a) \quad (4.2)$$

où  $A(s) = \{1 = \text{accepter}, 0 = \text{rejeter}\}$ .

En particulier, (4.2) implique les procédures suivantes : lorsqu'un appel arrive, la Q-valeur d'acceptation ainsi que la Q-valeur de rejet de l'appel sont déterminées soit à partir de la table (TRL-CAC), soit à partir du réseau de neurones (NRL-CAC). Si le rejet a une plus grande valeur, l'appel est alors est rejeté. Dans le cas contraire, l'appel est accepté.

Dans ces deux cas, et pour apprendre les Q-valeurs optimales  $Q^*(s,a)$ , la fonction est mise à jour à chaque transition du système d'un état  $s$  vers un état  $s'$ . Pour les deux algorithmes, ceci se fait de la façon suivante :

- TRL-CAC: C'est la formule (3.3) qui est utilisée pour mettre à jour la Q-valeur appropriée dans la table ;
- NRL-CAC: Quand un réseau de neurones est utilisé pour stocker la fonction  $Q$ , une seconde procédure d'apprentissage est nécessaire pour apprendre les poids du réseau de neurones. Cette procédure utilise l'algorithme de *rétro-propagation* (BP – Back Propagation) [45]. Dans ce cas,  $\Delta Q$  définie par la formule (3.4) est employée comme signal d'erreur qui est rétro-propagé dans les différentes couches du réseau de neurones (cf. Figure 4.3).

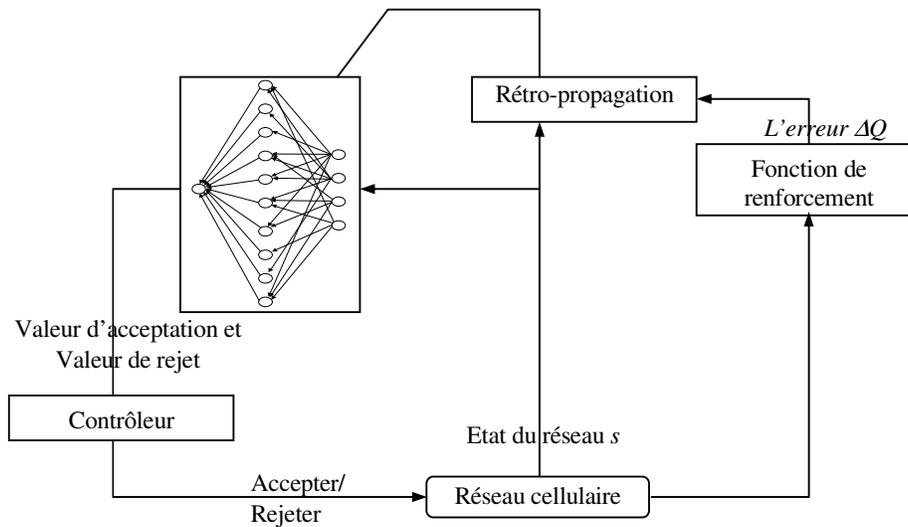


Figure 4.3 La rétro-propagation de l'erreur dans l'algorithme NRL-CAC.

Nous comparons nos politiques à celle que nous avons appelée politique *gloutonne* (ou *greedy*) [62] (politique qui accepte un nouvel appel ou un appel de handoff si la contrainte de capacité n'est pas violée en acceptant cet appel).

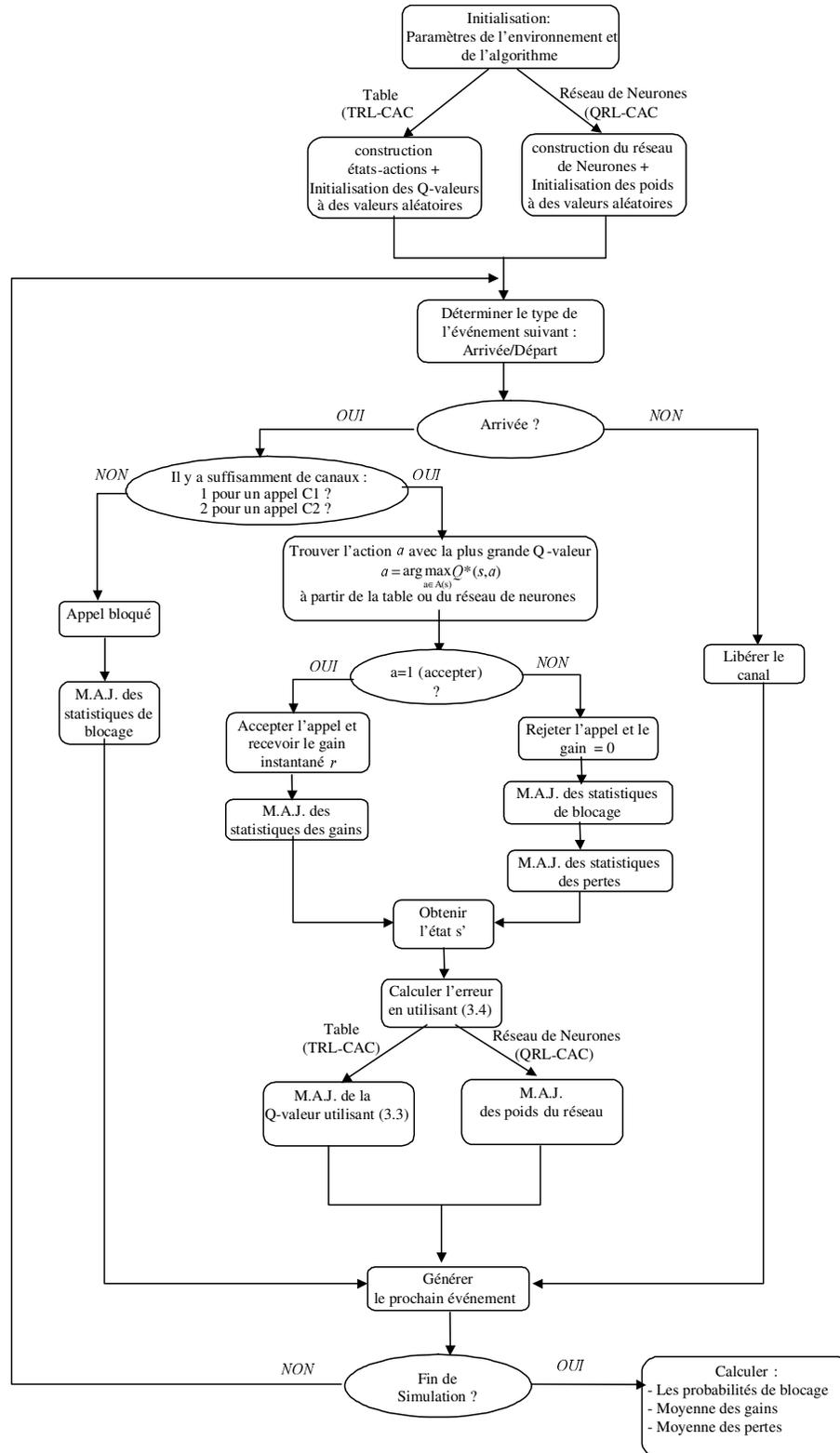


Figure 4.4 Les algorithmes TRL-CAC et NRL-CAC.

### 2.2.4 Exploration

Nous avons vu dans le Chapitre 3 que pour une exécution correcte et efficace de l'algorithme du Q-learning, toutes les paires potentiellement importantes des état-action  $(s,a)$  doivent être explorées. Pour cela, pendant une période d'apprentissage assez longue, l'action est choisie non selon la formule (4.2), mais selon la formule (4.3) suivante avec une probabilité d'exploration  $\varepsilon$  :

$$a = \arg \min_{a \in A(s)} \text{visites}(s,a) \tag{4.3}$$

où  $\text{visites}(s,a)$  désigne le nombre de fois où la configuration  $(s,a)$  a été visitée. Cette heuristique, appelée  $\varepsilon$ -gloutonne (ou  $\varepsilon$ -directed), accélère significativement la convergence des Q-valeurs. Ces valeurs sont, dans un premier temps, calculées en utilisant cette heuristique avec les paramètres indiqués dans le Tableau 4.3. Dans un second temps, ces valeurs seront employées pour initialiser les Q-valeurs dans nos deux algorithmes de CAC.

## 2.3 Étude analytique du mécanisme de réservation de canaux

Dans cette section, nous comparons nos solutions au mécanisme de réservation des canaux (ou "*guard channel*") présenté dans le Chapitre 2. La technique de réservation de canaux permet de partager la capacité de la cellule entre les nouveaux appels et les handoffs, en donnant une certaine priorité à ces derniers. Ceci est réalisé en réservant, dans chaque cellule, des canaux à l'usage exclusif des handoffs (canaux de garde).

Le nombre de canaux réservés pour les handoffs est donc un paramètre de grande importance dans ce genre de mécanisme. Dans un contexte multi-service, la question qui se pose est la suivante : combien de canaux peuvent être réservés aux handoffs pour chacune des classes de trafic, afin de maximiser l'ensemble des revenus ?

Pour répondre à cette question, nous proposons de développer un modèle mathématique classique : une cellule est représentée par une file multi-serveur classique de type M/M/N/N où chacun des serveurs représente un canal de communication (cf. Figure 4.5). Les arrivées des nouveaux appels et des handoffs dans la cellule suivent une distribution de Poisson avec un taux  $\lambda_j$ ,  $1 \leq j \leq K$  ( $K$  étant le nombre de types de trafic considérés). Le temps de séjour d'un appel de classe  $j$  dans la cellule est exponentiel de paramètre  $\mu_j$ . Nous supposons que chaque appel de type  $i$  génère un revenu  $\eta_i$ . Comme précédemment, nous supposons que  $\eta_1 \leq \eta_2 \leq \dots \leq \eta_K$

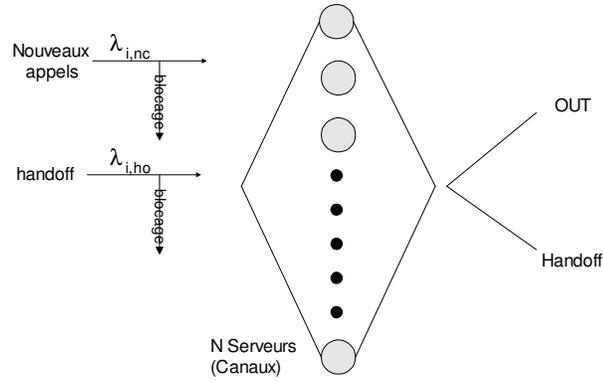


Figure 4.5 Modèle d'une cellule FCA.

La politique d'accès à la file est de type "*trunk reservation*", dont le mécanisme de réservation est caractérisé par un vecteur  $s = (s_1, s_2, \dots, s_K)^6$  correspondant aux différents seuils. Par conséquent, pour un vecteur de seuils  $s$  donné, l'ensemble des états possibles  $\Omega_s$  du système peut être décrit comme suit :

$$\Omega_s = \left\{ n = (n_1, n_2, \dots, n_K) \mid \forall 1 \leq j \leq K, \quad n \times b = \sum_{i=1}^K n_i \cdot b_i \leq s_j \right\}$$

où,  $b = (b_1, b_2, \dots, b_K)$ , et  $b_i$  représente le nombre de canaux demandés par les appels de type  $i$ .

Étant donné les caractéristiques de notre modèle, on sait que  $s_1 \leq s_2 \leq \dots \leq s_K = N$ . Le seuil  $s_i$  dépend de la classe de trafic  $i$  et de sa nature (nouvel appel/handoff). Par exemple, dans le cas de 2 classes de trafic seulement  $C_1$  et  $C_2$ , nous aurons  $K = 4$  types de trafic (nouvel appel/handoff pour chacune des classes de trafic  $C_1$  et  $C_2$ ).

Notre objectif est de déterminer le vecteur  $s^*$  des seuils optimaux permettant de maximiser le revenu moyen du système. Pour cela, notons que la moyenne temporelle des revenus reçus à chaque acceptation d'un appel peut être déterminée comme suit :

$$F(T) = \frac{1}{T} \sum_{i=1}^K \eta_i A_i(T)$$

où,  $A_i(t)$  est le nombre d'appels de type  $i$  acceptés entre 0 et  $T$ .

6.  $s_j$  représente le nombre de canaux réservés pour les types de trafic 1, 2, ...,  $j$ .

Le système étant ergodique (simple file multi-classes de type M/M/N/N), nous obtenons alors :

$$\lim_{T \rightarrow \infty} \frac{A_i(T)}{T} = \Lambda_i^s = \lambda_i (1 - P_{b,i}^s)$$

où  $\Lambda_i^s$  est le débit écoulé du trafic de type  $i$  à l'état stationnaire,  $\lambda_i$  le débit offert et  $P_{b,i}^s$  la probabilité de blocage du trafic de type  $i$  pour un vecteur  $s$  donné. De ce fait, l'ensemble des revenus reçus par un tel système sera égal à :

$$F^s = \lim_{T \rightarrow \infty} F(T) = \sum_{i=1}^K \eta_i \Lambda_i^s = \sum_{i=1}^K \eta_i \lambda_i (1 - P_{b,i}^s) \quad (4.4)$$

Or si on note  $\pi_n$ , la probabilité stationnaire d'être dans l'état  $n$ , alors :

$$P_{b,i}^s = \sum_{n \in \Omega_i} \pi_n \cdot 1_{\{n \times b + b_i > s_i\}} \quad (4.5)$$

Il convient donc de déterminer la distribution limite à l'état stationnaire  $\pi_n$ . A cette fin, notons que sous les conditions de trafic considérées, nous pouvons facilement démontrer que le processus  $\{N(t) = (N_1(t), N_2(t), \dots, N_K(t)), t \in \mathfrak{R}\}$ , où  $N_i(t)$  représente le nombre d'appels de type  $i$  à l'instant  $t$ , constitue un processus markovien à temps continu. Le nombre d'états  $\Omega_s$  étant fini, le graphe associé étant fortement connexe, le processus est ergodique. En revanche, contrairement à d'autres cas, la mise en place des seuils de réservation fait perdre la solution à forme produit et ne permet pas de relaxer la contrainte de durée d'appels exponentielle (par rapport au cas de la file M/G/C/C multiclasse, la région d'admission des appels n'est pas convexe [63]). Dans le cas où l'on considère un plus grand nombre de classes de trafic, ce problème pourrait être résolu par une solution approximative à forme produit [63]. Cependant, dans notre cas (4 types de trafic seulement), une solution numérique exacte peut être trouvée.

Le vecteur optimal  $s^*$  est déterminée selon la formule ci-après :

$$s^* = \arg \max_s \{F^s\} \quad (4.6)$$

où,  $F^s$  est calculée en utilisant l'équation (4.4). Les différentes probabilités de blocage pour chaque vecteur  $s$  candidat peuvent être calculées en utilisant la formule (4.5). Toutes les

configurations  $s$  possibles ont été examinées. Le vecteur optimale  $s^*$  est ainsi calculée hors-ligne, pour chaque configuration de trafic.

## 2.4 Résultats expérimentaux

Afin d'évaluer les avantages de nos algorithmes, nous avons utilisé une simulation à événements discrets pour représenter notre réseau cellulaire. La cellule FCA possède  $N = 24$  canaux. Les performances des algorithmes ont été évaluées en terme de : (i) revenus reçus de la part des clients acceptés (les gains), (ii) de revenus reçus de la part des clients rejetés (les pertes), et (iii) de probabilités de blocage des handoffs. La constante de propagation temporelle a été fixée à  $\gamma = 0.5$ , et la probabilité d'exploration  $\varepsilon = 1$ . Les principales procédures impliquées dans la simulation sont récapitulées dans la Figure 4.4.

Nous avons effectué un ensemble de simulations, incluant : (i) le cas d'une charge de trafic constante pour toutes les classes de trafic, (ii) le cas de charges de trafic variables, et (iii) le cas d'une charge de trafic variable dans le temps.

### 2.4.1 Charge de trafic constante

La première expérience considère une charge de trafic constante pour les deux classes  $C_1$  et  $C_2$ . Les paramètres de simulation utilisés sont les mêmes que ceux utilisés pendant la période d'apprentissage et sont récapitulés dans le Tableau 4.3. Dans la cellule, les nouveaux appels et les handoffs arrivent suivant une loi de Poisson de paramètres  $\lambda_{i,nc}$  et  $\lambda_{i,ho}$ , respectivement (pour chaque classe de trafic  $C_i$ ). La durée d'un appel est exponentiellement distribuée de paramètre  $\mu_C$  ( $1/\mu_C = 120s$ ). La période de séjour d'un mobile dans une cellule est également exponentiellement distribuée de paramètre  $\mu_H$  ( $1/\mu_H = 60s$ ). Par conséquent, le temps de séjour d'un appel est exponentiel de paramètre  $\mu = \mu_C + \mu_H$  ( $1/\mu = 40s$ ). Pour le calcul de la charge en Erlangs, la formule (4.10) peut être utilisée.

	Type de l'appel	
	$C_1$	$C_2$
Nombre de canaux ( $b_i$ )	1	2
Temps de séjour ( $1/\mu$ )	40 s	40 s
Débit des arrivées ( $\lambda_{i,nc}$ )	$\lambda_1=180$ appels/heure	$\lambda_2=\lambda_1/2=90$ appels/heure

Tableau 4.3 Paramètres expérimentaux.

La Figure 4.6 (a) représente les gains reçus en suivant les politiques générées par les deux algorithmes TRL-CAC et NRL-CAC. Ils sont calculés toutes les dix minutes durant une heure

de simulation. Ces gains, dus à l'acceptation de nouveaux appels ou de handoffs pour les deux classes de trafic ( $C1$  ou  $C2$ ) dans la cellule, ne sont pas très importants comparés à ceux de la politique *gloutonne* (la différence est de l'ordre de 10%). La Figure 4.6 (b), quant à elle, montre que les pertes dus au rejet de nouveaux appels ou à l'échec d'un transfert inter-cellulaire ont été amplement réduites (la différence est de l'ordre de 71%).

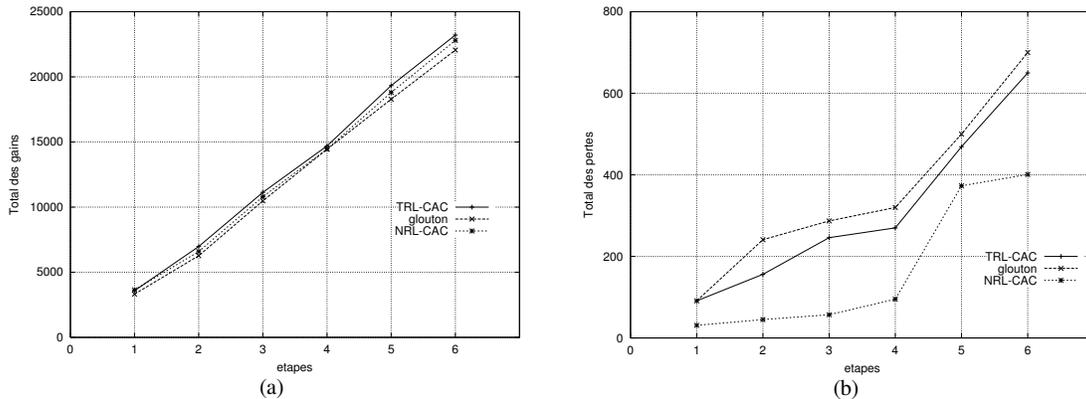


Figure 4.6 Total par heure (a) des gains (b) des pertes.

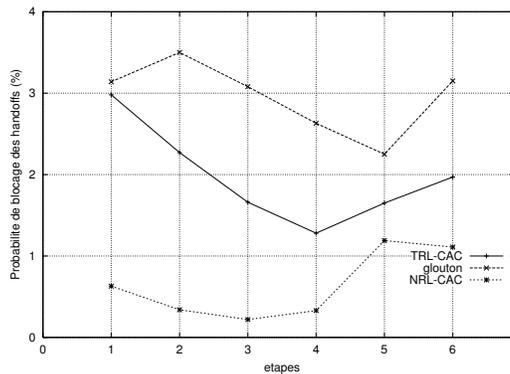


Figure 4.7 Probabilités de blocage des handoffs.

À partir des résultats de la Figure 4.7, nous pouvons remarquer que les probabilités de blocage des handoffs décroissent significativement en utilisant le Q-learning, en comparaison avec ceux de la politique gloutonne. La métrique probabilité de blocage des handoffs est donnée par la formule suivante :

$$P_{HO} = \frac{\text{nombre d'échecs de handoff dans le système}}{\text{nombre de tentatives de handoff dans le système}} \tag{4.7}$$

Nous avons aussi comparé les probabilités de blocage de handoff des appels de type  $C_1$  par rapport à ceux de type  $C_2$  dans les cas où l'algorithme TRL-CAC est utilisé (Figure 4.8 (a)),

ainsi que dans le cas où l'algorithme NRL-CAC est utilisé (Figure 4.8 (b)). Cette métrique est calculée pour chaque classe de trafic  $C_i$ , en utilisant la formule (4.8) suivante :

$$P_{HO(C_i)} = \frac{\text{nombre d'échecs de handoff de type } C_i \text{ dans le système}}{\text{nombre de tentatives de handoff de type } C_i \text{ dans le système}} \quad (4.8)$$

Étant donné que les appels de type  $C_1$  sont plus prioritaires<sup>7</sup> que les appels de type  $C_2$ , nous constatons que les probabilités de blocage de handoff de  $C_1$  ( $P_{HO(C_1)}$ ) sont plus petites que celles de  $C_2$  ( $P_{HO(C_2)}$ ).

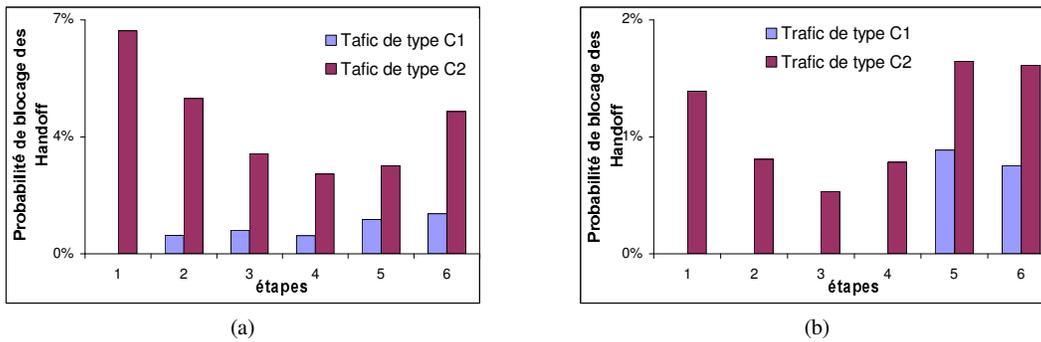


Figure 4.8 Probabilité de blocage des handoffs des appels de  $C_1$  vs.  $C_2$  en utilisant (a) TRL-CAC (b) NRL-CAC.

## 2.4.2 Charges de trafic variables

Dans cette deuxième expérience, nous avons utilisé les mêmes politiques que dans l'expérience précédente (charge de trafic constante), mais avec six charges de trafic différentes (pour les deux classes  $C_1$  et  $C_2$ ). Les charges sont données en terme de nombre d'appels/heure. Cependant, pour le calcul de la charge en Erlangs, la formule (4.10) peut être utilisée. Nous remarquons également, sur l'ensemble des figures ci-dessous, que les algorithmes proposés donnent de meilleurs résultats que d'autres heuristiques et ceci pour toutes les charges de trafic considérées, et plus particulièrement quand la charge de trafic est grande. Nous notons que lorsque la charge de trafic est petite, les avantages de TRL-CAC et NRL-CAC deviennent négligeables. Ceci peut facilement être expliqué par le fait que lorsque la charge de trafic est petite, il y a assez de canaux pour tous les appels et ainsi tous les appels sont acceptés. Ceci explique pourquoi, dans la Figure 4.10, la probabilité de blocage est presque égale à 0 quand la charge de trafic est fixée à 120 appels/h pour  $C_1$  et à 60 appels/h pour  $C_2$ . À partir de la Figure 4.11, nous pouvons également facilement remarquer que les probabilités de blocage de handoff

7. Les revenus associés aux appels de type  $C_1$  sont plus grands que ceux associés aux appels de type  $C_2$  (cf. Tableau 4.1).

de  $C_1$ , calculées en utilisant la formule (4.8) sont plus petites que celles de  $C_2$  pour toutes les charges du trafic considérées (les gains atteignent facilement les 85%).

Deux expériences ont été réalisées. La première compare les performances de nos algorithmes de contrôle d'admission (TQ-CAC et NQ-CAC). La seconde expérience compare ces deux algorithmes à la politique de réservation de canaux.

A. TQ-CAC vs. NQ-CAC

Nous remarquons que les résultats du premier algorithme TRL-CAC sont plus performants que ceux de NRL-CAC (cf. Figure 4.9). Ceci s'explique par le fait que NRL-CAC utilise une fonction d'approximation (un réseau de neurones) pour représenter les Q-valeurs, et donc a besoin de plus de temps pour converger. Les différences entre ces deux algorithmes, exprimées en terme de complexité de calcul et de besoins de stockage mémoire, sont discutées plus en détails dans la section 2.2.

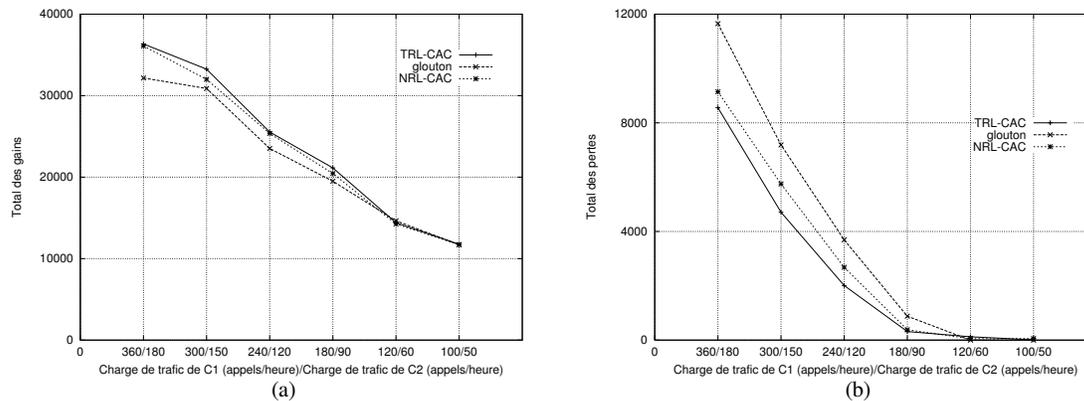


Figure 4.9 Total par heure (a) des gains (b) des pertes.

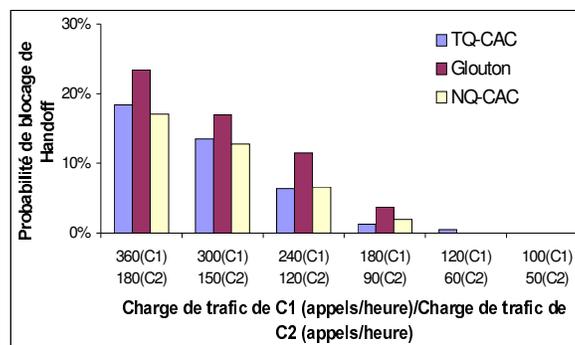


Figure 4.10 Probabilités de blocage des handoffs pour six différentes charges de trafic.

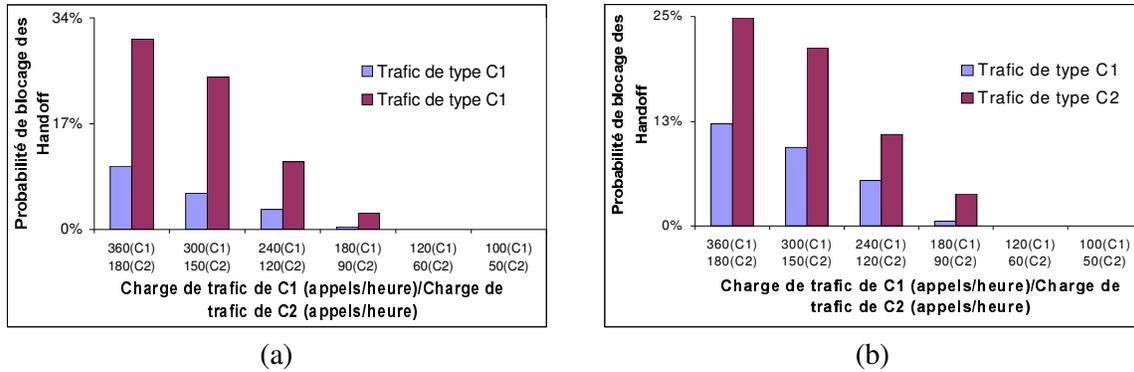


Figure 4.11 Probabilité de blocage des handoffs des appels de  $C_1$  vs.  $C_2$  en utilisant (a) TRL-CAC (b) NRL-CAC.

Cet ensemble de résultats illustre clairement que les deux algorithmes TRL-CAC et NRL-CAC offrent des possibilités intéressantes, pour améliorer significativement la tâche de contrôle d'admission sur un grand intervalle de charges de trafic. Il est intéressant d'observer que ni la table ni le réseau de neurones n'ont été réappris pour ces nouvelles charges, ce qui indique que le système possède des facultés d'adaptation et de généralisation.

### B. *Q-learning vs. techniques de réservation*

Nous avons également comparé nos résultats à : (1) guard channel avec seuils fixes – ces seuils ont été calculés pour une charge de trafic constante donnée par le Tableau 4.3 ; (2) guard channel avec seuils optimisés (le vecteur optimal  $s^*$  qui maximise  $F^s$ ) – ces seuils ont été calculés pour chaque valeur de la charge de trafic ; (3) glouton analytique – calculé par un modèle analytique exact (file M/M/N/N multi-classes). À partir des résultats de la Figure 4.12 (a), nous avons eu des profits allant jusqu'à 16%, et jusqu'à 35% dans la Figure 4.12 (b).

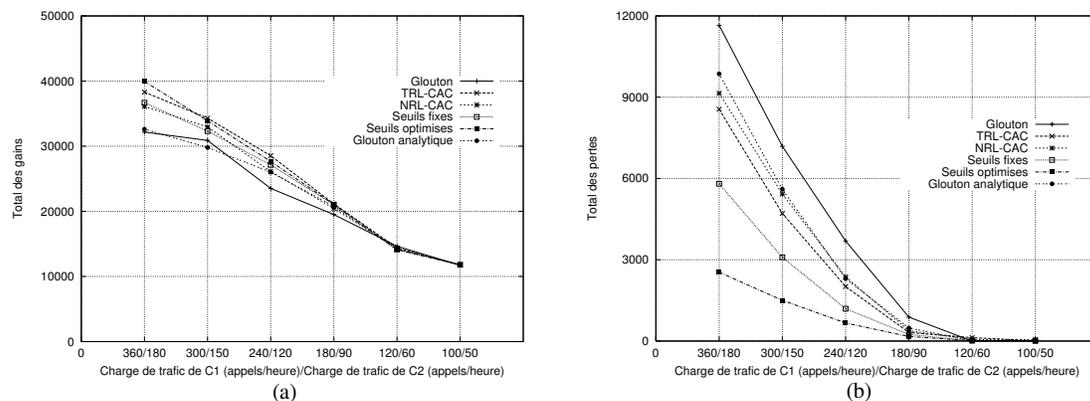


Figure 4.12 Total par heure (a) des gains (b) des pertes.

Nous remarquons dans la Figure 4.12, que la méthode de réservation de canaux avec seuils optimisés donne de meilleurs résultats que les deux algorithmes basés sur le Q-learning. En revanche, cette méthode suppose que les seuils optimisés soient calculés, hors ligne, pour chaque valeur de la charge de trafic. Au contraire, et comme mentionné auparavant, les algorithmes de contrôle d'admission basés sur le Q-learning possèdent des capacités d'adaptation et de généralisation, et donc les valeurs optimales des Q-valeurs ne sont pas recalculées pour chaque charge de trafic.

### 2.4.3 Charge de trafic variable dans le temps

La charge de trafic dans un système cellulaire est variable dans le temps. Dans cette dernière expérience, nous utilisons encore une fois la même politique que dans la première expérience (charge de trafic constante). Nous utilisons, comme dans [56], le modèle de trafic donné dans la Figure 4.13. Ce modèle décrit la variation des taux d'arrivées pendant une journée ouvrable. Dans une journée ouvrable, les heures de pointe sont à 11 heures et à 16 heures. La Figure 4.14 donne les résultats de simulation, avec l'hypothèse que les deux classes de trafic suivent le modèle de trafic donné dans la Figure 4.13. La charge de trafic maximum est fixée à 180 appels/h pour la classe C1 et à 90 appels/h pour la classe C2. Pour le calcul de la charge en Erlangs, la formule (4.10) peut être utilisée. Les probabilités de blocage ont été calculées sur une base d'une heure.

Les améliorations des algorithmes proposés sont bien apparentes par rapport à la politique *gloutonne*, et particulièrement pendant les pics de trafic (vers 11 heures et vers 16 heures). Comme dans l'expérience précédente, la méthode guard channel avec seuils optimisés donne de meilleurs résultats que les deux algorithmes reposant sur le Q-learning. En revanche, nous observons que NQ-CAC est plus performant que TQ-CAC dans ce cas. Ceci implique que l'approche utilisant les réseaux de neurones (NQ-CAC) a un potentiel d'adaptation plus important pour des charges de trafic variable dans le temps.

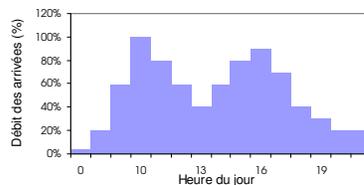


Figure 4.13 Modèle de trafic d'une journée ouvrable.

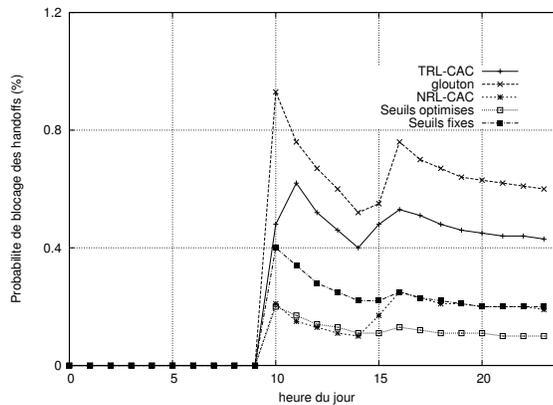


Figure 4.14 Probabilité de blocage des handoffs avec une charge de trafic variable dans le temps.

#### 2.4.4 Variation des valeurs des revenus immédiats

Le Q-learning fait partie des méthodes de résolution des problèmes d'apprentissage par renforcement, où il s'agit d'apprendre une fonction d'évaluation à partir d'un ensemble d'états et d'actions. En particulier, cette fonction d'évaluation  $Q(s,a)$  représente la moyenne des gains prévue, dans le cas où l'action  $a$  est exécutée lorsque l'état courant du système est  $s$ . Ainsi, le choix du revenu immédiat  $r$  associé à chacune des actions  $a$  est un point très important. Lors de cette étude, les revenus immédiats utilisés (cf. Tableau 4.1) ne représentent pas le coût réel de l'acceptation de chacune des classes de trafic. Ces revenus immédiats représentent une simple spécification de préférences entre ces classes de trafic. Nous avons donné de plus grandes valeurs de revenus pour les appels de type  $C_1$ , puisque nous supposons que les appels qui consomment moins (1 canal seulement) sont plus prioritaires. Nous fixons, dans le Tableau 4.1, le revenu immédiat des appels des handoff dix fois plus grand que ceux des nouveaux appels. Ce ratio est un paramètre décisif pour la comparaison des performances des différents algorithmes proposés. Dans le Tableau 4.4, nous pouvons voir l'impact de ce ratio, en comparant les performances des différents algorithmes proposés pour des valeurs de ratio égales à 3, 5, et 10.

Valeurs des Revenus immédiats		Ratio	Gains en terme de probabilités de blocage des handoffs pour $C_1$ vs. $C_2$
$\eta_1 = 2$	$\eta_3 = 6$	3	70%
$\eta_2 = 1$	$\eta_4 = 3$		
$\eta_1 = 2$	$\eta_3 = 10$	5	75%
$\eta_2 = 1$	$\eta_4 = 5$		
$\eta_1 = 2$	$\eta_3 = 20$	10	83%
$\eta_2 = 1$	$\eta_4 = 10$		

Tableau 4.4 Impact de la valeur de revenus immédiats.

### 2.4.5 Apprentissage en ligne

Pour la dernière expérience, nous avons observé la performance du Q-learning en ligne, quand les deux opérations d'apprentissage et de contrôle d'admission sont réalisées simultanément. Nous avons utilisé la politique apprise lors de la première expérience (charge de trafic constante), mais en continuant le processus d'apprentissage. Nous supposons que les débits d'arrivée varient dans le temps, et sont répartis selon le modèle de la Figure 4.13 décrivant la variation des taux d'arrivées pendant une journée ouvrable. La Figure 4.15 montre les probabilités de blocage des handoffs calculées pour une période de 2 jours (48 heures). Nous pouvons remarquer dans cette figure, quelques améliorations en raison de l'apprentissage en ligne, dans le sens où le cumul des probabilités de blocage durant le deuxième jour est moins important que celui du premier jour.

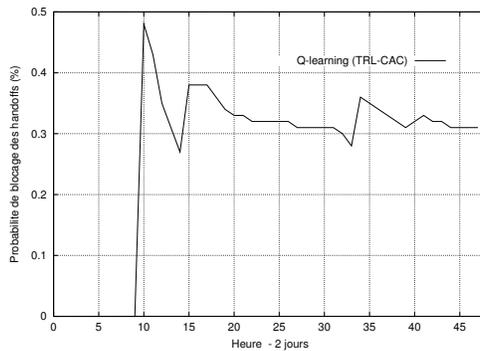


Figure 4.15 Performance de l'exécution en ligne du Q-learning.

## 3. Système FCA

La partie précédente suppose une seule cellule de type FCA. En revanche, dans cette partie, nous considérons tout un système cellulaire de type FCA avec plusieurs cellules. L'algorithme proposé est un algorithme distribué sur l'ensemble des cellules.

Le système étudié est un système FCA avec  $N$  cellules et  $L$  canaux disponibles. Afin de partager l'ensemble des canaux disponibles du système entre un plus grand nombre d'utilisateurs,  $L$  est divisé selon un certain *modèle de réutilisation* (cf. encadré ci-après).

<p><b>Modèle de réutilisation dans FCA</b></p>	<p>L'ensembles des canaux <math>L</math> est divisé en <math>S</math> groupes, chacun composé de <math>L/S</math> canaux. Cette allocation est faite au moyen d'une découpe géographique basée sur le paradigme hexagonal, permettant la réutilisation des fréquences sur des cellules éloignées. <math>S</math> est appelé "motif" et contient une et une seule fois l'ensemble des canaux radio <math>L</math>. Ce motif est répété sur toute la surface à couvrir.</p> <div data-bbox="527 619 1201 892" style="text-align: center;"> </div> <p>Plus le motif est grand, plus la distance de réutilisation est grande. L'ensemble des <math>L/S</math> canaux est alloué de façon permanente à chaque cellule. Un nouvel appel dans une de ces cellules ne peut être servi que si un canal libre dans cet ensemble est disponible. Si tous les canaux sont utilisés, le nouvel appel sera bloqué et perdu même si d'autres canaux sont disponibles dans les cellules du motif <math>S</math>. Cette méthode présente l'avantage d'être évolutive en fonction du trafic : (i) réduction de la taille des cellules en cas de saturation du réseau (zones géographiques à forte densité de population - zone urbaine); (ii) augmentation de la taille des cellules (zones géographiques à faible densité de population - zone rurale). Cette méthode permet d'accepter de nombreux utilisateurs, mais elle implique de résoudre les phénomènes d'interférences, qui pénalisent fortement la qualité d'un service de transmission de voix et/ou de données. Afin d'éviter qu'un niveau d'interférence trop élevé perturbe les communications, la réutilisation des fréquences est faite en respectant une distance minimum (appelée distance de résolution).</p>
--	---

Nous proposons une nouvelle approche pour résoudre le problème du contrôle d'admission des appels CAC dans un réseau cellulaire de type FCA QCAC-FCA<sup>8</sup>. Les politiques optimales sont obtenues en utilisant l'algorithme d'apprentissage par renforcement (RL).

8. QCAC-FCA : Implémentation de l'algorithme du Q-learning pour résoudre le problème CAC dans un système de type FCA.

L'ensemble des états " $s$ ", des actions " $a$ " et des revenus " $r$ " ne changent pas par rapport à ceux définis pour le cas d'une seule cellule. Nous avons utilisé, pour les études de performances, les paramètres expérimentaux récapitulés dans le Tableau 4.5.

### 3.1 Implantation de l'algorithme QCAC-FCA

Nous allons décrire l'implantation de l'algorithme QCAC-FCA capable de résoudre le SMDP associé au problème de contrôle d'admission dans un système FCA. Ici, la même question se pose : où stocker les valeurs de la Q-fonction ?

#### 3.1.1 Représentation des Q-valeurs

Comme dans TRL-CAC, nous avons utilisé une simple table pour représenter les valeurs de la fonction  $Q$ . Cette dernière requiert  $264 \times 5 = 1320$  unités mémoires uniquement pour le stockage de ses valeurs<sup>9</sup>. Dans QCAC-FCA, tous les appels nécessitent un seul canal.

#### 3.1.2 Implantation

Pour les mêmes raisons que dans le cas d'une seule cellule, nous ne tenons pas compte des états associés aux départs des appels. Nous avons initialisé les Q-valeurs à des valeurs aléatoires très petites. Lorsqu'un appel arrive (nouveau ou handoff) dans la cellule  $i$ , l'algorithme détermine si la qualité de service est maintenue en acceptant cet appel (en vérifiant tout simplement s'il y a assez de canaux disponibles dans la cellule  $i$ ). Si cette qualité de service est violée, l'appel est rejeté ; sinon l'action est choisie selon la formule (4.2).

Le nombre de cellules est de  $N = 36$  cellules hexagonales, et le nombre de canaux disponibles dans un motif de  $S = 7$  cellules est de  $L = 70$  canaux (cf. Figure 4.19). Chaque cellule lui est attribuée  $L/S = 70/7 = 10$  canaux de façon permanente. La constante de propagation temporelle a été fixée à  $\gamma = 0.5$ . Nous avons utilisé pendant la période d'apprentissage un taux d'apprentissage fixe  $\alpha = 0.1$  qui semble donner de bons résultats quoique les théorèmes de convergence exigent que  $\alpha$  soit décroissant avec le temps.

#### 3.1.3 Exploration

Pour une exécution correcte et efficace de l'algorithme du QCAC-FCA, l'action est choisie non selon la formule (4.2), mais selon une distribution de Boltzmann [50] donnée par le formule (4.9), ceci pendant une période d'apprentissage assez longue.

---

9. La taille réelle de l'espace des états a été réduite de 800 ( $10 \times 10 \times 4 \times 2$ ) états à 264 états, en éliminant les combinaisons impossibles.

Des études comparatives des différentes méthodes d'exploration (aléatoire,  $\epsilon$ -gloutonne et Boltzmann), en termes qualitatifs et quantitatifs, ont été faites dans [48].

$$p_s(a) = \frac{e^{\frac{Q(s,a)}{T}}}{\sum_{b \in A} e^{\frac{Q(s,b)}{T}}} \quad (4.9)$$

L'idée est de favoriser, au début, l'exploration (la probabilité d'exécuter des actions autres que celles avec la Q-valeur la plus élevée), en utilisant toutes les actions possibles avec une même probabilité. L'idée ensuite est de tendre, au fur et à mesure, vers l'utilisation de l'action de plus grande Q-valeur. La température  $T$  contrôle ainsi cette quantité d'exploration. Si  $T$  est grande ou si les Q-valeurs sont toutes identiques, l'agent va sélectionner une action aléatoirement. Si  $T$  est petite et les Q-valeurs sont différentes, il tendra à sélectionner l'action avec la Q-valeur la plus élevée. La température  $T$  doit être choisie assez élevée au début, pour décroître vers 0 quand l'apprentissage sera terminé. Les valeurs apprises seront employées après pour initialiser les Q-valeurs dans QCAC-FCA.

### 3.2 Résultats expérimentaux

Afin d'évaluer les avantages de notre solution, nous avons simulé un réseau cellulaire en utilisant une simulation à événements discrets. Nous l'avons comparée à une politique que nous avons appelée politique *gloutonne* [65] (glouton-FCA<sup>10</sup>). Les performances des algorithmes ont été évaluées en terme de : (i) revenus reçus de la part des clients acceptés - les gains, (ii) de revenus reçus de la part des clients rejetés - les pertes, et (iii) de probabilités de blocage des handoffs.

Les paramètres utilisés pendant la période d'apprentissage sont récapitulés dans le Tableau 4.5. Les appels arrivent suivant une loi de Poisson de paramètres  $\lambda_1$  (=120 appels/heure) pour les appels de la classe  $C_1$  et  $\lambda_2$  (=60 appels/heure) pour les appels de la classe  $C_2$ . Nous supposons que la durée d'un appel est exponentiellement distribuée de paramètre  $\mu_C$  ( $1/\mu_C = 180s$ ). Le temps de séjour d'un mobile dans une cellule est également exponentiellement distribuée de paramètre  $\mu_H$  ( $1/\mu_H = 60s$ ). La charge de trafic correspondante est égale à  $\rho=9$  Erlangs et est calculée comme suit :

---

10. Glouton-FAC: Politique qui accepte un appel si la contrainte de capacité n'est pas violée en acceptant cet appel .

$$\rho = \sum_{i=1}^2 \frac{\lambda_i}{\mu_i} \times b_i \tag{4.10}$$

	Type de l'appel	
	C <sub>1</sub>	C <sub>2</sub>
Nombre de canaux ( <i>b<sub>i</sub></i> )	1	1
Durée de l'appel ( $1/\mu_C$ )	180 s	180 s
Temps de séjour dans une cellule ( $1/\mu_H$ )	60 s	60 s
Débit des arrivées	$\lambda_1 = 120 \text{ appels / heure} \quad \lambda_2 = \lambda_1 / 2 = 60 \text{ appels / heure}$	

Tableau 4.5 Paramètres expérimentaux pendant la période d'apprentissage.

Un ensemble de simulations a été effectué, incluant : (i) le cas d'une charge de trafic uniformément répartie sur toutes les cellules, (ii) le cas d'une charge non-uniformément répartie, (iii) et le cas d'une charge de trafic variable dans le temps. Les résultats expérimentaux sont présentés dans les figures allant de la Figure 4.16 à la Figure 4.24.

### 3.2.1 Répartition uniforme du trafic

La première expérience considère une charge de trafic constante dans les 36 cellules pour les deux classes de trafic. Nous avons utilisé les politiques apprises pendant la période d'apprentissage, mais avec cinq charges de trafic différentes (pour les deux classes C<sub>1</sub> et C<sub>2</sub>). Les différentes charges sont récapitulées dans le Tableau 4.6. Les charges sont données en terme de nombre d'appels/heure. Cependant, pour le calcul de la charge en Erlangs, la formule (4.10) peut être utilisée.

	Type de l'appel	
	C <sub>1</sub>	C <sub>2</sub>
	60	30
Charges de	80	40
trafic	100	50
(appels/heure)	120	60
	140	70

Tableau 4.6 Paramètres expérimentaux.

À partir des résultats de la Figure 4.16, nous pouvons remarquer que les probabilités de coupure diminuent de manière significative en utilisant le Q-learning (QCAC-FCA), par rapport aux probabilités de coupure des politiques gloutonnes (glouton-FCA). Cette observation est confirmée même pour les autres charges de trafic listés dans le Tableau 4.6, et plus particulièrement lorsque la charge de trafic est grande.

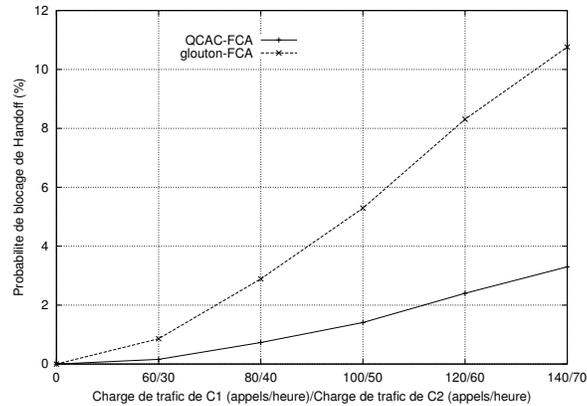


Figure 4.16 Probabilités de blocage des handoffs pour cinq charges de trafic différentes.

La Figure 4.17(a) représente l'ensemble des revenus reçus en suivant les politiques générées par QCAC-FCA, pour toutes les charges de trafic considérées dans le Tableau 4.6 et calculés durant une heure de simulation. Ces gains sont dus à l'acceptation de nouveaux appels ou de handoffs pour les deux classes de trafic dans le système cellulaire. Les gains générés par QCAC-FCA sont plus importants que ceux générés par glouton-FCA. La Figure 4.17(b), quant à elle, montre que les pertes dus au rejet de nouveaux appels ou à l'échec d'un transfert inter-cellulaire ont été amplement réduites, en utilisant le Q-learning. Cette réduction est particulièrement importante lorsque la charge de trafic est grande.

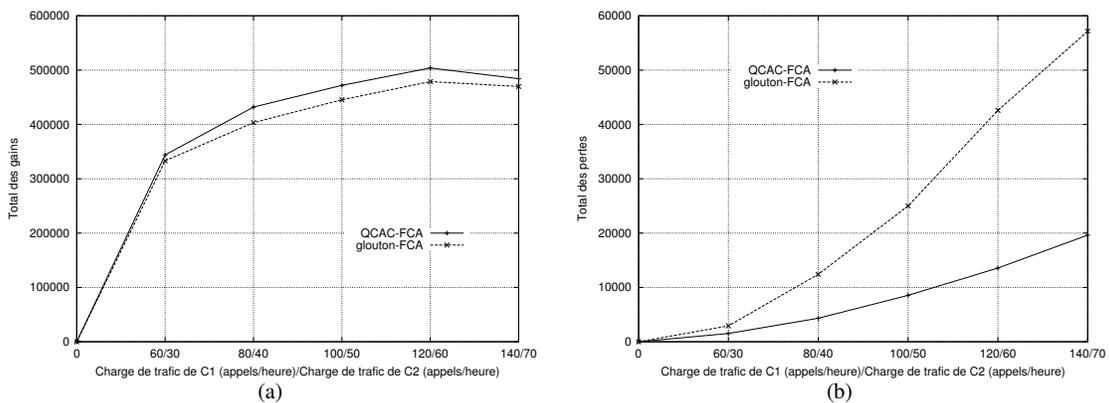


Figure 4.17 Total par heure (a) des gains (b) des pertes.

Nous avons aussi comparé les probabilités de blocage de handoff des appels de la classe  $C_1$  par rapport à ceux de la classe  $C_2$  dans les cas où l'algorithme QCAC-FCA est utilisé (Figure 4.18 (a)). Étant donné que les appels de la classe  $C_1$  sont plus prioritaires<sup>11</sup> que les appels de la

11. Les revenus associés aux appels de type  $C_1$  sont plus grands que ceux associés aux appels de type  $C_2$  (cf. Tableau 4.1).

classe  $C_2$ , nous remarquons que les probabilités de blocage des handoffs de  $C_1 (P_{ho(C_1)})$  sont plus petites que celles de  $C_2 (P_{ho(C_2)})$  lorsque QCAC-FCA est employé.

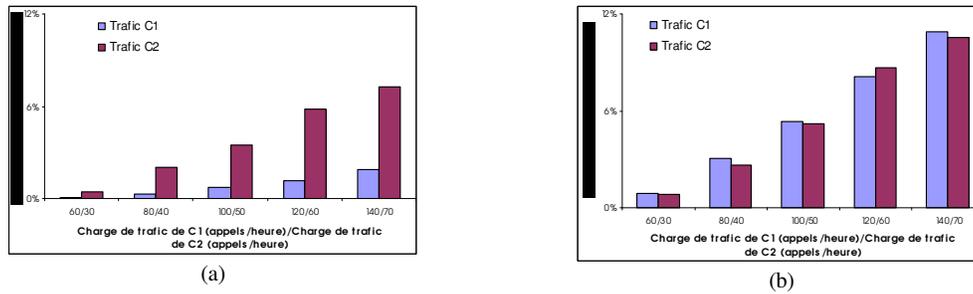


Figure 4.18 Probabilité de blocage de handoff de  $C_1$  vs.  $C_2$  en utilisant (a) QCAC-FCA (b) glouton-FCA.

Tout ceci illustre clairement que QCAC-FCA offre des possibilités intéressantes pour améliorer significativement les performances du système sur un large intervalle de charges de trafic. Il est intéressant d'observer que les Q-valeurs n'ont pas été réappries pour ces nouvelles charges, indiquant que le système possède des facultés d'adaptation et de généralisation.

### 3.2.2 Répartition non uniforme du trafic

Dans cette deuxième expérience, nous avons utilisé les politiques apprises pendant la période d'apprentissage mais, comme le montre Figure 4.19, les charges de trafic dans ce cas ne sont plus uniformément réparties sur les 36 cellules. La moyenne des arrivées des appels de type  $C_1$  est de 100 appels/heure et la moyenne des arrivées des appels de type  $C_2$  est de 50 appels/heure. La charge de trafic correspondante est égale à  $\rho=7.5$  Erlangs.

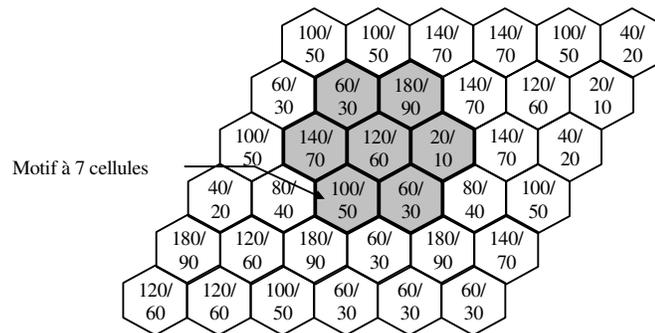


Figure 4.19 Répartition non uniforme du trafic (charge du trafic de type  $C_1$ /charge du trafic de type  $C_2$ ).

La Figure 4.20 donne les probabilités de blocage des handoffs dans diverses configurations. Ces configurations sont obtenues en multipliant les taux d'arrivée de base, donnés dans la Figure 4.19, par 60%, 80%, 100%, 120% et 140%. Les figures ci-dessous indiquent des améliorations significatives des politiques de Q-learning, par rapport aux politiques traditionnelles.

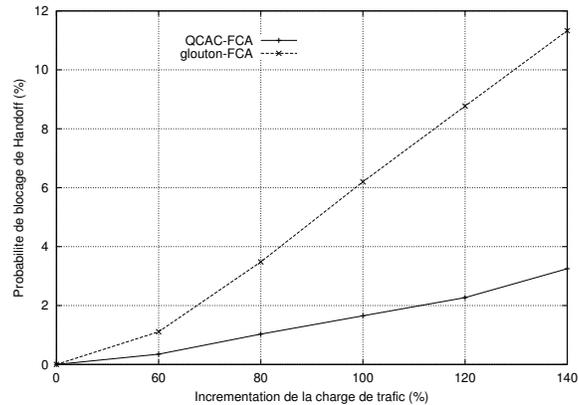


Figure 4.20 Probabilités de blocage des handoffs pour des charges de trafic différentes.

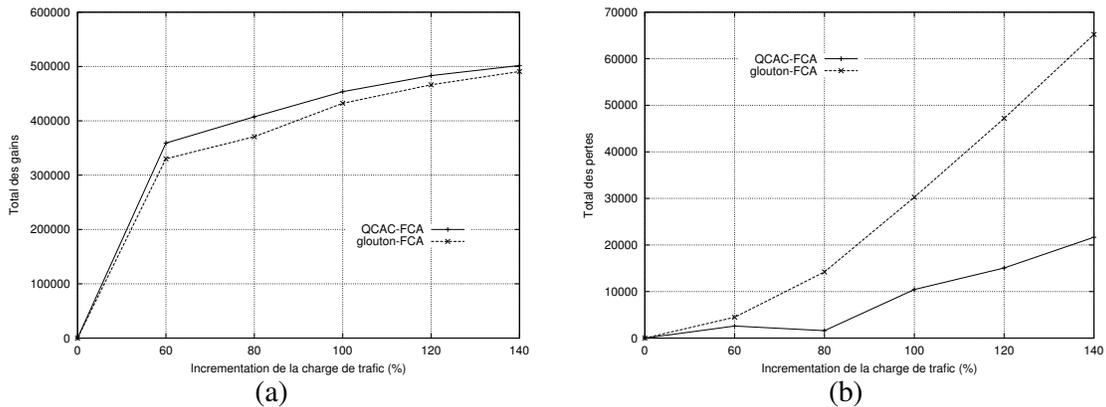


Figure 4.21 Total par heure (a) des gains (b) des pertes.

À partir de la Figure 4.22, nous pouvons aisément constater que les probabilités de blocage de handoff des appels de type  $C_1$ , calculées en utilisant la formule (4.8) sont plus petites que celles de  $C_2$  pour QCAC-FCA.

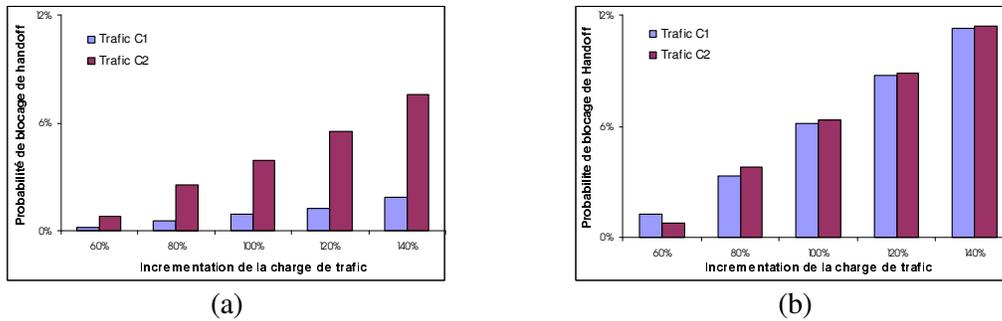


Figure 4.22 Probabilité de blocage de handoff de  $C_1$  vs.  $C_2$  en utilisant (a) QCAC-FCA (b) glouton-FCA.

### 3.2.3 Charge de trafic variable dans le temps

En utilisant la même politique apprise lors d'une charge de trafic constante, nous testons notre algorithme pour une charge de trafic variable dans le temps. La Figure 4.23 donne les résultats de simulation avec l'hypothèse que les deux classes de trafic sont uniformément réparties et suivent le même modèle de trafic que celui de la Figure 4.13. La charge de trafic maximum est fixée à 120 appels/h pour la classe  $C_1$ , et à 60 appels/h pour la classe  $C_2$  ( $\rho=9$  Erlangs). Les probabilités de blocage ont été calculées sur une base d'une heure. Les améliorations des algorithmes d'apprentissage proposés sont pleinement visibles par rapport aux politiques gloutonnes, et particulièrement pendant les pics de trafic (à 11 heures et à 16 heures). Les gains, par rapport à la politique gloutonne, sont de l'ordre de 71% pour QCAC-FCA.

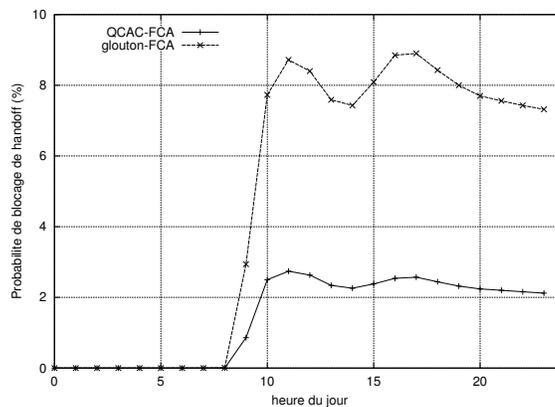


Figure 4.23 Probabilité de blocage des handoffs avec une charge de trafic variable dans le temps et un trafic uniformément réparti.

Nous avons également examiné le cas où la charge de trafic est variable dans le temps, et non uniformément répartie sur les 36 cellules. Nous faisons les mêmes hypothèses que dans le

cas uniforme. Dans la Figure 4.24, les améliorations qu'apportent les algorithmes d'apprentissage par renforcement, par rapport aux politiques traditionnelles, sont visibles. Néanmoins, une amélioration plus significative a été constatée dans le cas uniforme. Les gains dus aux algorithmes RL sont d'environ 71% pour QCAC-FCA.

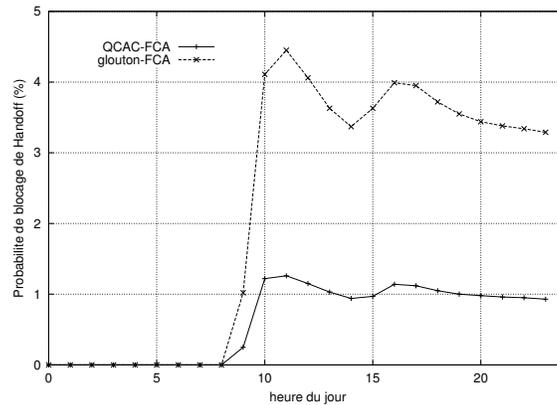


Figure 4.24 Probabilité de blocage des handoffs avec une charge de trafic variable dans le temps et un trafic non uniformément réparti.

## 4. Conclusion

Dans ce chapitre, nous avons proposé une nouvelle approche pour résoudre le problème du contrôle d'admission d'appel CAC dans un réseau cellulaire supportant plusieurs classes de trafic. Nous avons formulé le problème sous forme d'un processus de décision semi-markovien (SMDP). Les solutions optimales sont obtenues en utilisant un système autonome basé sur deux implantations de l'algorithme d'apprentissage par renforcement 'Q-Learning'.

Outre l'originalité de notre proposition, les avantages acquis, en utilisant une telle approche, peuvent être récapitulés comme suit. Tout d'abord, l'apprentissage fournit une voie simple pour obtenir une solution optimale au processus de décision semi-markovien. Outre la nature intelligente de notre solution, et contrairement aux autres approches permettant de résoudre ce SMDP (programmation dynamique, par exemple), elle est adaptable aux variations de l'état du réseau (i.e. variations des conditions de trafic, incluant une répartition uniforme et non uniforme du trafic, un trafic variable dans le temps, etc.). Ainsi, n'importe quel événement imprévu, dû aux variations dans les conditions de trafic, peut être considéré comme une nouvelle expérience pour améliorer l'adaptation et l'apprentissage du système. En second lieu, et comme les algorithmes proposés sont exécutés dans un environnement temps réel, il est donc possible d'effectuer en ligne la tâche d'apprentissage tout en exécutant en temps réel le contrôle

d'admission. Troisièmement, la politique d'acceptation peut être déterminée rapidement et avec peu d'efforts de calcul. Elle est facilement implantable dans chaque station de base, ce qui la rend plus attrayante. Nous démontrons, également, que les algorithmes de CAC proposés donnent de meilleurs résultats que d'autres heuristiques.

Comme perspectives à ces travaux, il nous semble intéressant d'intégrer les mécanismes proposés (QCAC-FCA et Q-DCA) à la technique d'allocation hybride HCA. Le premier algorithme QCAC-FCA serait utilisé, dans chacune des cellules, afin de gérer l'ensemble des canaux alloués de façon permanente à la cellule. Le second algorithme Q-DCA, quant à lui, serait utilisé pour gérer dynamiquement le reste des canaux contenus dans un pool commun, et utilisable par toutes les cellules.



# Chapitre 5

## Allocation dynamique et adaptative des ressources

### 1. Motivations

Nous avons vu dans le Chapitre 2 que dans les stratégies d'allocation dynamiques des ressources (DCA), tous les canaux disponibles dans le système sont mis dans un pool commun (ou *common pool*) utilisable par toutes les stations de base. Lors d'une demande de communication, une cellule choisit un canal du pool commun<sup>12</sup>, qui sera restitué à la fin de l'appel.

Différents algorithmes d'allocation dynamiques de ressources ont été comparés au Chapitre 2, en termes de performance, flexibilité et complexité. L'une des stratégies la plus utilisée dans la littérature, appartient à une classe d'algorithmes appelée "*DCA à recherche exhaustive - exhaustive searching DCA*" [17][18][19][20][21][31]. Dans ce type d'algorithmes, une récompense (coût) est associée à chaque canal disponible. Quand un nouvel appel arrive, le système recherche *exhaustivement* le canal avec la récompense maximale (coût minimal) et l'affecte à l'appel. La majorité de ces propositions ne considèrent pas les politiques de contrôle d'admission CAC en tant que moyens pour empêcher la congestion [66]. Le rapport entre l'allocation de ressources et le contrôle d'admission ont été étudiés précédemment [38][67][68].

Ce chapitre propose une nouvelle approche pour résoudre le problème de l'allocation dynamique des ressources, considérant également le contrôle d'admission d'appels (CAC) dans les systèmes DCA supportant plusieurs classes de trafic. Pour les mêmes raisons citées dans le

---

12. tout en respectant le rapport signal/interférence *C/I*.

chapitre précédent, le contrôle d'admission d'appels est indispensable lorsque le réseau supporte plusieurs classes de clients. Les politiques d'allocation sont obtenues en utilisant l'algorithme d'apprentissage par renforcement (RL) : Q-learning.

Les fonctions principales de cet algorithme, que nous avons appelé Q-DCA<sup>13</sup> et que nous allons présenter dans ce chapitre, sont : (i) accepter des clients et d'en rejeter d'autres, et (ii) allouer le meilleur canal disponible aux clients acceptés. L'objectif consiste à maximiser le cumul des revenus reçus au cours du temps.

Nous allons maintenant décrire, la formulation du problème par un SMDP ainsi que l'implantation de l'algorithme du Q-learning capable de résoudre ce SMDP.

## 2. Formulation du problème sous forme de SMDP

Ce travail est une extension du travail de *Nie et Haykin* [56], présenté plus en détail dans le Chapitre 3, et faisant partie des algorithmes DCA à recherche exhaustive. Nous envisageons de résoudre non seulement le problème d'allocation dynamique de ressources, mais également le problème du contrôle d'admission d'appels dans un réseau cellulaire. Ce système contient  $N$  cellules et  $M$  canaux disponibles maintenus dans un pool commun. Ce système supporte deux classes de trafic (contrairement à [56] où une seule classe de trafic –appel téléphonique– était prise en compte). Chaque canal peut être temporairement alloué à n'importe quelle cellule, à condition que la contrainte sur la distance de réutilisation soit satisfaite (une qualité donnée de signal soit maintenue).

Les appels arrivent dans la cellule, puis partent. Le réseau devra donc, choisir d'accepter ou de rejeter ces demandes de connexion. Si l'appel est accepté, le système lui alloue un des canaux disponibles du pool commun. L'objectif pour l'opérateur du réseau est de trouver une politique d'allocation dynamique de ressources capable de maximiser l'ensemble des gains à long terme et de réduire les probabilités de blocage de handoff (contrairement à [56] qui n'accorde aucune priorité aux handoffs ni dans les simulations ni dans la formulation sous forme de programme dynamique).

---

13. Q-DCA : Implémentation de l'algorithme du Q-learning pour résoudre les problèmes d'allocation dynamique de ressources et de CAC dans un système DCA.

Les paramètres expérimentaux utilisés sont récapitulés dans le Tableau 4.5. L'ensemble des états "s", des actions "a" et des revenus "r" ont été définis comme suit :

**Etats :** Nous définissons l'état du système  $s=(i,A(i),x,e)$  où :

–  $i \in \{1, \dots, N\}$ , est l'indice de la cellule où se produit l'événement  $e$  ;

–  $A(i) \in \{1, 2, \dots, M\}$ , est le nombre de canaux disponibles dans la cellule  $i$ . Il dépend des conditions d'utilisation des canaux dans cette cellule et dans l'ensemble des cellules interférentes<sup>14</sup>  $I(i)$  à cette cellule. Pour obtenir  $A(i)$ , nous avons défini un vecteur de "disponibilité"  $u_q$  de dimension  $M$  pour chaque cellule  $q, q= 1,2, \dots, N$  comme suit :

$$u_{qk} = \begin{cases} 0, & \text{si le canal } k \text{ est libre dans la cellule } q \\ n, & \text{sinon} \end{cases}, \quad \text{où } n > 0.$$

En utilisant  $u_{ik}$ ,  $A(i)$  peut être calculée à partir de la formule suivante

$$A(i) = \sum_{k=1}^M 1_{\{u_{ik}=0\}}$$

–  $x=(x_1, x_2)$ , où  $x_1$  et  $x_2$  représentent le nombre d'appels de chacune des classes de trafic ( $C_1$  et  $C_2$ , respectivement) dans la cellule.

–  $e=\{1,2,3,4\}$  est le même que dans la formule (4.1).

Pour les mêmes raisons que celles mentionnées dans le chapitre précédent (système FCA), nous ne tenons pas compte des états associés au départ des appels.

**Actions :** Nous combinons les notions d'allocation dynamique de ressources et de contrôle d'admission. Ainsi, appliquer une action consiste à accepter/rejeter la demande de connexion d'un appel, et à lui allouer un des  $A(i)$  canaux disponibles dans la cellule  $i$ , en cas d'acceptation. L'ensemble des actions possibles est donc  $A=\{0,1,2,..,M\}$  où :

$$a_i = \begin{cases} k, & \text{accepter l'appel et lui allouer le canal } k \\ 0, & \text{rejeter l'appel} \end{cases}$$

14. L'ensemble des cellules voisines, qui sont à une distance inférieure à distance de réutilisation  $D$ .

où  $k = 1, 2, \dots, M$  et  $u_{ik} = 0$ .

**Revenus :** Le revenu  $r(s,a)$  représente le gain instantané obtenu après l'exécution d'une action  $a$  lorsque le système est dans un état  $s$ .

$$r(s,a) = \begin{cases} (n_1(a)r_1 + n_2(a)r_2 + n_3(a)r_3) \times \eta_i, & \text{si } a = 1, \dots, M \text{ et } e = e_i \\ 0, & \text{si } a = 0 \end{cases} \quad (5.1)$$

La formule (5.1) implique les opérations suivantes. Quand un appel arrive dans une cellule  $i$  et qu'il est rejeté ( $a=0$ ), le revenu associé sera *nul*. En revanche, s'il est accepté ( $a \neq 0$ ), ce revenu sera égal au coût du choix du canal  $a$  pour servir la demande de connexion. Il existe différentes manières pour définir  $r$ . Ici, nous prenons en compte le type de l'appel et comme dans [56] les conditions d'utilisation dans les cellules co-canaux liées à la cellule  $i$ . Pour obtenir  $n_1(k)$ ,  $n_2(k)$  et  $n_3(k)$  à un instant donné  $t$ , nous avons défini un vecteur de dimension  $M$  représentant le *statut* du canal  $k$  pour chaque cellule  $q$  comme suit :

$$s_{qk} = \begin{cases} 1, & \text{si le canal } k \text{ est utilisé dans la cellule } q \\ 0, & \text{sinon} \end{cases}$$

Le paramètre  $\eta_i$  permet de prendre en compte le type de l'appel courant (Nouvel ou handoff). Les valeurs de  $\eta_i$  sont définies dans le Tableau 4.1.

En résumé, nous avons choisi l'ensemble des états comme étant  $\{s = (i, A(i), (x_1, x_2), e)\}$ , où  $A(i)$  représente le nombre de canaux disponibles dans la cellule  $i$ ,  $x_k$  représente le nombre d'appels en cours de la classe  $C_k$ , et  $e \in \{1,2,3,4\}$  indique l'arrivée d'un nouvel appel ou d'un appel de handoff dans une cellule. Quand un événement se produit, l'agent doit choisir une des actions possibles  $A(s) = \{0 = \text{rejeter}\} \cup \{1, \dots, M\}$ . Quand un appel se termine, aucune mesure ne doit être prise. L'agent aura à déterminer les politiques permettant d'accepter ou de rejeter un appel et à lui allouer, dans le cas d'acceptation, le canal qui permette de maximiser le cumul des gains reçus, en ne connaissant que l'état actuel du réseau  $s$ . Ce système constitue un SMDP avec comme espace d'états un ensemble fini  $S = \{(i, A(i), x, e)\}$ , et comme espace d'actions possibles un ensemble fini  $A = \{0, 1, \dots, M\}$ . Contrairement à d'autres travaux, Q-DCA résout non seulement le problème d'allocation dynamique de ressources mais également le problème de contrôle d'admission dans un réseau cellulaire supportant plus d'une seule classe de trafic. Q-DCA prend en considération le type de l'appel et accorde, ainsi, une priorité aux handoffs.

### 3. Implantation de l'algorithme Q-DCA

Après avoir formulé le problème sous forme de SMDP, nous allons décrire l'implantation de l'algorithme Q-learning capable de le résoudre.

#### 3.1 Représentation des Q-valeurs

L'ensemble des couples état-action  $(s,a)$  étant très grand, l'utilisation d'une simple table devient rédhibitoire à cause des immenses besoins de stockage. L'utilisation d'une fonction d'approximation, pour représenter les Q-valeurs, devient alors indispensable. Dans le chapitre précédent, nous avons utilisé un réseau de neurones pour représenter les Q-valeurs, mais cette fois-ci nous avons choisi une approximation à base d'agrégation d'états.

Nous avons partitionné l'espace d'états  $S$  en sous-ensembles  $S_0, S_1, \dots, S_K$ , et avons défini un vecteur  $\phi$  de dimension  $K$  dont le  $m^{\text{ème}}$  composant est une approximation de la fonction  $Q$  pour tout état  $s \in S_m$  sous l'action  $a$ . En d'autres termes

$$\tilde{Q}(s, a, \phi) = \phi(m, a) \quad \text{si } s \in S_m \quad (5.2)$$

Dans Q-DCA, l'agrégation d'états peut être interprétée comme une architecture où une valeur  $\phi(s,.)$  est attribuée à tous les états  $s$  partageant les mêmes caractéristiques du trafic. Ainsi, au lieu de spécifier précisément le nombre d'appels  $x_i$  pour chaque classe de trafic  $C_i$ , nous avons choisi de caractériser le trafic comme suit : (1) bas, (2) moyen, (3) élevé. L'espace des états est ainsi réduit ; et une simple table peut être utilisée pour représenter les états agrégés. Etant donné que des états identiques (i.e. états ayant un nombre d'appels en cours similaire) ont les mêmes Q-valeurs, la perte en performance liée à l'agrégation devient négligeable [65].

#### 3.2 Implantation

L'organigramme de la Figure 5.1 récapitule les différentes phases de l'algorithme Q-DCA. Nous avons initialisé les Q-valeurs à des valeurs aléatoires très petites. Lorsqu'un appel arrive (nouvel appel ou handoff) dans la cellule  $i$ , l'algorithme détermine si la qualité de service n'est pas violée en acceptant cet appel (en vérifiant tout simplement s'il y a des canaux libres dans le pool commun). Si cette qualité de service est violée, l'appel est rejeté ; sinon l'action est choisie selon la formule :

$$a = \arg \max_{a \in A(s)} Q^*(s, a) \quad (5.3)$$

où  $A(s) = \{0 = \text{rejeter}, 1, 2, \dots, M\}$ .

La formule (5.3) implique les procédures suivantes. Quand il y a une tentative de connexion d'un appel dans la cellule  $i$ , la Q-valeur de rejet ( $a = 0$ ) ainsi que les Q-valeurs d'acceptation ( $a = 1, 2, \dots, M$ ) sont déterminées à partir de la table des Q-valeurs. Les Q-valeurs d'acceptation incluent les différentes Q-valeurs correspondant aux choix de chacun des canaux  $a$  ( $a = 1, 2, \dots, M$ ), pour servir l'appel. Si le rejet a la plus grande valeur, l'appel est alors rejeté. Dans le cas contraire, si l'une des valeurs d'acceptation a la plus grande valeur, l'appel est accepté et le canal  $a$  lui est alloué.

Nous avons fixé les paramètres de la formule (5.1) calculant le revenu immédiat à  $r_1 = +5$ ,  $r_2 = +1$  et  $r_3 = -1$ . Ce choix a comme conséquence une situation dans laquelle les canaux déjà utilisés dans les cellules compactes (et donc liées à  $r_1$ ) auront les plus grandes Q-valeurs, et vont ainsi être les candidats facilement choisis.

Le système cellulaire étudié est composé de  $N = 36$  cellules hexagonales et  $M = 70$  canaux disponibles dans un pool commun. Nous prenons une distance de réutilisation  $D = \sqrt{21}R$  ( $R$  représente le rayon de la cellule). Ceci implique que si un canal est alloué à une cellule  $i$ , alors il ne peut pas être réutilisé dans les deux rangées adjacentes à  $i$  en raison des interférences co-canaux. Ainsi, il y a au plus 18 cellules interférentes avec chaque cellule du système. La constante de propagation temporelle  $\gamma$  a été fixée à 0.5, et le taux d'apprentissage  $\alpha$  à 0.1.

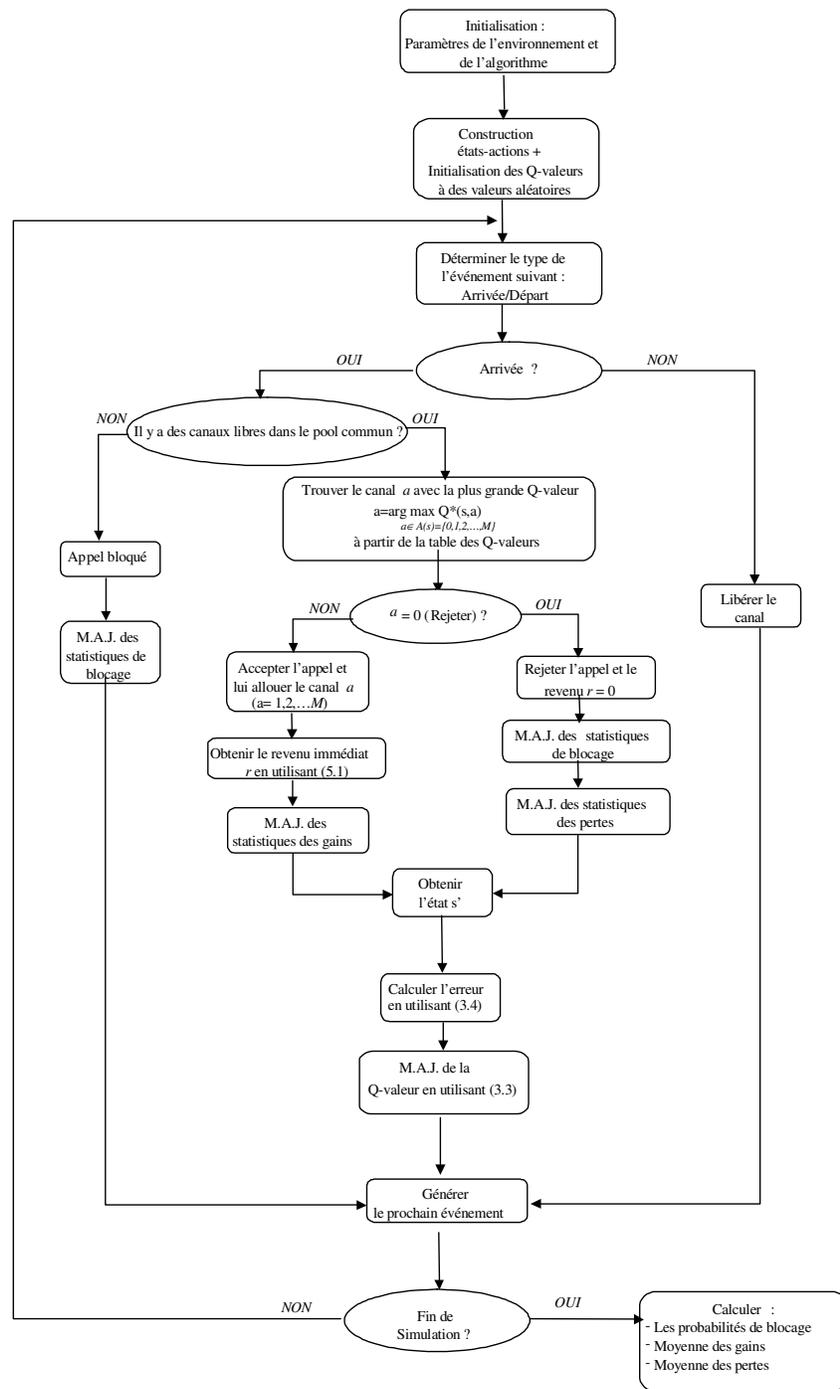


Figure 5.1 L'algorithme Q-DCA.

### 3.3 Exploration

Pour que toutes les état-action  $(s,a)$  soient explorées, l'action est choisie selon la distribution de Boltzmann définie dans le chapitre précédent.

## 4. Résultats expérimentaux

Afin d'étudier les performances de Q-DCA, un ensemble de simulations a été réalisé. Nous avons comparé Q-DCA à la politique glouton-DCA<sup>15</sup> [65], ainsi qu'à l'algorithme DCA-Nie<sup>16</sup> [56]. Les performances des algorithmes ont été évaluées en terme de gains, de pertes, ainsi que de probabilités de blocage des handoffs. Les mêmes paramètres et le même ensemble de simulations ont été utilisés, que dans le cas d'un système FCA. De plus, nous avons pu vérifier la robustesse de Q-DCA dans le cas d'une panne d'équipements.

### 4.1 Répartition uniforme du trafic

Dans cette expérience, la charge de trafic est considérée constante dans les 36 cellules, pour les deux classes de trafic. Les résultats de la Figure 5.2 montrent clairement que, malgré la charge de trafic, les probabilités de coupure diminuent de manière significative pour Q-DCA, comparées aux probabilités de coupure générées en utilisant les politiques gloutonnes. Q-DCA donne des résultats légèrement meilleurs que ceux de DCA-Nie.

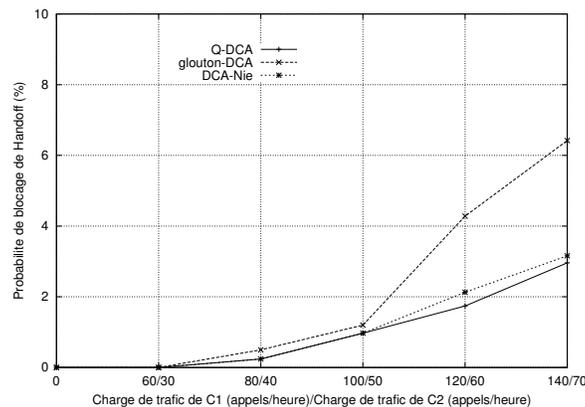


Figure 5.2 Probabilités de blocage des handoffs pour cinq différentes charges de trafic.

Les gains reçus en utilisant l'algorithme Q-DCA calculés durant une heure de simulation sont récapitulés dans la Figure 5.3(a). L'ensemble des pertes est présenté lui aussi dans la Figure 5.3(b). Nous remarquons que les politiques de contrôle d'admission et d'allocation dynamique de ressources utilisant le Q-learning (Q-DCA) sont particulièrement plus performantes que glouton-DCA, quand la charge de trafic est grande.

15. Glouton-DAC: Politique qui choisit aléatoirement un canal pour servir un appel sans aucune mesure d'interférences. Chacun des  $M$  canaux a la même probabilité d'être choisi pour servir le nouvel appel.

16. L'algorithme DCA-Nie est détaillé dans le Chapitre 3.

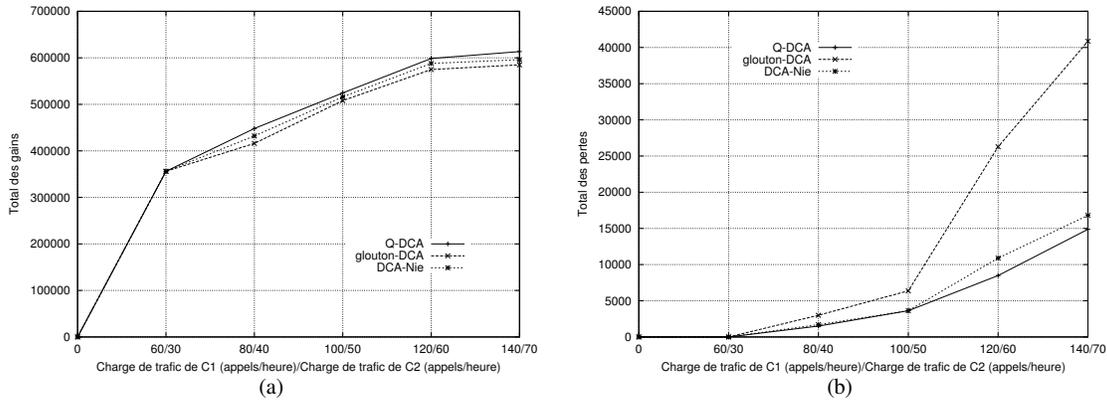


Figure 5.3 Total par heure (a) des gains (b) des pertes.

Nous avons aussi comparé les probabilités de blocage de handoff des appels de type  $C_1$  par rapport à ceux de type  $C_2$  dans les cas où l’algorithme Q-DCA serait utilisé (Figure 5.4 (a)).

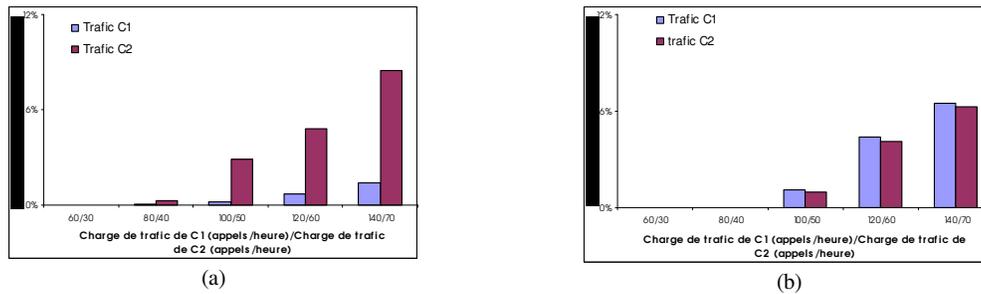


Figure 5.4 Probabilité de blocage de handoff de  $C_1$  vs.  $C_2$  en utilisant (a) Q-DCA (b) glouton-DCA.

## 4.2 Répartition non uniforme du trafic

En utilisant les mêmes politiques apprises pendant la période d’apprentissage et en considérant des charges de trafic non uniformément réparties sur les 36 cellules (cf. Figure 4.19 dans le Chapitre 4), nous avons voulu tester les capacités d’adaptation de Q-DCA. Les résultats des politiques utilisant Q-DCA indiquent des améliorations significatives par rapport aux autres politiques.

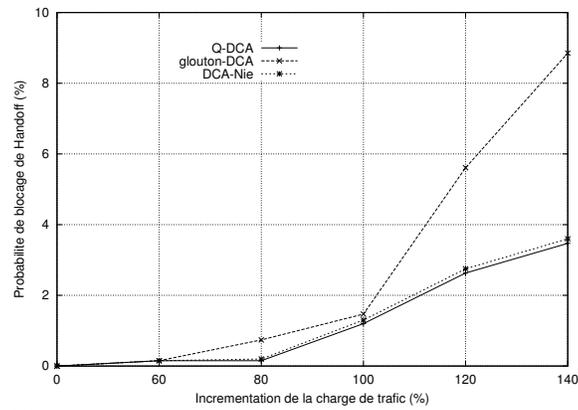


Figure 5.5 Probabilités de blocage des handoffs pour des charges de trafic différentes.

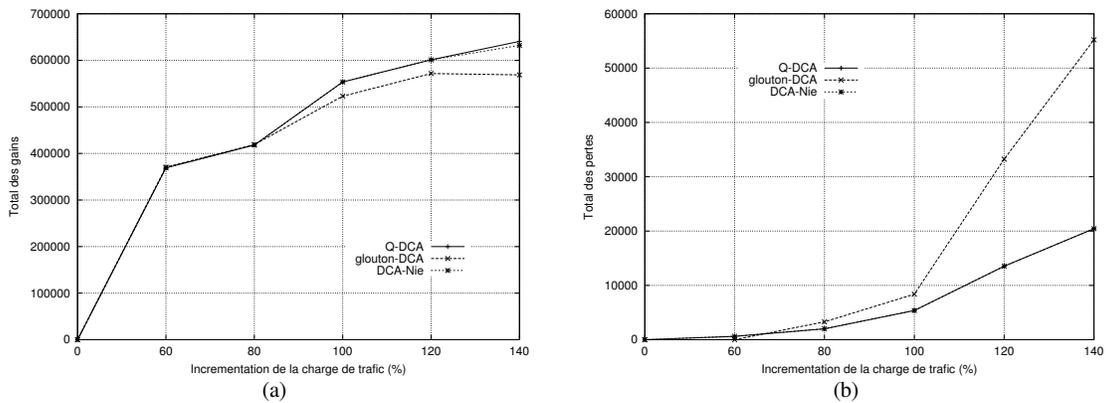


Figure 5.6 Total par heure (a) des gains (b) des pertes.

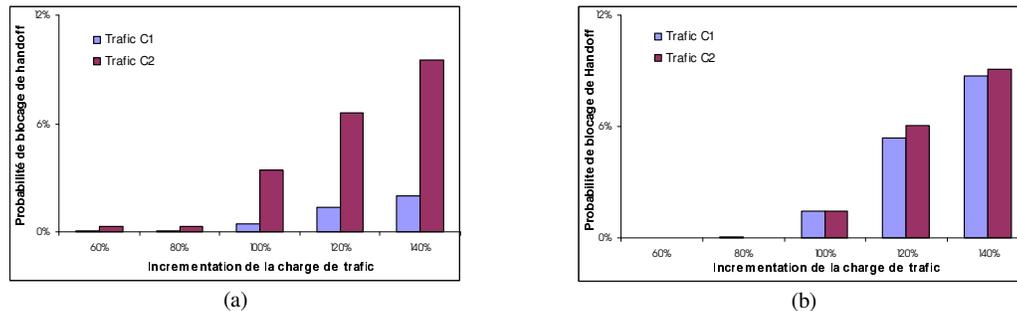


Figure 5.7 Probabilité de blocage de handoff de  $C_1$  vs.  $C_2$  en utilisant (a) Q-DCA (b) glouton-DCA.

Cet ensemble de résultats montre les possibilités qu'offre l'apprentissage par renforcement afin d'apprendre la meilleure politique d'admission et d'allocation dynamique de ressources. Q-DCA possède une aptitude à généraliser et à s'adapter aux changements dans les conditions de trafic.

### 4.3 Charge de trafic variable dans le temps

En utilisant le modèle de trafic de la Figure 4.13, nous avons voulu tester, encore une fois, les capacités d'adaptation de Q-DCA lorsque la charge du trafic change dans le temps. Les résultats des politiques utilisant Q-DCA indiquent des améliorations significatives par rapport aux autres politiques et, particulièrement, pendant les pics de trafic (à 11 heures et à 16 heures). Nous avons examiné le cas où la charge de trafic est uniformément répartie sur les 36 cellules du système (Figure 5.8(a)), ainsi que le cas où elle est non-uniformément répartie (Figure 5.8(b)).

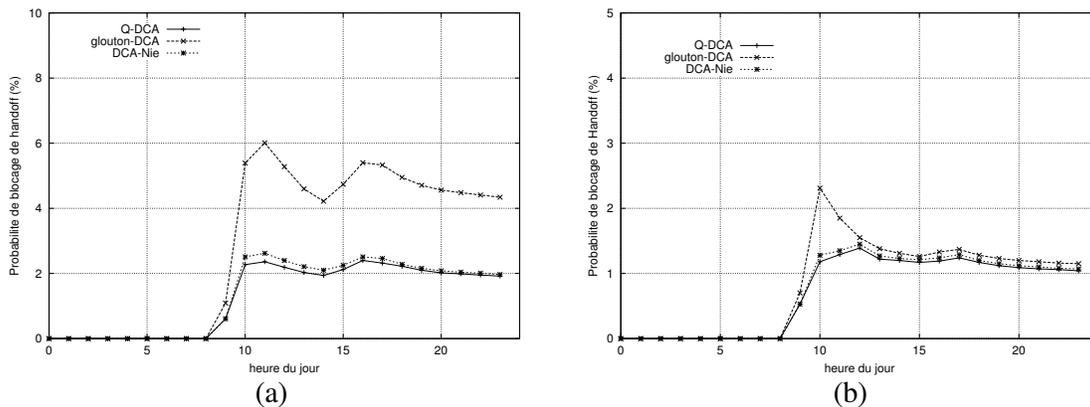


Figure 5.8 Probabilité de blocage des handoffs avec une charge de trafic variable dans le temps et un trafic (a) uniformément réparti et (b) non uniformément réparti.

### 4.4 Panne d'équipement dans un système DCA

Nous avons simulé, comme dans [56], un échec d'équipement par le fait que quelques canaux deviennent temporairement indisponibles. Au début de la simulation, il y a 70 canaux disponibles dans le système. Mais entre 10 heures et 15 heures, nous avons temporairement suspendu 0, 3, 5 et 7 canaux. La Figure 5.9 montre l'impact de l'échec des canaux sur les probabilités de coupure en utilisant l'algorithme Q-DCA. Les deux classes de trafic sont supposées uniformément réparties et utilisent les paramètres donnés dans le Tableau 4.5. Nous pouvons clairement voir que l'algorithme d'allocation dynamique de ressource Q-DCA possède une certaine robustesse contre les situations de panne d'équipements.

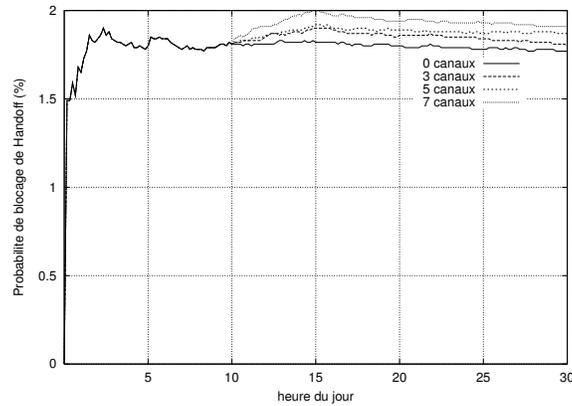


Figure 5.9 Performances de Q-DCA lors de pannes de canaux.

## 5. Conclusion

Dans ce chapitre, nous avons proposé un algorithme de recherche de la politique optimale de contrôle d'admission des appels dans un réseau cellulaire supportant plusieurs classes de trafic, utilisant notre propre stratégie d'allocation dynamique des canaux. Notre solution, appelée Q-DCA, repose sur l'apprentissage par renforcement.

Contrairement aux études que nous avons trouvées dans la littérature, notre proposition est originale, puisqu'elle combine à la fois la recherche de la politique de CAC optimale et de la meilleure stratégie d'allocation dynamique de canaux. Outre la nature intelligente de notre solution, et contrairement à d'autres travaux (études basées sur des modèles mathématiques ou sur des simulations, supposant des paramètres expérimentaux fixes), elle est adaptable aux variations de l'état du réseau (i.e. variations des conditions de trafic, pannes d'équipements, etc.). De part sa nature distribuée, elle est facilement implantable dans chaque station de base, ce qui la rend plus attrayante. Les évaluations de performances ont montré que les politiques obtenues en utilisant Q-DCA fournissaient de meilleurs résultats que les solutions classiques.

# Chapitre 6

## Réseaux ad hoc : Vue d'ensemble

### 1. Introduction

Depuis l'apparition et le grand succès commercial de la téléphonie cellulaire, le développement d'équipements mobiles n'a cessé de prendre de l'importance. Grâce aux assistants personnels et aux ordinateurs portables, l'utilisateur devient de plus en plus nomade. Dans un tel contexte, il n'est pas surprenant de voir apparaître des solutions de communication sans fil de plus en plus performantes. Les réseaux ad hoc répondent à ce besoin en éliminant le besoin d'infrastructure fixe pour communiquer. L'activité du groupe MANET de l'IETF [1] montre que le développement de ces réseaux sans fil et sans infrastructure est en plein essor. Cependant, cette absence d'infrastructure fixe pose un certain nombre de problèmes non triviaux, que nous allons discuter dans cette partie. Ce chapitre donne une vision d'ensemble sur les réseaux ad hoc, nécessaire pour la compréhension de la suite du document.

### 2. Présentation des réseaux Ad hoc

Les réseaux ad hoc sont caractérisés par un ensemble de stations, pouvant être mobiles, interconnectées et communiquant entre elles via une interface radio (uni ou bidirectionnelle). Ce type de réseaux ne nécessite aucune infrastructure pré-existante et aucune administration centralisée, formant ainsi un réseau temporaire. Chaque station peut être par ailleurs, mise à contribution par d'autres stations pour effectuer le routage de données. De ce fait, lorsqu'une station émettrice est hors de portée de la station destinatrice, la connectivité du réseau est maintenue par les stations intermédiaires. Un exemple de réseau ad hoc est illustré par la Figure

6.1 dans laquelle une station émettrice communique avec une station réceptrice en utilisant les nœuds intermédiaires.

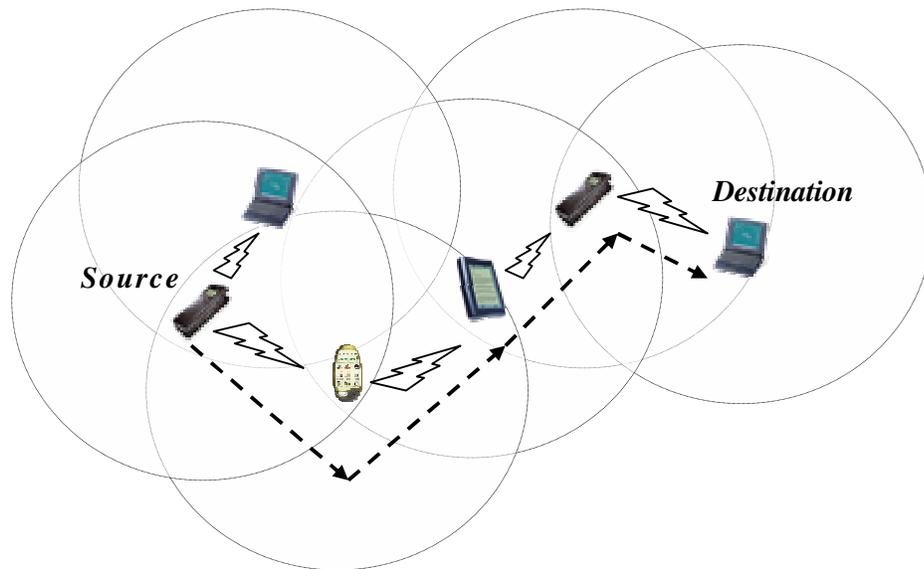


Figure 6.1 Exemple de réseau ad hoc.

(Les cercles représentent la portée radio de chaque station ad hoc)

## 2.1 Applications cibles

Les réseaux ad hoc peuvent trouver des champs d'application à travers trois catégories de scénarios :

- dès qu'il y a une absence ou une non fiabilité des infrastructures pré-existantes, c'est-à-dire typiquement lors d'opérations de secours (par exemple des incendies), de tactiques militaires (incursions en territoires ennemis) ou encore d'explorations (fouilles archéologiques) ;
- une exploitation intéressante pourrait être envisager par l'extension des réseaux filaires (par exemple dans les aéroports, les gares et les régions peu habitées, etc.) ;
- éventuellement, et à plus long terme, on peut imaginer étendre cette technologie aux applications commerciales de tout type, avec la garantie d'un coût extrêmement bas ;

Un autre exemple applicatif des réseaux ad hoc est celui que nous avons proposé pour le projet ITEA Ambience [69], que nous avons appelé "musée interactif". Les personnes visitant ce musée seront munies de PDA ou d'ordinateurs portables. Dès leurs entrées au musée, un

réseau ad hoc est mis en place. L'objectif principal de ce réseau est de véhiculer des informations utiles aux visiteurs du musée. Ces informations peuvent être de différentes formes (texte, image, vidéo, audio) et vont guider les visiteurs dans leurs visite du musée, et leurs donner des explications sur les différentes œuvres exposées. Le musée dispose, aussi, d'un certain nombre de robots participant à la mise en place du réseau ad hoc. Ces robots interviennent comme routeurs mobiles et permettent de donner accès au réseau d'infrastructure. Ils permettent, également, de préserver la connectivité du réseau, en se plaçant à des endroits particuliers du musée (endroits fortement chargés par le trafic, groupe de visiteurs non joignables directement par un simple saut).

## 2.2 Avantages

Les avantages de cette technologie sont nombreux du fait qu'il n'y a pas besoin d'infrastructure pré-existante :

- les réseaux ad hoc peuvent être déployés dans un environnement quelconque ;
- le coût d'exploitation du réseau est nul : aucune infrastructure n'est à mettre en place initialement et surtout aucun entretien n'est à prévoir ;
- le déploiement d'un réseau ad hoc est : (i) simple et ne nécessite aucun pré requis puisqu'il suffit de disposer d'un certain nombre de terminaux dans un espace pour créer un réseau ad hoc, et (ii) rapide puisqu'il est immédiatement fonctionnel dès lors que les terminaux sont présents ;
- la souplesse d'utilisation est importante puisque les seuls éléments pouvant tombés en panne sont les terminaux eux mêmes. Autrement dit, il n'y a pas de panne "pénalisante" de manière globale (une station qui sert au routage peut être remplacée par une autre si elle tombe en panne).

## 2.3 Inconvénients

Même si les perspectives pour les réseaux ad hoc sont prometteuses, plusieurs contraintes restent encore à traiter :

- la connectivité limite les possibilités de communication. Ainsi, deux stations ne sont joignables que s'il existe un ensemble de stations pouvant assumer la fonction de routeur afin de faire suivre les paquets de données échangées entre les deux stations.

Dans l'architecture filaire, les possibilités de communication sont prévisibles avant sa mise en place et les bornes d'accès d'une architecture cellulaire telle que GSM ou UMTS permettent de manière similaire de connaître avec exactitude les zones de couverture (cf. Chapitre 2). Ce n'est plus le cas avec les réseaux ad hoc où une communication n'est possible que si la collaboration entre stations est suffisante pour lier l'émetteur au récepteur ;

- les liens entre les stations ne sont pas isolés les uns des autres et polluent le voisinage, par diffusion, lors de chaque émission/réception de données. Par conséquent, tout paquet de diffusion émis vers une station en cours de communication (que le paquet lui soit destiné ou pas) va altérer la communication de cette station. La diffusion est un facteur qui alourdit aussi d'autres paramètres tels que la bande passante et la consommation de batterie ;
- la sécurité dans les réseaux ad hoc est difficile à contrôler, notamment parce que sur l'interface air l'écoute clandestine est très simple à réaliser ;
- l'absence de centralisation rend les stations toutes semblables ; il devient alors difficile d'adopter des politiques de gestion globale du réseau. En effet, mettre en place un système de facturation est techniquement délicat, et offrir des qualités de service (QoS) différentes aux utilisateurs est aussi difficilement contrôlable dans ce contexte ;

<b>Proposition [70] :</b>	La tendance actuelle pour résoudre ce genre de problèmes dans les réseaux fixes, est réalisée par l'intermédiaire d'un modèle par politiques [71]. Ce modèle repose sur l'utilisation d'un serveur de politique (PDP - Policy Decision Point) pour mettre en place, gérer et configurer les différents services offerts par le réseau. Nous proposons dans [70] d'adapter ce modèle par politique pour les réseaux ad hoc afin de résoudre ces problèmes de sécurité et de QoS. Cette adaptation est nécessaire vu que ce type de réseau ne repose sur aucune infrastructure et n'est composé que d'équipements mobiles, contrairement aux réseaux fixes qui reposent eux sur une infrastructure et qui contiennent des serveurs dédiés. Notre idée consiste à proposer un mécanisme permettant l'élection du meilleur terminal pouvant supporter le rôle de serveur de politiques. Cette élection prend en considération plusieurs critères liés aux capacités des terminaux mobiles. La solution proposée est une solution robuste traitant tous les cas de figures qui peuvent survenir dans un réseau ad hoc
---------------------------	--

	(contention entre deux mobiles pour le rôle de serveur, le serveur quitte le réseau, transfert du serveur vers un autre terminal présentant de meilleures capacités, etc.). Cette proposition ne rentre pas dans le contexte de cette thèse. Nous invitons les lecteurs intéressés, à se référer à l'article [70] pour plus de détails.
--	---

- L'absence totale d'administration centralisée rend complexe et coûteuse l'utilisation des techniques de multiplexage des communications utilisées dans les réseaux avec point d'accès (FDMA, TDMA, etc.) définies dans le Chapitre 2. En effet, pour utiliser ce type de mécanismes, il faut d'une part, concevoir un protocole distribué permettant aux nœuds de se partager les fréquences, les codes ou les unités de temps. D'autre part, la mobilité des nœuds d'un tel réseau provoque des changements fréquents de topologie. Le routage étant aussi distribué, les routeurs sont les mobiles et un transfert de données peut aisément être interrompu par le départ d'un nœud de la route utilisée.
- enfin, la faible autonomie des batteries constitue un frein à une utilisation longue du terminal et à la mise en place de nouveaux services. C'est une contrainte qui existe certes dans les réseaux de type GSM ou UMTS, mais qui est plus forte dans les réseaux ad hoc, puisque les ressources énergétiques sont mises en commun même pour les besoins du routage. Nous nous intéressons dans cette thèse, plus spécialement, à ce dernier point. Nous proposons, dans le Chapitre 8, des solutions permettant de mieux gérer cette consommation des batteries.

### 3. Principaux algorithmes de routage ad hoc

Globalement, on distingue deux familles de protocoles de routage ad hoc : les protocoles de routage dits "proactifs", qui anticipent la demande d'acheminement de paquets et les protocoles de routage "réactifs" qui réagissent à la demande. Entre ces deux familles, une autre approche commence à émerger, il s'agit des protocoles dits "hybrides" qui s'inspirent à la fois des protocoles proactifs et des protocoles réactifs. La liste des protocoles de routage qui suit est loin d'être exhaustive ; il en existe bien d'autre mais cette sélection couvre les protocoles les plus classiques et les plus étudiés. Nous donnons, ci-après, une vue globale de ces protocoles et de leurs caractéristiques essentielles.

### 3.1 Protocoles proactifs

Les protocoles de routage proactifs pour les réseaux mobiles ad hoc, sont basés sur le même principe des protocoles de routage utilisés dans les réseaux filaires. Les deux principales méthodes utilisées sont la méthode "état de lien" (ou Link State) et la méthode "vecteur de distance" (ou Distance Vector). Ces deux méthodes exigent une mise à jour périodique des données de routage qui doit être diffusée par les différents nœuds de routage du réseau. Nous allons décrire dans ce qui suit, les protocoles les plus importants de cette classe LSR, OLSR et DSDV.

#### 3.1.1 Le protocole LSR

Dans le protocole LSR (Link State Routing) [72], chaque station diffuse périodiquement à ses voisines l'état de ses liens. Ceux-ci diffusent à leur tour, et de proche en proche, de manière récursive les informations qui leurs parviennent, jusqu'à les faire converger pour qu'elles soient connues de toutes les stations. De cette manière, chaque station va pouvoir constituer ainsi sa propre table de routage, qui va être utilisée lorsque la station souhaitera joindre un destinataire : une simple recherche dans la table suffira pour localiser le récepteur. Ce protocole illustre parfaitement le concept de routage proactif, et cumule les défauts inhérents à cette technologie (une diffusion parfois excessive des données de routage, et un gaspillage de la bande passante). En faible mobilité, ce protocole fournit de bons résultats, mais qui s'affaiblissent progressivement quand la mobilité des stations augmente.

#### 3.1.2 Le protocole OLSR

Comme son nom l'indique, OLSR (Optimized Link State Routing) [73], optimise le protocole LSR en réduisant les diffusions grâce à la notion de nœuds multi-points (ou MPR - Multi-Point Relay). Les MPR sont des nœuds élus par chaque station de manière à ce que tout voisin de cette station soit joignable en un maximum de deux sauts à travers les nœuds MPR. Chaque nœud émet périodiquement la liste de ces voisins, mais seuls les voisins MPR vont diffuser cette liste à leur tour pour minimiser les diffusions. Autrement dit, seul les nœuds MPR ont la connaissance de la topologie du réseau, et peuvent assumer le rôle de routeur. Les autres stations ayant pour seules possibilités de diffuser vers leurs voisins MPR. Globalement, ce protocole améliore réellement LSR en évitant l'inondation totale du réseau.

### 3.1.3 Le protocole DSDV

Le protocole DSDV (Dynamic destination-Sequenced Distance Vector) [74] se base sur l'algorithme distribué de Bellman-Ford (DBF), qui utilise les vecteurs de distance. Chaque station maintient une table de routage contenant toutes les destinations qu'elle peut atteindre et le coût (en nombre de saut) pour atteindre la destination, ainsi qu'un numéro de séquence lié à chaque destination dont le but est d'éviter la formation de boucle de routage. Cette table est constituée par l'intégration des données de mise à jour émises par chaque station. Ces mises à jour s'effectuent en fonction du temps ou bien en fonction d'événements liés à une modification de la topologie du réseau (lien rompu, nouvelle station, etc.). Elles se font soit de manière incrémentale (les seules données qui ont changé par rapport à la dernière mise à jour), soit intégralement (la table toute entière), ceci selon l'importance des modifications constatées.

## 3.2 Protocoles réactifs

Les protocoles de routage appartenant à cette catégorie, créent et maintiennent les routes selon les besoins. Lorsque le réseau a besoin d'une route, une procédure de découverte de routes est lancée. Dans ce qui suit, nous allons décrire les protocoles les plus importants de cette classe : DSR et AODV. Nous nous intéressons particulièrement, au protocole AODV puisqu'il fait l'objet d'un ensemble d'extensions abordé dans le Chapitre 8. Nous présentons également un protocole réactif multi-chemins appelé AOMDV, utilisé comme base de l'une de nos propositions.

### 3.2.1 Le protocole DSR

Le principe du protocole DSR (Dynamic Source Routing) [75], se décompose en deux étapes complémentaires : "découverte de route" et "maintien de la route". La première étape de découverte de route se déroule comme suit : la source émet une requête RREQ (Route Request), contenant un champ 'enregistrement de route', en diffusion à ses voisins. Ces derniers diffusent à leur tour cette requête, en complétant le champ 'enregistrement de route' de leur propre identifiant, jusqu'à atteindre la destination cible. Le nœud destinataire répond alors par un paquet de réponse RREP (Route Reply), en diffusion récursive vers la source. Ce paquet de réponse contient le champ 'enregistrement de route' servant aux nœuds intermédiaires pour le routage et indiquant un chemin à la source.

Suit alors l'étape de maintien de route qui se déroule comme suit : si une station route un message en utilisant une liaison devenue non valide, la station émet un message d'erreur de

route RERR (Route Error), vers les nœuds qui sont à l'origine du message. De cette façon, tous les récepteurs retirent le lien erroné de tous les chemins utilisés, et effectuent de nouvelles découvertes de route pour chacun de ces chemins.

Une judicieuse exploitation des messages RREQ et RREP permet la constitution de caches dans les stations recevant de tels messages qu'elles soient source, destinataire ou simplement intermédiaire. Ceci permet non seulement d'éviter d'avoir à effectuer une nouvelle découverte de route tant que la route reste valide, mais fournit également aux nœuds intermédiaires des routes potentiels. Globalement, DSR est un protocole simple et robuste, qui réduit considérablement la surcharge et fournit différents chemins possibles. On peut cependant lui reprocher l'accroissement de la taille des paquets par la donnée du chemin, posant ainsi un problème de passage à l'échelle (scalabilité).

### 3.2.2 Le protocole AODV

L'algorithme AODV (Ad hoc On demand Distance Vector) [76], reprend le protocole DSR en le modifiant afin d'éviter d'avoir à transporter la route à suivre dans tous les paquets émis. Pour cela, chaque nœud tient à jour une table de routage spécifiant pour chaque requête la source, la destination et le nœud précédent constitué durant le message de requête RREQ. Cette dernière information est utilisée, par lecture inversée, lors du retour par le message de réponse RREP. Comme pour DSR, les tables de routage peuvent servir de caches de façon à ce qu'un nœud intermédiaire qui connaît la destination réponde à la source directement. AODV utilise en outre un numéro de séquence pour comparer les chemins les plus récents, afin de forcer les mises à jour si nécessaires et d'éviter la formation de boucles de routage. L'utilisation de 'timeout' permet de purger les entrées de la table de routage pour forcer une mise à jour de chemins jugés non valides avec la mobilité. Par ailleurs, des messages "Hello" sont périodiquement émis pour s'assurer de la validité des liens (ce contrôle pouvant être couplé avec les acquittements de la couche MAC) de façon à émettre un message d'erreur RERR en l'absence de réponse à ce message, signalant la rupture du lien. La découverte et le maintien de la route dans AODV sont détaillés ci-dessous.

### 3.2.3 Découverte de route

AODV construit les routes avec l'utilisation d'un cycle de requêtes "route request / route reply". Lorsqu'un nœud source désire établir une route vers une destination pour laquelle il ne possède pas encore de route, il diffuse un paquet RREQ à travers le réseau. Les nœuds recevant le paquet mettent à jour leur information relative à la source, et établissent des pointeurs de

retour vers la source dans les tables de routage. Outre l'adresse IP de la source, le numéro de séquence courant et l'identificateur (ID) de diffusion, le RREQ contient également le numéro de séquence de la destination le plus récent connu de la source. Un nœud recevant un RREQ émettra un paquet RREP (route reply) soit s'il est la destination, soit s'il possède une route vers la destination avec un numéro de séquence supérieur ou égal à celui repris dans le RREQ. Si tel est le cas, il envoie (en unicast) un paquet RREP vers la source. Sinon, il rediffuse le RREQ. Chaque nœud conserve une trace des adresses IP sources et des identificateurs de diffusion des RREQ. S'ils reçoivent un RREQ qu'ils ont déjà traité, ils l'écartent et ne le transmettent pas.

Alors que les paquets de réponses RREP reviennent vers la source, les nœuds intermédiaires construisent des pointeurs vers la destination dans leurs tables de routage. Une fois que la source a reçu le RREP, elle peut commencer à émettre des paquets de données vers la destination. Si ultérieurement, la source reçoit un RREP contenant un numéro de séquence supérieur ou le même mais avec un nombre de sauts plus petit, elle mettra à jour son information de routage vers cette destination et commencera à utiliser la nouvelle route.

### **3.2.4 Maintien de la route**

Une route est maintenue aussi longtemps qu'elle continue d'être active. Elle est considérée active tant que des paquets de données y transitent périodiquement. Lorsque la source arrêtera d'émettre des paquets de données, le lien expirera et sera alors effacé des tables de routage des nœuds intermédiaires. Si un lien se rompt alors qu'une route est active, le nœud extrémité du lien rompu émet un paquet d'erreur RERR vers le nœud source, pour lui notifier que la destination est désormais injoignable. Après réception de RERR, si la source souhaite toujours émettre, elle peut re-initier un processus de découverte de route.

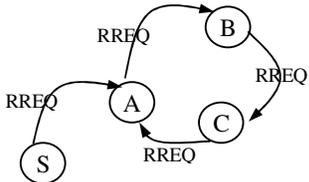
### **3.2.5 Le protocole AOMDV**

L'objectif primaire du protocole réactif AOMDV (Ad hoc On demand Multipath Distance Vector) [77] est de fournir une tolérance efficace aux fautes, dans le sens où le rétablissement en cas d'échec d'une route est plus rapide. Pour réaliser cet objectif, AOMDV construit plusieurs routes sans boucles de routage allant de la source jusqu'à la destination; contrairement à AODV qui construit et maintient une seule route seulement pour chaque couple (source, destination). Les protocoles de routage à la demande multi-routes, telles que AOMDV, essaient d'atténuer la latence élevée de la procédure de découverte de route pouvant affecter les performances défavorablement. AOMDV calculent de multiples chemins lors d'une simple tentative de découverte de route. Ces multiples chemins vont être formés aussi bien dans les

sources de trafic que dans les nœuds intermédiaires. Une nouvelle découverte de route n'est nécessaire, que si toutes les routes deviennent non valides. Ceci réduit considérablement les latences et la surcharge de la procédure de découverte. Dans ce qui suit, nous allons décrire la découverte et le maintien de la route dans AOMDV.

### 3.2.6 Découverte de route

L'idée principale dans AOMDV consiste à calculer différentes routes, allant de la source de trafic jusqu'à la destination, tout en évitant la formation de boucles de routage. Notons que dans AODV, chaque requête (respectivement réponse) reçue par un nœud pendant le processus de découverte de route définit potentiellement une route alternative vers la source (respectivement destination). Par exemple, chaque copie du paquet de requête RREQ arrivant à un nœud définit une route alternative vers la source. Cependant, accepter naïvement toutes ces copies va mener à la formation de boucles dans les routes.

<p><b>Exemple de formation de boucles :</b></p>	<p>Pour voir comment les boucles peuvent se produire, considérez le simple exemple ci-contre : la source <i>S</i> diffuse des paquets de requêtes RREQ, qui seront interceptés par le nœud intermédiaire <i>A</i>. Le nœud <i>A</i> le diffuse à nouveau aux nœuds voisins (nœud <i>B</i> incluse). Le nœud <i>B</i>, envoie cette requête au nœud <i>C</i>. Ce dernier qui est un voisin du nœud <i>A</i>, va rediffuser cette requête, qui sera entendue par le nœud <i>A</i> aussi. Si le nœud <i>A</i> accepte cette copie de RREQ, une boucle va se former (cf. la figure ci-contre).</p>	
---	--	--

Afin d'éliminer la possibilité de formation de boucles, les auteurs proposent de maintenir la même condition que dans AODV, et qui consiste à n'accepter une nouvelle requête que si : (i) le nombre de sauts est inférieur ou que (ii) le numéro de séquence est supérieur.

La construction de différentes routes dans AOMDV est basée sur la notion de "advertised\_hopcount". Pour chaque nœud sur la route, ce paramètre représente le nombre de sauts qui ont été nécessaires à la première requête pour atteindre ce nœud. Ce nombre de sauts est alors considéré comme l'advertised\_hopcount, et ne peut changer pour le même numéro de séquence. AOMDV n'acceptera par la suite, que les routes alternatives dont le nombre de sauts

est inférieur ou égal à l'advertised\_hopcount. Cette condition est suffisante pour garantir la non formation de boucles.

destination
numéro de séquence
hopcount
nexthop
expiration_timeout

destination
numéro de séquence
advertised_hopcount
liste_de_routes {(nexthop <sub>1</sub> , hopcount <sub>1</sub> ), (nexthop <sub>2</sub> , hopcount <sub>2</sub> ),..}
expiration_timeout

(a) AODV
(b) AOMDV

Figure 6.2 Structure des tables de routage dans AODV et AOMDV.

La Figure 6.2 montre la structure des tables de routage pour AODV et AOMDV. Dans AOMDV, l'advertised\_hopcount remplace le nombre de sauts (hopcount) de AODV. Une liste appelée "liste\_de\_routes" remplace le prochain saut (nexthop). Cette liste définit pour chaque prochain saut (nexthop<sub>k</sub>), le nombre de sauts (hopcount<sub>k</sub>) nécessaire pour atteindre la destination en passant par ce nœud. Notons que tous les prochains sauts pour une même destination disposent du même numéro de séquence. Notons également, que le paramètre *advertised\_hopcount* est initialisé à chaque fois que le numéro de séquence est mis à jour.

### 3.2.7 Maintien de la route

L'idée principale dans AOMDV est de calculer différents chemins pendant la phase de découverte de route. Ainsi, une nouvelle découverte de route est nécessaire uniquement lorsque aucune des routes précédemment établies n'est valide.

Lorsque aucune route vers la destination n'est valide, le nœud envoie un paquet d'erreur RERR à tous ses voisins qui utilisent ce nœud comme prochain saut dans la route vers cette destination. Ces routes sont effacées des tables de routage des nœuds récepteurs de ce paquet d'erreur. En revanche, contrairement au protocole AODV, ces paquets d'erreur ne sont pas relayés jusqu'à la source. Si la source de trafic reçoit un paquet RERR, elle lance une nouvelle découverte de route dans le cas où elle souhaite toujours émettre.

## 3.3 Protocoles hybrides

Le protocole ZRP ( Zone Routing Protocol ) [78] est un exemple de protocole hybride, à mi-chemin entre les deux familles de protocoles (proactif et réactif). Ainsi, chaque nœud maintient une table de routage, dont les données sont régulièrement émises en diffusion pour

tous les nœuds qui lui sont distants de moins d'une valeur  $d$  prédéfinie (routage proactif dans cette zone). Pour atteindre tout autre nœud qui n'apparaîtrait pas dans sa table de routage (une distance supérieure à  $d$ ), un nœud a recours à un protocole de routage de type réactif similaire au protocole DSR. Ce type de protocole fournit un assez bon compromis en terme de diffusion pour les mises à jour. Cette tentative pour cumuler les qualités des deux approches se place en intermédiaire plus qu'en solution, parce qu'elle est moins efficace que les algorithmes de routage de base, en forte mobilité ou avec beaucoup de stations.

## 4. Conclusion

Les principaux défis à relever dans les réseaux ad hoc sont fondamentalement liés au problème de routage. Nous avons vu, dans la section 3 de ce chapitre, que plusieurs protocoles de routage ont été conçus. Cependant, très vraisemblablement, le protocole de routage optimal dans toutes les situations n'existe pas. En effet, les études de performance de ces algorithmes montrent que leurs performances sont souvent moyennes. Les meilleurs résultats n'étant atteignables que dans certaines situations bien précises. Afin de choisir le meilleur protocole de routage le plus approprié, il convient de définir les conditions d'utilisation (mobilité, nombre de stations, QoS, etc.) souhaitées, et de choisir ensuite l'algorithme qui satisfera au mieux ces conditions. La priorité principale étant de garder les terminaux mobiles en marche. Dans le chapitre suivant, nous allons traiter ce dernier aspect, souvent considéré comme étant l'un plus des grands challenges pour les réseaux ad hoc.

# Chapitre 7

## Consommation de l'énergie dans les réseaux ad hoc

### 1. Motivations

L'objectif pour un réseau ad hoc est que les terminaux mobiles soient utilisés au maximum n'importe où et n'importe quand "any-where and anytime". Cependant, l'une des grandes limitations de cet objectif concerne le support énergétique. En effet, la principale contrainte dans les communications sans fil est la durée de vie limitée des terminaux mobiles dont le support énergétique représente souvent une batterie dont la capacité est limitée.

Nous donnons, ci-dessous, quelques exemples des principaux facteurs affectant la consommation de batteries dans un réseau sans fil :

- *la transmission radio* : la consommation la plus importante est due à la transmission radio. Il convient de noter qu'une émission est davantage coûteuse qu'une réception et que comparativement, une veille est la moins coûteuse ;
- *gestion du terminal* : une autre utilisation des batteries est due à l'alimentation des ressources propres au terminal, qu'il s'agisse de CPU, de disque dur, des mémoires, d'affichage, etc. ;
- *fonctionnalités des protocoles permettant d'assurer la liaison entre les stations pendant une communication* : nous citons, par exemple, les protocoles de contrôle de collision, de congestion, de routage ;

- *les applications* : les applications qui commencent à émerger pour les réseaux ad hoc consomment également les ressources de la batterie et plus spécialement les applications utilisant des calculs pour la compressions ou le cryptage de données, par exemple.

Plusieurs études ont montré que les plus grands facteurs consommant la batterie d'un ordinateur portable sont le microprocessor (CPU), l'écran à cristaux liquides (LCD), le disque dur, le lecteur de CDROM, le lecteur de disquette, le système d'E/S, et la carte réseau sans fil [79]. La conservation de l'énergie a donc, été pendant longtemps traitée au niveau de la couche physique. Cependant, nous pouvons remarquer, dans la liste énumérée ci-dessus, qu'il existe d'autres facteurs, en dehors de la couche physique, pouvant affecter la consommation de l'énergie des batteries. Ainsi, la conception de nouveaux protocoles de communication réseaux devrait aider à réduire cette consommation. La section 2 donne un aperçu de ce qui se fait dans ce domaine. Nous nous concentrons, dans la section 3, sur la couche réseau et nous allons prouver l'impact du choix du protocole de routage sur la consommation de l'énergie dans les réseaux ad hoc.

## 2. Protocoles minimisant la consommation des batteries

Chacune des couches du modèle de référence a fait l'objet de propositions pour réduire la consommation des batteries selon différentes approches [80]. Notons que le problème de consommation de l'énergie ne peut être traité séparément dans une des couches protocolaires. Nous allons exposer ici quelques propositions parmi les plus intéressantes concernant ces différentes couches.

### 2.1.1 Couche Physique

Dans le passé, les recherches sur la consommation de l'énergie se sont concentrées sur la couche physique étant donné que la consommation de l'énergie dans un ordinateur portable était considérée comme origine directe du matériel utilisé. Ces travaux concernent l'augmentation de la capacité de la batterie, une fréquence d'horloge variable pour le CPU, l'utilisation des mémoires flash, etc. Notons qu'à la différence d'autres domaines de l'informatique, la technologie de batterie n'a pas subi un avancement significatif au cours des 30 dernières années.

### 2.1.2 Sous-couche MAC

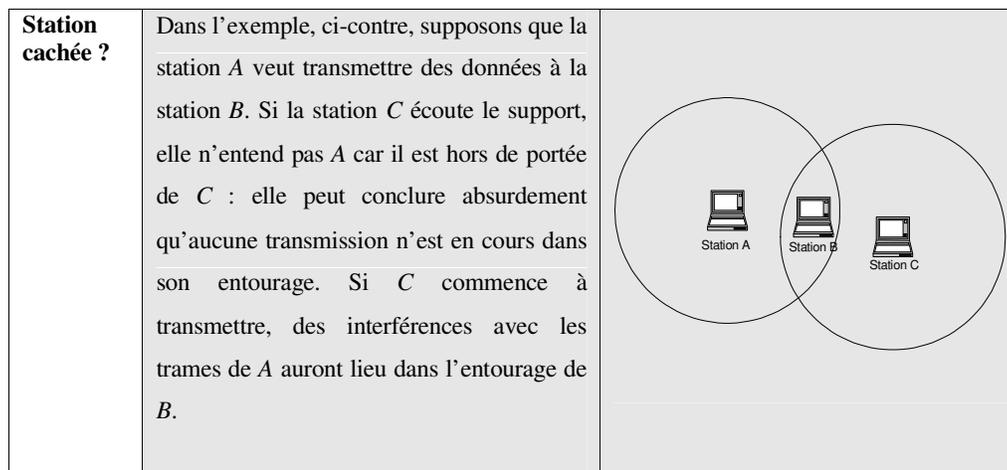
La couche MAC (Medium Access Control) est une sous-couche de la couche liaison de données. Cette couche fait l'interface avec la couche physique et comprend des protocoles

définissant la manière d'allouer les canaux radios partagés entre l'ensemble des nœuds mobiles. L'objectif de la couche MAC consiste à éliminer, autant que possible, les collisions puisque les collisions provoquent des retransmissions, et les retransmissions mènent à une consommation inutile de l'énergie. Notons que, dans un réseau sans fil, les retransmissions ne peuvent pas être complètement évitées à cause du taux d'erreurs très élevés et de la mobilité des nœuds.

Il existe plusieurs protocoles MAC : IEEE 802.11, EC-MAC [81] et PAMAS [82]. Nous choisissons de présenter le protocole PAMAS parce que c'est un protocole conçu spécialement dans le but de réduire la consommation de l'énergie dans les réseaux ad hoc.

**A. Le protocole PAMAS**

Le protocole PAMAS (Power Aware Multi-Access Protocol with Signalling) [82], est un protocole à accès multiples au médium radio basé sur le protocole MACA (Multiple Access Collision Avoidance) [83] permettant de palier aux problèmes de station cachée.



PAMAS rajoute par rapport à MACA, l'utilisation de deux canaux distincts : (i) l'un utilisé pour la signalisation, transportant donc les messages de contrôle RTS/CTS<sup>17</sup> ; (ii) l'autre utilisé pour le transport des données. Ainsi, toute station ne peut perturber que le canal de signalisation, n'affectant pas le transport des données. Bien sûr, les messages RTS peuvent toujours être soumis aux problèmes de collision mais leur taille est relativement courte et, de manière générale, la perte subie en demandant une re-émission ultérieure d'un tel paquet est nettement inférieure à celle subie pour la re-émission d'un paquet de données.

---

17. Request ToSend/Clear To Send.

Par ailleurs, le principal apport de PAMAS est la conservation des batteries des nœuds. Ceci est réalisé grâce à une optimisation de la couche MAC, en permettant à toutes les stations d'éteindre leur interface radio afin de réduire leur consommation d'énergie, à chaque fois que leur interface radio n'est pas utilisable. En effet, le protocole se base sur le constat qu'une station ne peut recevoir de message tant qu'elle écoute une autre communication, et ne peut pas émettre non plus si l'une de ses voisines est en train de recevoir des messages (pour causes d'interférences). Toute station pourra éteindre son interface radio (s'endormir) au cas où : (i) elle ne veut pas émettre et au moins un voisin est en train d'émettre vers une autre station ; (ii) elle veut émettre et au moins un voisin est en train de recevoir ; (iii) elle veut émettre mais tous ces voisins sont déjà en train d'émettre. Pour déterminer la durée pendant laquelle la station peut rester éteinte, il faut que le paquet RTS indique la longueur du message de façon à ce que le paquet CTS le diffuse au voisinage, qui en est ainsi avertie. Un problème se pose cependant lorsque les stations se réveillent et trouvent le canal occupé, ignorant la durée de cette nouvelle communication et surtout augmentant la probabilité de collision des messages RTS. Une solution à ce problème est apportée en utilisant un algorithme de backoff. Ainsi, une station, avant toute émission, se voit affecté un quota de temps (différent pour chaque station) au bout duquel elle est autorisée à émettre son paquet de requête RTS. La station dont le délai, avant d'émettre, est le plus court va donc pouvoir contacter sa cible, qui lui répondra par diffusion que le canal est occupé. Cette réponse va être entendue de tous les voisins directs de la cible qui seront ainsi avertis du message et par conséquent, abandonneront le processus d'émission de leur paquet RTS. Ce mécanisme va donc permettre de disperser dans le temps les messages RTS pour réduire les possibilités de collisions des stations qui se réveillent en utilisant PAMAS. PAMAS est donc un bon protocole réduisant les collisions, et minimisant la consommation des batteries. En outre, son intégration aux protocoles de routage classique est assez aisée.

### **2.1.3 Sous-couche LLC**

Dans ce paragraphe, nous nous intéressons aux fonctionnalités de contrôle d'erreur de la sous-couche de la couche liaison de données : LLC (Logical Link Control). Les deux techniques les plus utilisées pour le contrôle d'erreur, dans un environnement filaire, sont ARQ (Automatic Repeat Request) et FEC (Forward Error Correction). Ces méthodes consomment beaucoup d'énergie en raison des retransmissions et de la surcharge nécessaires pour la correction d'erreurs. Utiliser ce genre de protocoles, dans un environnement sans fil, nécessite une grande adaptation vu le taux d'erreur qui est plus élevé (bruit, évanouissement du signal, mobilité, etc.). Quelques travaux récents ont proposé des algorithmes d'adaptation de ces

méthodes de contrôle d'erreurs permettant également de réduire la consommation de l'énergie [84][85].

Dans [84] par exemple, les auteurs proposent d'incorporer un protocole de sondage (probing) permettant de ralentir la transmission des données dès lors que l'état du canal est dégradé. Ainsi, le protocole ARQ est utilisé normalement jusqu'à ce que l'émetteur détecte une erreur sur les données ou sur le canal de contrôle qui est due à la non réception d'acquittement (ACK). À cet instant, le protocole passe en mode sondage dans lequel un paquet de sondage (probe) est envoyé chaque  $t$  slots. Le paquet de sondage contient seulement un en-tête avec peu ou pas de charge utile et consomme donc peu d'énergie. Ce mode est adopté jusqu'à ce qu'un acquittement ACK soit correctement reçu. Le protocole alors revient en mode normale et continue la transmission des données à partir du point auquel il a été interrompu.

#### 2.1.4 Couche réseau

Les protocoles de couche réseau composent la plus grande classe de protocoles d'économie d'énergie. Ils diffèrent selon le type du trafic : (i) trafic en unicast, ou (ii) trafic de diffusion.

##### 2.1.4.1 Trafic de diffusion

La diffusion est une importante opération dans les réseaux et plus spécialement dans réseaux ad hoc en raison de la mobilité des nœuds. Dans ces réseaux, les diffusions chargent considérablement le réseau en plus d'accroître les collisions, qui entraînent ensuite des rediffusions. C'est une source de consommation des batteries qui peut vite être importante. Des solutions à ce problème ont été suggérées, en utilisant différentes stratégies possibles [86] :

- des considérations probabilistes : la première fois qu'un message est reçu il est diffusé avec une probabilités  $p$ , en considérant qu'une absence de diffusion peut être compensée par les voisins, si la valeur de  $p$  est bien adaptée ;
- des considérations de comptage : ne pas diffuser un paquet s'il a déjà été diffusé. Cette solution est simple et évite les redondances mais elle n'est pas toujours efficace ;
- ou encore des considérations de proximité : ne pas diffuser un paquet, s'il vient d'une station proche d'une distance  $d$ , puisqu'une diffusion aura environ la même portée.

Ces solutions sont approximatives, mais les simulations en ont montré l'intérêt.

### 2.1.4.2 Trafic unicast

La principale métrique utilisée dans les protocoles de routage ad hoc traditionnels est le nombre minimum de sauts. Cependant, cette métrique a un effet négatif sur la consommation de l'énergie de certains nœuds. Par exemple, dans la Figure 7.1, le trafic allant du nœud *A* au nœud *D* va toujours passer par le nœud *E*, ce qui va provoquer la consommation de toute la batterie de ce dernier. Et si tel est le cas, le nœud *F* devient injoignable et le réseau devient partitionné. L'utilisation d'un algorithme de routage qui tient compte de la métrique énergétique va, au contraire, utiliser une autre route que celle passant par le nœud *E* (par exemple, la route passant par les nœuds *B* et *C*).

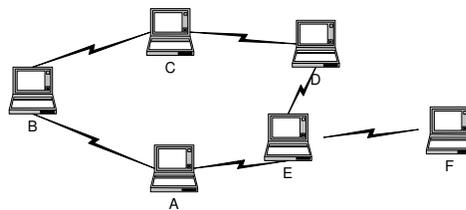


Figure 7.1 Exemple de topologie d'un réseau ad hoc.

Les protocoles de routage spécifiquement développés pour réduire la consommation des batteries ne fournissent pas de nouveaux algorithmes de routage mais proposent des améliorations à ceux déjà existants. L'idée fondamentale de ces protocoles est de router les paquets en fonction de la minimisation d'un critère relatif à la consommation des batteries, et qui peut être de trois types : (i) minimiser de manière globale, (ii) minimiser de manière locale, ou (iii) moduler la puissance d'émission.

#### A. Minimiser de manière globale

La première approche consiste à minimiser l'énergie consommée pour atteindre la destination, en cherchant le chemin de plus faible consommation. Les auteurs dans [87] fournissent un algorithme (PSR – Power-aware Source Routing), qui reprend le protocole DSR, et assigne à chacun des liens un poids qui est fonction de l'énergie nécessaire pour transmettre un paquet sur cette route. Le routage se fait donc préférentiellement suivant les routes de plus faible poids, en agrégeant l'ensemble des liens composant la route. Ce protocole a le mérite de réduire globalement la consommation du système.

Le protocole [89], appelé LPR (Lifetime Prediction Routing), tient compte de l'état des batteries de chaque terminal traversé et choisit les routes qui mettent à contribution les stations

les mieux chargées, de façon à améliorer la durée de vie du réseau. Dans cet algorithme, chaque nœud estime la durée de vie de sa batterie en se basant sur son ancienne activité. Lors d'une découverte de route, LPR choisit la route avec une durée de vie maximum et donc la route qui ne contient pas de nœuds dont la durée de vie estimée est faible.

### **B. Minimiser de manière locale**

La seconde approche tente justement d'augmenter la durée de vie du système. Le protocole LEAR (Local Energy-Aware Routing) proposé dans [88] met en cause la "bonne volonté" de chaque nœud mobile à participer au processus de sélection de route et de relayer des paquets de données au nom d'autres nœuds. À la recherche d'une route, chaque nœud utilise les informations locales sur le niveau de sa propre batterie pour décider de participer ou non au processus de sélection de la route. Ce travail étend le protocole DSR. Un nœud affamé peut conserver sa batterie en refusant de relayer des paquets qui ne lui sont pas destinés. La prise de décision dans LEAR est distribuée sur tous les nœuds et ne requiert pas des informations globales sur le réseau.

### **C. Moduler la puissance d'émission**

Enfin, la troisième approche propose un routage en fonction de la puissance de transmission. La Figure 7.2 montre un exemple de la façon dont la topologie d'un réseau ad hoc est définie par la puissance d'émission des nœuds. Les petites puissances d'émission réduisent la contention et augmentent la capacité de réseau, tout en consommant moins d'énergie. Ceci implique qu'une route avec un plus grand nombre de sauts peut être plus efficace qu'une route avec peu de sauts. Cependant, une route qui contient plus de sauts souffre d'une forte probabilité d'échec de route et de retransmission de paquets.

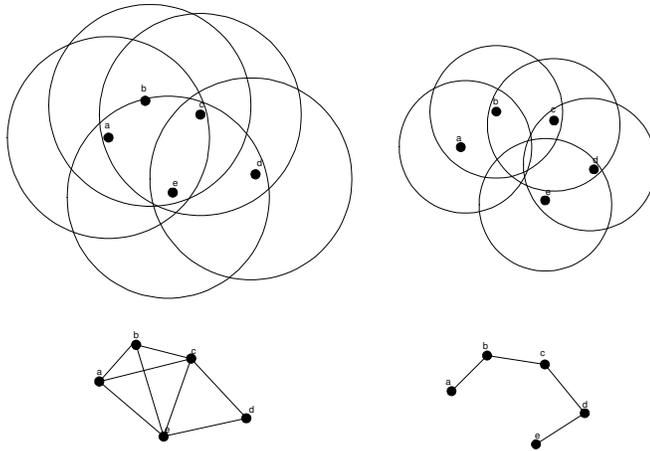


Figure 7.2 Effet de la puissance d'émission sur la topologie d'un réseau avec 5 nœuds.  
(Les cercles représentent la portée radio de chaque station ad hoc)

Dans [90], la portée de chacune des stations est contrôlée de manière à être réduite au maximum pour minimiser la consommation des batteries, tout en garantissant une connectivité globale. Pour cela, chaque nœud fixe sa puissance de façon à avoir au minimum un voisin dans chaque angle de  $2\pi/3$ . De même, dans [91] chaque nœud s'assure d'avoir un nombre fixé de voisins de manière à être ni isolé (tant que sa portée le permet), ni en interaction avec un trop grand nombre de stations (réduit en outre les interférences).

Le cas où la puissance de transmission est variable en fonction de la distance entre l'émetteur et le récepteur est traité dans [99]. Les auteurs proposent un nouvel algorithme de routage, qui suppose que la puissance de transmission et de réception requise est une fonction linéaire de  $d^\alpha$  ( $d$  est la distance entre deux nœuds voisins et  $\alpha$  est un paramètre qui dépend de l'environnement physique).

### 2.1.5 Couche Transport

La couche transport fournit un service de transport de données de bout en bout. Le protocole de transport le plus généralement utilisé pour les réseaux fixes, où les liens physiques sont assez fiables, est le protocole TCP (Transmission Control Protocol). Cependant, en raison des propriétés du lien sans fil, les performances de TCP se dégradent significativement. TCP fait appel à un grand nombre de retransmissions et à des mesures de contrôle de congestion lors d'erreur sur le lien sans fil ou de pertes dues aux handoffs. Cependant, et comme indiqué auparavant, les retransmissions consomment inutilement l'énergie des batteries.

Des propositions de protocoles telles que *Reno* et *New Reno* [92], ont été faites dans l'unique but de concevoir des protocoles de transport dans un environnement sans fil. Bien que ces protocoles aient mené, ou non, à une plus grande efficacité énergétique, ils n'ont pas directement étudié l'idée de réduire la consommation de l'énergie au niveau de la couche transport.

### 2.1.6 Couche Application

L'efficacité énergétique dans la couche application est devenue un important domaine pour les industriels ainsi que pour les chercheurs. En industrie, des API ont été développées [93] afin d'aider les réalisateurs de logiciel à créer des programmes qui sont moins consommateurs d'énergie. En recherche, les auteurs de [94] proposent une technique de transmission de vidéo encodée permettant de réduire la consommation des batteries. L'idée fondamentale de ce travail consiste à diminuer le nombre de bits transmis sur le lien sans fil afin de réduire la consommation de l'énergie tout en préservant une qualité visuelle acceptable de la vidéo.

## 3. Impact du routage sur la consommation de l'énergie

Dans cette section, nous nous positionnons au niveau 3 (couche réseau) du modèle de référence. L'objectif est d'observer l'impact du protocole de routage sur la consommation de l'énergie dans les réseaux ad hoc. Ainsi, nous cherchons à comparer quatre importants protocoles de routage ad hoc présentés dans le chapitre précédent : DSDV, OLSR, AODV et DSR. Nous évoquerons tout d'abord, les éléments importants inhérents aux algorithmes de routage étudiés qui peuvent influencer la consommation de l'énergie. Ceci nous servira ensuite, comme base pour présenter nos résultats de simulation et pour les analyser.

### 3.1 Caractéristiques énergétiques des protocoles de routage

La section suivante discute de certaines caractéristiques des protocoles étudiés qui pourraient affecter la consommation de l'énergie. En effet, quelques différences significatives, du point de vue consommation de l'énergie, entre le comportement proactif et réactif des algorithmes de routage étudiés sont listées ci-après :

- *la bande passante est différemment sollicitée* : le fait qu'il s'agisse d'un protocole de routage proactif ou réactif se traduit, en terme de consommation, par le fait que : (i) dans le cas proactif on doit consommer de l'énergie pour mettre à jour périodiquement la table de routage, et (ii) dans le cas réactif, on doit consommer de l'énergie uniquement

pour découvrir une route au cas où communication doit être établie. Dans les algorithmes de routage proactifs, les diffusions permanentes de paquets de contrôle ne seront probablement jamais utilisées ultérieurement. Clairement, les mécanismes réactifs ont un net avantage sur ce point (mais probablement pas en ce qui concerne les délais de routage) ;

- *la gestion de la mobilité* : un autre avantage des protocoles de type réactif concerne la gestion des scénarios de mobilité. Effectivement, dans le cas d'un protocole proactif, plus le nombre de nœuds en mouvement est nombreux, plus les mises à jour nécessaires à réparer les tables de routage sont nombreuses. Ces mises à jour gaspillent inutilement les batteries ;
- *Réduction des diffusions* : avec OLSR, seuls les stations multipoints MPR peuvent assumer la fonction de routeur, ce qui minimise les diffusions, et par conséquent la consommation de l'énergie.

Ces éléments précédents peuvent être résumés avec le tableau suivant :

	<b>DSDV</b>	<b>OLSR</b>	<b>AODV</b>	<b>DSR</b>
<i>Routes maintenue dans</i>	table de routage	table de routage	table de routage	cache de route
<i>Découverte de route nécessaire</i>	Non	Non	Oui	Oui
<i>Mise à jour périodique nécessaire</i>	Oui	Oui	Non	Non
<i>Mise à jour auprès de</i>	Tous les voisins	Seulement aux MPR	Pas de mise à jour	Pas de mise à jour
<i>Utilise des messages "Hello"</i>	Oui	Oui	Oui (aux voisins actifs seulement)	Non
<i>Chemin inséré dans l'en-tête du paquet</i>	Non	Non	Non	Oui
<i>Utilise des temporisateurs de route</i>	Non	Non	Oui	Non
<i>Multiple routes disponibles</i>	Non	Non	Non	Oui

Tableau 7.1 Comparaison des caractéristiques des quatre protocoles.

Ainsi, chaque approche a ses propres avantages et inconvénients que nous allons pouvoir évaluer via des simulations. Ces simulations permettront de comprendre, plus précisément, leur mode de fonctionnement relatif à l'utilisation de l'énergie.

### 3.2 Étude comparative des protocoles de routage

L'étude de performances de ces quatre protocoles de routage (DSDV, OLSR, DSR et AODV) a été réalisée grâce au simulateur ns-2 [95] développé par le groupe de recherche VINT de l'université de Berkeley. Le groupe de recherche Monarch de l'université de Carnegie Mellon - CMU a étendu le simulateur ns-2 pour inclure des scénarios de mobilité. La majorité des protocoles de routage ad hoc y sont implémentés. Les nouvelles versions de cette extension incluent, également, des modèles d'énergie pour les nœuds mobiles, et que nous avons utilisé pour cette étude. La couche MAC implémente l'interface IEEE 802.11 utilisant la fonction DCF (Distributed Coordination Function) comme méthode d'accès. La carte d'interface réseau utilise les valeurs spécifiées pour les cartes WaveLan de Lucent [96], et qui sont détaillés dans le prochain chapitre.

Nous avons considéré un scénario de base en le faisant varier par la suite. Dans le scénario de base, nous définissons le réseau ad hoc comme une surface de (600×600) m<sup>2</sup>, avec 20 nœuds aléatoirement dispersés sur cette surface. Les nœuds se déplacent pendant une durée de 1000 seconds, avec une vitesse maximum de 2m/s et un temps de pause de 35 secondes. Ils communiquent ensemble en générant chaque seconde 4 paquets CBR de 512 octets.

La consommation de l'énergie est principalement due à la transmission/réception des paquets de données et des paquets de contrôle. Enfin, nous avons apporté des modifications au scénario de base pour voir comment ceux ci peuvent influencer la consommation. Trois paramètres ont été considérés : la vitesse des nœuds, le nombre de nœuds et la quantité du trafic.

#### 3.2.1 Résultats de simulation

La consommation de l'énergie pour ces protocoles de routage en fonction de la vitesse des nœuds est présentée dans la Figure 7.3. Ces résultats indiquent que les protocoles réactifs tels que DSR et AODV consomment moins d'énergie que les protocoles proactifs. En effet, les protocoles réactifs ne consomment presque rien lorsqu'il n'y a aucun trafic dans le réseau, tandis que les protocoles proactifs consomment constamment de l'énergie par les calculs de routes, même si aucun paquet ne sera envoyé. Les protocoles réactifs sont donc moins sensibles au

déplacement des nœuds. En outre, la mise en place de nouveaux relais multipoint lorsque la topologie du réseau change, rend OLSR un peu plus consommateur que DSDV.

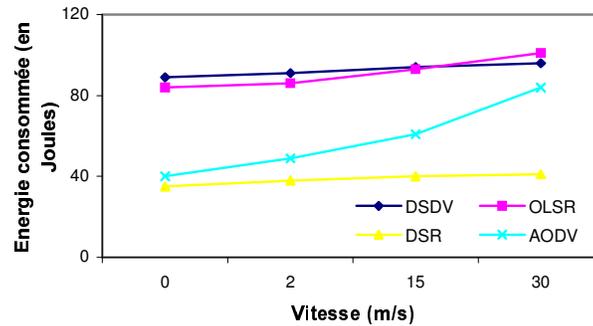


Figure 7.3 Consommation de l'énergie en fonction de la vitesse des nœuds.

Dans la Figure 7.4, on peut constater que les protocoles réactifs surpassent, encore fois, les protocoles proactifs lorsque le nombre de nœuds augmente. En fait, plus le nombre de nœuds est grand, plus les protocoles proactifs souffrent de leur mise à jour. Nous remarquons que AODV est moins stable que DSR. Nous remarquons également que si OLSR consomme beaucoup, DSDV consomme encore davantage et de manière irrégulière. En effet, OLSR réduit le nombre de diffusions aux seuls nœuds multipoints MPR. Mais globalement, les protocoles proactifs posent dans ce cas (un grand réseau avec beaucoup de nœud) un problème de passage à l'échelle (scalabilité).

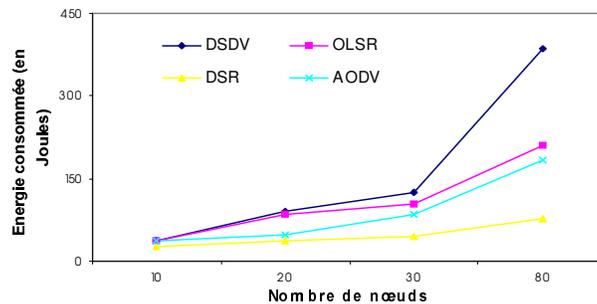


Figure 7.4 Consommation de l'énergie en fonction du nombre de nœuds.

Finalement, la Figure 7.5(a) et la Figure 7.5(b) montrent un comportement similaire des protocoles de routage, puisque les paramètres que l'on fait varier concernent le trafic et que les résultats obtenus sont semblables. À mesure que le trafic augmente, DSDV et OLSR voient leur énergie décroître de manière régulière, avec un net avantage pour OLSR. Ils rentabilisent en quelque sorte le travail de leur mise en place de la table de routage puisque les découvertes de routes ont déjà été faites de manière globale. En revanche, DSR même s'il est plutôt stable

quand le nombre de paquets croit, il réagit mal lorsque le nombre de sources augmente. Cela est dû au fait que dans le premier cas, les chemins sont les mêmes et seul le nombre de paquets varie ; alors que dans le second cas, de nouvelles découvertes de routes sont à initier pour établir la communication entre la source et la destination. À une plus grande échelle, le même problème apparaît avec AODV, puisque AODV montre aussi ses faiblesses lorsque le trafic augmente et de manière encore plus importante qu'avec DSR. En fait, AODV est contraint pour deux de ses caractéristiques : (i) les messages "hello", périodiquement envoyés aux voisins actifs dans les communications, ainsi que (ii) les re-découvertes de route forcées par un temporisateur qui permet de purger les entrées de la table.

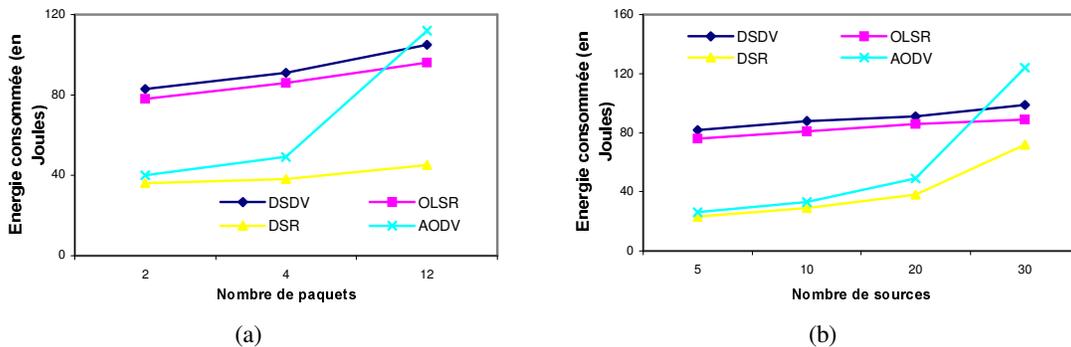


Figure 7.5 Consommation de l'énergie en fonction du (a) nombre de paquets (b) nombre de sources.

## 4. Conclusion

Dans ce chapitre, nous avons pu s'en assurer, grâce à un ensemble de résultats expérimentaux, que le choix du protocole de routage influe réellement sur le taux de consommation de l'énergie dans les réseaux ad hoc. Nous avons pu constater que, dans ce contexte, les protocoles réactifs surpassent les protocoles proactifs.

Par ailleurs, nous remarquons que ces protocoles (DSR, AODV, OLSR et DSDV) ainsi que tous les autres protocoles normalisés dans le groupe MANET de l'IETF, s'intéressent à découvrir uniquement le plus court chemin lors du processus de découverte de route. Cependant, une métrique de routage basée sur la consommation de l'énergie peut s'avérer plus efficace. D'autant plus, que cette métrique n'est pas forcément un frein à la rapidité du processus de transmission. Par exemple, une station qui assure une forte connectivité peut engendrer un délai de transmission plus long que celui prévu initialement. Ceci est dû au fait

que sa position en fait une cible de routage idéale, pouvant amener à une situation de congestion sur ces liens.

Autrement dit, le plus court chemin n'est probablement pas le critère le plus efficace, puisqu'un service de même qualité peut être apporté par d'autres critères. Nous proposons, dans la dernière partie de cette thèse, des extensions de l'un des plus importants protocoles de routage actuels qui est AODV. Ces extensions prennent en compte une métrique basée sur la consommation de l'énergie lors de découverte de route, permettant ainsi d'augmenter la durée de vie du réseau ; c'est probablement le facteur principal pour pouvoir communiquer.

# Chapitre 8

## Protocoles de routage minimisant la consommation de l'énergie

### 1. Motivations

Nous avons présenté dans le chapitre précédent quelques solutions existantes, permettant de réduire la consommation des batteries dans les réseaux ad hoc. Nous avons principalement mis le point sur les protocoles en rapport avec la couche 3 (routage) du modèle de référence, et avons démontré que la consommation de l'énergie devrait être une question cruciale lors de la conception d'un algorithme de routage.

Dans ce chapitre, nous allons présenter de nouveaux protocoles de routage ad hoc, dont la métrique de routage est basée sur la consommation de l'énergie. Ces nouveaux protocoles ont pour principaux objectifs de garantir que la connectivité du réseau soit maintenue aussi longtemps que possible, et que le niveau d'énergie du réseau entier soit du même ordre. Nous avons regroupé ces deux objectifs en un seul terme, qui est "*survivabilité*" d'un réseau ad hoc. Les protocoles que nous avons développés (LEAR-AODV, PAR-AODV, LPR-AODV et Q-AOMDV) garantissent cette survivabilité. Ce sont des protocoles réactifs et, comme l'indiquent leurs noms, ils sont basés sur l'un des plus importants protocoles de routage actuels qui est AODV (ce dernier est détaillé dans le Chapitre 6).

## 2. Le protocole LEAR-AODV

L'un des soucis majeurs des terminaux ad hoc réside dans leur caractéristique inhérente de mise à contribution de leurs ressources propres en batteries pour les besoins de routage des autres terminaux. De ce fait, le principe majeur sur lequel se base notre réflexion est l'équité. En effet, certaines situations géographiques (centre du réseau) sont relativement sollicitées ; tout comme un certain nombre de personnes qui sont des consommateurs gourmands, alors que d'autres sont des consommateurs moyens ou faibles. L'objectif consiste à ce que ces derniers (consommateurs moyens ou faibles) ne soient pas trop défavorisés lors de la sélection de route, tout en permettant aux premiers (consommateurs gourmands) d'exploiter pleinement leurs batteries. Autrement dit, l'objectif est de réduire tant que possible le problème de la station cible, dans lequel un terminal n'aura pas servi à son utilisateur, mais uniquement en tant que routeur, pour les besoins des autres utilisateurs.

Nous proposons, dans cette partie, un nouveau protocole de routage, appelé LEAR-AODV (Local Energy-Aware Routing based on AODV), permettant d'équilibrer la consommation de l'énergie entre tous les nœuds du réseau. À la recherche d'une route, chaque nœud utilise des informations locales concernant le niveau de sa propre batterie pour décider de participer ou non au processus de sélection de route. Nous utilisons le même mécanisme que celui proposé dans [88], et que nous avons présenté brièvement dans le chapitre précédent. Les auteurs proposent d'étendre le protocole DSR, en introduisant un nouveau mécanisme dans lequel un nœud affamé peut conserver sa batterie en refusant de relayer des paquets qui ne lui sont pas destinés. La prise de décision dans LEAR-AODV est distribuée sur tous les nœuds, et ne requiert pas des informations globales sur le réseau. La découverte et le maintien de la route dans LEAR-AODV sont décrits ci-dessous.

### 2.1 Découverte de route

Dans le protocole AODV, le nœud mobile n'a pas vraiment le choix, et a l'obligation de relayer des paquets pour d'autres nœuds. L'idée fondamentale du protocole LEAR-AODV est de mettre en cause cette "bonne volonté" de chaque nœud mobile à participer au processus de sélection de route, et de relayer des paquets de données au nom d'autres nœuds. Chaque nœud détermine à partir de son énergie résiduelle  $E_r$ , s'il va accepter et donc relayer le paquet de requête RREQ ou bien le rejeter. Si cette valeur est supérieure à un certain seuil  $\theta$  ( $E_r > \theta$ ), le paquet de requête RREQ est retransmis, mais dans le cas contraire, le paquet est simplement

rejeté. La destination recevra un paquet de requête seulement lorsque tous les nœuds intermédiaires au long de la route auront de bons niveaux de batterie. Ainsi, le premier message à arriver est considéré comme empruntant une route raisonnablement courte, et essentiellement avec une énergie suffisante.

## 2.2 Maintien de la route

Le maintien de la route est nécessaire soit lorsque les liaisons entre deux ou plusieurs nœuds sur la route sont perdues à cause de leur mouvements (mobilité des nœuds), soit lorsque les ressources énergétiques de certains nœuds sur la route sont épuisées trop rapidement (épuisement de l'énergie d'un nœud).

Dans le premier cas, et tout comme dans AODV, un nouveau paquet d'erreur RERR est renvoyé, et l'entrée de la table de routage correspondant au nœud qui s'est déplacé hors de portée est supprimée. Dans le deuxième cas, le nœud envoie un paquet d'erreur même lorsque la condition (8.1) est satisfaite. Pour cela, nous adoptons une approche locale, où chaque nœud intermédiaire dans la route surveille la diminution de son énergie. Ce nœud intermédiaire ne va relayer des paquets de données que si son énergie résiduelle est supérieure à un certain seuil  $\theta$ . Dans le cas contraire, le nœud envoie un paquet d'erreur RERR vers la source. La source lance alors une nouvelle découverte de route.

$$E_r \leq \theta \tag{8.1}$$

Cependant, le problème se pose dans le cas où tous les nœuds ont une énergie inférieure au seuil  $\theta$ . La source ne va recevoir, par conséquent, aucun paquet de réponse RREP, même si le chemin entre la source et la destination existe. Pour éviter ce problème, la source renvoie la même requête, mais avec un numéro de séquence incrémenté, et quand un nœud intermédiaire reçoit la requête, il diminue son seuil afin de permettre la transmission. Nous introduisons également, un nouveau message de contrôle *ADJUST\_Thr*: Quand un nœud *A* rejette une requête RREQ, il diffuse ce message *ADJUST\_Thr*, et chaque nœud qui le reçoit saura qu'un paquet RREQ a été rejeté, et donc diminue son seuil quand il reçoit un deuxième paquet de requête RREQ généré par la source.

Dès qu'une destination reçoit un paquet de requête RREQ, elle émet un paquet de réponse (RREP) vers la source. Le traitement des paquets de réponse par la source est identique à celui de AODV.

Nous insistons sur le fait que les décisions, pour un nœud ad hoc implémentant le protocole LEAR-AODV, ne dépendent que de la capacité résiduelle de sa batterie. Par conséquent, ce sont des décisions locales seulement. Le comportement de base de l'algorithme LEAR-AODV est détaillé ci-dessous :

<i>Nœud</i>	<i>Étapes</i>
<i>Nœud source</i>	<p>Diffuser un RREQ</p> <p><u>À la réception d'un RREP</u>  <i>Si premier paquet RREP reçu</i>  <i>Choisir cette route;</i>  <i>Sinon</i>  <i>Si la nouvelle route est meilleure comparée à celle déjà connue, Choisir la nouvelle ;</i></p>
<i>Nœud intermédiaire</i>	<p><u>À la réception d'un RREQ</u>  <i>Si <math>E_r \leq \theta</math></i>  <i>Ajuster (diminuer) <math>\theta</math> par <math>d</math>;</i>  <i>Diffuser ADJUST_Thr</i>  <i>Sinon</i>  <i>Si le nœud connaît une route vers la destination,</i>  <i>Envoyer un RREP vers la source</i>  <i>Sinon, Diffuser RREQ;</i></p> <p><u>À la réception d'un RREP</u>  <i>Si <math>E_r \leq \theta</math>, Ignorer toutes les futures réponses</i>  <i>Sinon</i>  <i>Si le message est le premier RREP reçu, Le renvoyer vers la source;</i>  <i>Sinon,</i>  <i>Si le message n'est pas le premier RREP reçu et il contient une meilleure route comparée à celle déjà connue, Choisir la nouvelle route et renvoyer le RREP vers la source;</i>  <i>Sinon, Rejeter le paquet RREP</i></p> <p><u>À la réception d'un ADJUST Thr</u>  <i>Si <math>E_r \leq \theta</math>, Ajuster (diminuer) <math>\theta</math> par <math>d</math>;</i>  <i>Si le nœud ne connaît aucune route vers la destination</i>  <i>Diffuser ADJUST_Thr</i>  <i>Sinon, Renvoyer (unicast) ADJUST_Thr ;</i></p>
<i>Nœud destination</i>	<p><u>À la réception d'un RREQ, Envoyer un RREP vers la source;</u></p>

### 3. Le protocole PAR-AODV

Le protocole PAR-AODV (Power-Aware Routing based on AODV) est un protocole de routage à la demande basé sur AODV. L'objectif principal est de prolonger la durée de vie d'un réseau ad hoc. Le principe de fonctionnement de ce protocole consiste à résoudre le problème de la recherche d'une route  $\pi$  à un instant de découverte de route  $t$ , de telle sorte que la fonction coût suivante [87] soit réduite au minimum :

$$C(\pi, t) = \sum_{i \in \pi} C_i(t) \quad (8.2)$$

$$\text{où } C_i(t) = \rho_i \left( \frac{F_i}{E_{r,i}(t)} \right)^\alpha \quad (8.3)$$

$\rho_i$  : puissance de transmission du nœud  $i$  ;  
 $F_i$  : capacité totale de la batterie du nœud  $i$  ;  
 $E_{r,i}(t)$  : capacité résiduelle de la batterie du nœud  $i$  à l'instant  $t$  ;  
 $\alpha$  : facteur de pondération.

Puisque la découverte et le maintien de la route dans PAR-AODV sont plus compliqués comparés à ceux utilisés dans AODV, ces deux étapes sont décrites en détail ci-après.

### 3.1 Découverte de route

Dans PAR-AODV, l'activité commence par une diffusion de paquets de requête RREQ, par la source, quand elle a des données à envoyer. Tous les nœuds, à l'exception du nœud source et du nœud destination, calculent leur coût  $C_i$  en utilisant la formule (8.3) et l'ajoutent au coût de la route dans l'en-tête du paquet RREQ (cf. formule (8.2)). Quand le nœud destination reçoit un paquet de requête, il envoie un paquet de réponse RREP vers la source.

Quand un nœud intermédiaire reçoit un paquet RREQ, il garde le coût dans l'en-tête de ce paquet comme étant un coût minimal. Si d'autres paquets RREQ arrivent avec la même destination et le même numéro de séquence, le coût du paquet RREQ nouvellement arrivé est comparé au coût minimal :

- (i) Si le nouveau paquet a un coût inférieur et si ce nœud intermédiaire ne connaît aucune route valide vers la destination, le coût minimal est mis à jour par cette nouvelle valeur et le paquet RREQ est diffusé ;
- (ii) Si le nouveau paquet a une valeur inférieure mais que le nœud intermédiaire connaît une route vers la destination, le nœud renvoie un message unicast<sup>18</sup> *COMPUTE\_Cost*. Le message de contrôle *COMPUTE\_Cost* calcule le coût de cette route ;
- (iii) Si, en revanche, le nouveau paquet RREQ a un coût plus grand, ce paquet est tout simplement éliminé.

Quand la destination reçoit un paquet RREQ ou un paquet *COMPUTE\_Cost*, elle renvoie un message de réponse RREP vers la source. Le paquet RREP est renvoyé à la source par

---

18. Message en unicast, et non en diffusion, afin de diminuer la surcharge du réseau.

l'intermédiaire du chemin inverse. Ce message de réponse contient le coût du chemin choisi. Le nœud source choisira la route avec le coût minimum.

### 3.2 Maintien de la route

Le maintien de la route dans PAR-AODV est identique à celui de l'algorithme LEAR-AODV décrit ci-dessus. Par conséquent, dans PAR-AODV, quand la capacité résiduelle de la batterie d'un nœud intermédiaire  $i$  est inférieur à un certain seuil ( $E_{r,i} \leq \theta$ ), toute requête est simplement abandonnée.

## 4. Le protocole LPR-AODV

Le routage LPR-AODV (Lifetime Prediction Routing based on AODV) est un protocole de routage à la demande, qui utilise une nouvelle métrique consistant à prédire la durée de vie des batteries. L'objectif principal est toujours le même : prolonger la survivabilité d'un réseau ad hoc. Ce protocole favorise la route dont la durée de vie est maximum. Nous entendons par là, que la route ne contient pas de nœuds dont la durée de vie est faible. Le principe du routage LPR-AODV est de résoudre le problème de la recherche d'une route  $\pi$  à un instant de découverte de route  $t$ , de telle sorte que la fonction coût suivante [89] soit maximisée :

$$\text{Max}_{\pi} \left( T_{\pi}(t) \right) = \text{Max}_{\pi} \left( \text{Min}_{i \in \pi} \left( T_i(t) \right) \right) \quad (8.4)$$

où :  $T_{\pi}(t)$ , durée de vie de la route  $\pi$  ;

$T_i(t)$ , durée de vie prévue du nœud  $i$  dans la route  $\pi$ .

Dans LPR-AODV, chaque nœud tente d'estimer la durée de vie de sa batterie en se basant sur son ancienne activité. Ceci est réalisé grâce à un récent historique sur l'activité du nœud : à l'envoi d'un paquet de données, le nœud  $i$  enregistre l'instant de l'envoi ( $t$ ) ainsi que l'énergie résiduelle à cet instant ( $E_i(t)$ ). Ces deux valeurs sont enregistrées et stockées dans le nœud. Au bout de  $N$  paquets envoyés/relayés, le nœud  $i$  récupère la valeur de l'instant de l'envoi du  $N^{\text{ème}}$  paquet ( $t'$ ) ainsi que la valeur de l'énergie résiduelle correspondante ( $E_i(t')$ ). Ce récent historique  $\{(t, E_i(t)), (t', E_i(t'))\}$  est un bon indicateur du trafic passant par le nœud. Par conséquent, nous l'avons choisi afin de prédire la durée de vie restante pour le nœud. Cette

approche est une approche distribuée permettant d'équilibrer la charge sur l'ensemble des nœuds, en choisissant les routes les moins chargées. La découverte et le maintien de la route dans LPR-AODV sont décrits ci-dessous.

#### 4.1 Découverte de route

Dans LPR-AODV, tous les nœuds à l'exception du nœud source et du nœud destination calculent leur durée de vie prévue,  $T_i$ , en utilisant la formule (8.5). Chaque requête est munie d'un champ supplémentaire représentant la durée de vie minimum de la route. Un nœud  $i$ , dans la route, remplace la valeur de cette durée de vie minimum par sa propre durée de vie,  $T_i$ , dans le cas où  $T_i$  est inférieur à cette valeur.

$$T_i(t) = \frac{E_i(t)}{\text{Vitesse\_décharge}_i(t)} \quad (8.5)$$

$$\text{où } \text{Vitesse\_décharge}_i(t) = \frac{E_i(t') - E_i(t)}{t - t'}$$

$E_i(t)$  : est l'énergie résiduelle à l'instant  $t$  ;

$t$  : temps courant correspondant au moment où le nœud envoie/relaie le paquet courant ;

$t'$  : valeur du temps enregistré dans le nœud correspondant au moment où le nœud avait envoyé/relayé le  $N^{\text{ème}}$  'prédécesseur' du paquet courant.

Plus précisément, lorsqu'un nœud intermédiaire reçoit un premier paquet de requête RREQ, il conserve la durée de vie contenue dans l'en-tête de ce paquet comme étant la durée de vie minimum de la route. Si d'autres paquets RREQ arrivent avec le même numéro de séquence et la même destination, la durée de vie du paquet nouvellement arrivé est comparée à la durée de vie minimum de la route :

- (i) Si le nouveau paquet a une plus grande durée de vie et si le nœud intermédiaire ne connaît pas de route valide vers la destination, la durée de vie minimum est mise à jour par cette nouvelle valeur et le nouveau paquet de requête RREQ est diffusé ;
- (ii) Si le nouveau paquet a une plus grande durée de vie, mais le nœud intermédiaire connaît une route vers la destination, le nœud renvoie un message unicast<sup>19</sup> *COMPUTE\_lifetime*. Le message de contrôle *COMPUTE\_lifetime* calcule la durée de vie de cette route ;

---

19. Message en unicast, et non en diffusion, afin de diminuer la surcharge du réseau.

(iii) Si, au contraire, le nouveau paquet a une durée de vie inférieure, le nouveau paquet RREQ est éliminé.

Quand la destination reçoit un paquet RREQ ou un paquet *COMPUTE\_lifetime*, elle renvoie un paquet de réponse RREP. Le paquet de réponse RREP est renvoyé à la source par l'intermédiaire du chemin inverse. Ce message de réponse contient la durée de vie de la route. Le nœud source choisira la route dont la durée de vie est la plus grande.

## 4.2 Maintien de la route

Le maintien de la route est nécessaire dans le cas où un nœud devient hors de portée d'un nœud émetteur, ou dans le cas où il détecte un changement dans sa durée de vie prévue. Dans le premier cas (mobilité des nœuds), le mécanisme est le même que dans AODV. Dans le deuxième cas, le nœud envoie un paquet d'erreur même lorsque la condition (8.6) est satisfaite. Pour ceci, chaque nœud surveille la diminution de la durée de vie de sa batterie. Quand cette diminution dépasse un certain seuil ( $\delta$ ), le nœud décide d'envoyer un paquet d'erreur RERR vers la source comme si la route était devenue non valide. Ce message d'erreur de route force la source à lancer une nouvelle découverte de route. Cette décision dépend uniquement de la capacité résiduelle de la batterie ainsi que de la vitesse de décharge de celle-ci. Par conséquent c'est une décision totalement locale.

Plus précisément, le nœud  $i$  émet un paquet d'erreur de route RERR à l'instant  $t$  quand la condition suivante est satisfaite :

$$T_i(t) \geq \delta \tag{8.6}$$

où  $t$  : temps courant.

Mais il reste toujours, comme dans LEAR-AODV, le problème où la condition (8.6) est satisfaite pour tous les nœuds. Dans ce cas, la source ne va recevoir aucun paquet de réponse RREP, même si le chemin entre la source et la destination existe. Pour éviter ce problème et comme dans LEAR-AODV, la source renvoie la même requête, mais avec un numéro de séquence incrémenté. Quand un nœud intermédiaire reçoit cette requête, il diminue son seuil  $\delta$  pour permettre la transmission. Un nouveau message *ADJUST\_Thr* est utilisé pour éviter ce problème. Quand un nœud  $A$  élimine une requête RREQ, il diffuse le message *ADJUST\_Thr*. Chaque nœud qui reçoit ce nouveau message saura qu'un paquet RREQ a été éliminé, et donc diminue son seuil quand il reçoit un deuxième paquet de requête RREQ généré par la source.

## 5. Le protocole Q-AOMDV

Nous avons développé un protocole de routage à la demande multi-routes basé sur AOMDV<sup>20</sup>. Nous l'avons appelé Q-AOMDV (Q-routing for Ad hoc on-demand Multipath Distance Vector). Le principal objectif de l'algorithme consiste à équilibrer la consommation de l'énergie à travers tous les nœuds mobiles du réseau, en envoyant les paquets de données sur les différentes routes calculées en utilisant AOMDV. Pour réaliser cet objectif, Q-AOMDV choisit l'une des routes disponibles, en se basant sur l'algorithme Q-routing défini ci-après. Ce dernier permet de trouver la route optimale que le nœud source devrait prendre, afin d'atteindre la destination avec un coût minimal.

Grâce à l'algorithme du Q-routing, chaque nœud dans le réseau ad hoc tente d'apprendre une fonction de routage de bout en bout, adaptative et permettant de prolonger la survivabilité du réseau. Dans ce qui suit, nous allons décrire la découverte, et le maintien de la route dans Q-AOMDV.

### 5.1 Découverte de route

#### 5.1.1 Calcul de routes multiples sans boucles de routage

La découverte de route dans Q-AOMDV utilise le même mécanisme de découverte de route que celui du protocole AOMDV, détaillé dans le Chapitre 6. Ce mécanisme consiste à calculer différentes routes, allant de la source de trafic jusqu'à la destination, tout en évitant la formation de boucles de routage.

destination
numéro de séquence
hopcount
liste_de_routes {(nexthop <sub>1</sub> , hopcount <sub>1</sub> ), (nexthop <sub>2</sub> , hopcount <sub>2</sub> ),...}
expiration_timeout

Destination
numéro de séquence
advertised_hopcount
liste_de_routes {(nexthop <sub>1</sub> , hopcount <sub>1</sub> , Q-value <sub>1</sub> ), (nexthop <sub>2</sub> , hopcount <sub>2</sub> , Q-value <sub>2</sub> ),...}
expiration_timeout

(a) AOMDV
(b) Q-AOMDV

Figure 8.1 Structure des tables de routage dans AOMDV et Q-AOMDV.

La Figure 8.1 montre la structure des tables de routage pour AOMDV et Q-AOMDV. La seule différence réside dans la liste appelée "liste\_de\_routes". À chacun des prochains sauts,

20. Le protocole AOMDV est présenté dans le Chapitre 6.

dans les routes candidates, est associé une Q-valeur<sub>k</sub> (Q-value<sub>k</sub>). Nous donnons plus d'explications en ce qui concerne ce paramètre Q-valeur dans la prochaine section.

### 5.1.2 Q-routing

Après un calcul des différentes routes, la source peut maintenant commencer à transmettre les données. Dans Q-AOMDV, la source (et par la suite chaque nœud intermédiaire) choisit une de ces différentes routes valides en se basant sur l'algorithme du Q-routing présenté dans le Chapitre 3. Cet algorithme essaie de trouver la route optimale que le nœud source ou le nœud intermédiaire devrait choisir, afin d'atteindre la destination avec un coût associé minimal<sup>21</sup>.

Dans Q-AOMDV, la politique de routage dans chaque nœud  $x$  se sert d'une table de valeurs  $Q_x(y,d)$ , où chaque valeur est une estimation, pour un voisin  $y$  et une destination  $d$ , du cumul de la charge des batteries des nœuds faisant partie de la route  $\pi$ , et allant du nœud  $y$  jusqu'au nœud  $d$ . Ainsi, quand un nœud doit faire un choix de routage, il choisit tout simplement le voisin  $y$  pour lequel  $Q_x(y,d)$  est minimum :

$$y = \arg \min_{\omega \in N(x)} Q_x(\omega, d) \quad (8.7)$$

où  $N(w)$  représente l'ensemble des voisins du nœud  $w$ .

L'apprentissage est utilisé afin de mettre à jour les valeurs de Q. À l'envoi de  $M$  paquets au nœud  $y$ ,  $x$  récupère immédiatement l'estimation de  $y$  concernant le cumul de la charge des batteries des nœuds, qui sont au long de la route restante (du nœud  $y$  jusqu'au nœud destination  $d$ ), à savoir :

$$Q_y(\bar{z}, d) = \min_{z \in N(y)} Q_y(z, d) \quad (8.8)$$

où  $N(y)$  représente l'ensemble des voisins du nœud  $y$ .

---

21. contrairement à [53] nous ne considérons pas le délai de transmission comme coût, mais plutôt la survivabilité du réseau.

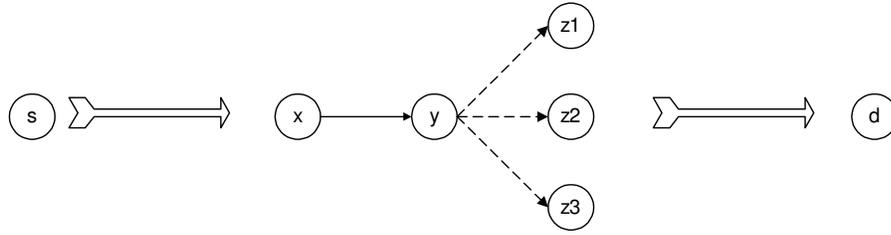


Figure 8.2 Exemple explicatif.

Le nœud  $x$  peut mettre à jour son estimation en se basant sur cette rétroaction :

$$\Delta Q_x(y, d) = \eta \left( \overbrace{Q_y(\bar{z}, d) + C_y(t)}^{\text{nouvelle estimation}} - \overbrace{Q_x(y, d)}^{\text{ancienne estimation}} \right) \quad (8.9)$$

où  $\eta$  est le taux d'apprentissage et  $C_y(t)$  est défini par la formule (8.3). Ainsi, les informations sur la route restante sont employées pour mettre à jour la Q-valeur du nœud émetteur.

## 5.2 Maintien de la route

Le maintien de la route dans Q-AOMDV utilise le même mécanisme de maintien de la route que celui du protocole AOMDV, détaillé dans le Chapitre 6. Ce mécanisme consiste à ne lancer une procédure de maintien de la route que lorsque aucune des routes pré-établies par Q-AOMDV n'est valide.

## 6. Résultats expérimentaux

L'étude de performances des différents algorithmes présentés dans ce chapitre ont été réalisées grâce à l'environnement de simulation GloMoSim 2.0 [97][98]. GloMoSim implémente une couche physique qui inclut un modèle de propagation radio, des interfaces radio, et utilise DCF<sup>22</sup> comme méthode d'accès au médium. Le modèle radio prend en compte les collisions, les délais de propagation et l'atténuation du signal. La bande passante totale considérée est de 2Mbps et la portée radio de chaque nœud est de 250 mètres.

Nous simulons un réseau ad hoc composé de 36 nœuds répartis sur une surface de (800×800) m<sup>2</sup>. Des connexions aléatoires sont établies entre les différents nœuds. Ces

22. DCF est une méthode d'accès au médium définie par le groupe IEEE 802.11 pour sa norme de réseaux locaux sans fil.

connexions sont du type CBR à 4 paquets/seconde et la taille des paquets est de 1024 octets. La capacité initiale de la batterie de chaque nœud est fixée à 10 unités. Cette énergie initiale est réduite au fur et à mesure par la transmission et la réception de données. Quand elle atteint le niveau zéro, le nœud correspondant ne peut plus participer à la communication et est considéré comme 'décédé' ou 'mort'. Pour chacun des nœuds, la consommation de l'énergie est mesurée à la couche radio au cours de la simulation. Nous considérons le cas simple où la puissance de transmission est fixe. Dans ce cas-là, chaque paquet transmis ou relayé consomme une quantité fixe d'énergie. Le cas où la puissance de transmission est variable en fonction de la distance entre l'émetteur et le récepteur est traité dans [99]. Ce travail est présenté dans le chapitre précédent.

Selon les spécifications des cartes de Lucent utilisant la norme IEEE 802.11 [96], la puissance de transmission varie entre 0.045 Watts en mode sommeil et 1.25~1.50 Watts en mode réception/transmission, respectivement. Pour obtenir l'énergie consommée instantanée, la puissance d'émission est multipliée par le temps de transmission. Par exemple, la transmission d'un paquet de données de 1024 octets consomme  $6.14 \times 10^{-3}$  Joules ( $1.50 \text{ Watts} \times 1024 \times 8 \text{ bits} / 2000000 \text{ bps}$ ). Nous faisons les deux hypothèses suivantes :

- (i) Nous supposons que le temps nécessaire pour la réception de données est semblable au temps nécessaire pour la transmission de données dans chacun des nœuds intermédiaires. Le problème est qu'un nœud doit toujours pouvoir être à l'écoute d'un éventuel émetteur qui chercherait à le contacter, et que cette écoute est coûteuse en terme d'énergie. Cette écoute reste souvent inutilisée, et est donc purement 'parasite'. Elle peut toujours être réduite par des protocoles tels que PAMAS [82], présenté dans le chapitre précédent. Les normes de réseaux sans fil telles que l'IEEE 802.11 [100] et Bluetooth [101] fournissent aussi un mécanisme pour que chaque nœud puisse savoir quand se réveiller et recevoir des paquets, et dormir le reste du temps. Sans cette supposition, la consommation de l'énergie est dominée par une réception ou écoute abusive, et les algorithmes proposés deviennent moins avantageux ;

- (ii) En second lieu, la consommation de l'énergie pendant la mise en veille a été ignorée. Puisqu'un nœud peut rester en veille pendant longtemps<sup>23</sup>, une idée pour conserver l'énergie est de mettre le nœud en mode sommeil quand il est en mode veille.

La principale métrique de performances dans cette étude est la durée de vie du réseau ; c'est ce que nous avons appelé la survivabilité. Cette métrique peut être définie de différentes manières :

- 1- temps nécessaire pour le décès de  $K$  nœuds dans le réseau ;
- 2- temps nécessaire pour le décès d'un premier nœud dans le réseau ;
- 3- temps nécessaire pour le décès de tous les nœuds dans le réseau.

Les deux premières définitions ont été adoptées pour notre étude. La durée de vie du réseau pour les algorithmes proposés est comparée pour différents scénarios. Cette comparaison est souvent faite par rapport à AODV puisque ces algorithmes y sont dérivés.

Nous avons testé les deux cas où : (i) les nœuds sont fixes et donc ont une mobilité nulle, et (ii) les nœuds sont mobiles et se déplacent sur la surface de simulation avec différentes vitesses de déplacement.

## 6.1 Nœuds fixes

La Figure 8.3 montre les moments auxquels un certain nombre de nœuds décèdent à cause de l'épuisement de leurs batteries, dans le cas où tous les nœuds du réseau ad hoc sont fixes. Nous choisissons la valeur du nombre de premiers nœuds décédés ( $K$ ) entre 1 et 7. Pour Q-AOMDV, nous ne représentons que le temps de décès des deux premiers nœuds, puisqu'à la fin de la simulation (8000 secondes) seuls deux nœuds ont un niveau de batterie nulle. Nous constatons que pour AODV, le premier nœud meurt environ 2056 secondes plus tôt que dans LEAR-AODV, 2572 secondes plus tôt que dans PAR-AODV, 3244 secondes plus tôt que dans LPR-AODV, et 3976 secondes plus tôt que dans Q-AOMDV. Pareil, que pour 4 nœuds, ceux-ci

---

23. Dans notre cas par exemple, la source de trafic CBR transmet 4 paquets de 1024 octets par seconde pendant 8000 secondes de temps de simulation. Ce qui fait un temps de transmission total de 131.072 secondes ( $1024 \times 8/2000000$  bps  $\times$  8000 secondes  $\times$  4 paquets/seconde), et un même temps pour la réception. Pendant le temps qui reste ( $7737.856$  secondes =  $8000 - 2 \times 131.072$ ), le nœud ne fera absolument rien.

décèdent environ 888 secondes plus tôt que dans LEAR-AODV, 1132 secondes plus tôt que dans PAR-AODV, et 1832 plus tôt que dans LPR-AODV.

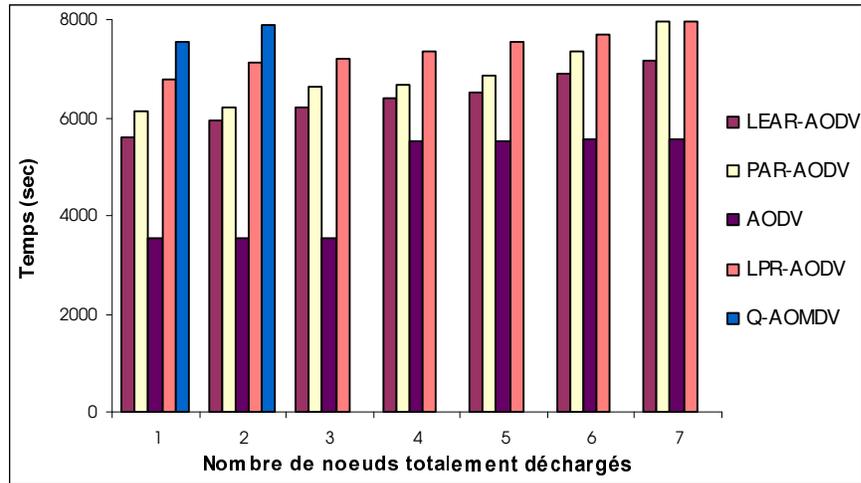
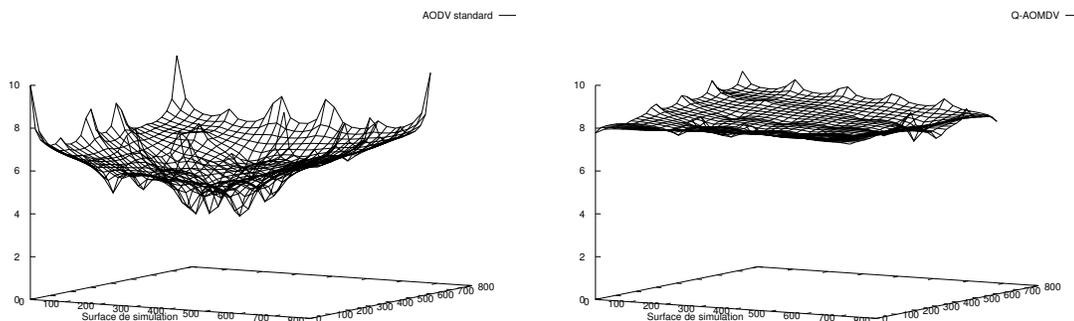


Figure 8.3 Le nombre de nœuds morts/temps.

Nous remarquons que LPR-AODV donne de meilleures performances que PAR-AODV. En effet, LPR-AODV prend en compte non seulement la capacité résiduelle de la batterie mais également sa vitesse de décharge. D'un autre côté, l'algorithme PAR-AODV n'utilise pas la 'prévision' et utilise seulement la capacité résiduelle de la batterie. Cette différence rend la fonction coût dans LPR-AODV plus précise. Effectivement, la fonction coût dans LPR-AODV utilise une prévision sur le temps de vie restant pour le nœud. Cette prévision est basée sur un récent historique de la quantité du trafic passant par ce nœud. Cet historique utilisé par LPR-AODV est un bon indicateur du trafic passant par le nœud. Nous avons soigneusement<sup>24</sup> fixé la taille de l'historique  $N$  à 5.

24. Nous avons testé plusieurs valeurs pour  $N$ , et c'est la valeur 5 qui donne les meilleurs résultats.



(a)

(b)

Figure 8.4 Niveau des batteries de tous les nœuds dans le réseau ad hoc en utilisant (a) AODV (b) Q-AOMDV (temps de simulation de 1000 secondes).

Afin d'améliorer la survivabilité du réseau, la variance des énergies résiduelles entre tous les nœuds devrait être réduite au minimum. La Figure 8.4 donne le niveau des batteries de tous les nœuds du réseau ad hoc après un temps de simulation de 1000 secondes. Nous considérons, pour cette expérience, que les nœuds sont fixes. Nous représentons uniquement les résultats de Q-AOMDV, comparés à ceux de AODV. Nous avons fait ce choix car nous avons remarqué que c'est Q-AOMDV qui fournit les meilleurs résultats par rapport aux autres. Dans Q-AOMDV (cf. Figure 8.4(b)), les nœuds consomment l'énergie plus équitablement. Ainsi, les nœuds au centre du réseau continuent à fournir la connectivité plus longtemps, et le réseau n'est pas rapidement partitionné. Ceci va mener à une dégradation plus 'gracieuse' du réseau : c'est l'idée de la survivabilité du réseau. En revanche, pour AODV (cf. Figure 8.4(a)) le niveau d'énergie pour les nœuds du centre est largement inférieur à la moitié du niveau initial.

## 6.2 Nœuds mobiles

L'effet de la mobilité est présenté par la Figure 8.5. Nous avons comparé les performances des trois algorithmes LEAR-AODV, PAR-AODV et LPR-AODV par rapport à AODV. Comme nous pouvons le constater, nos algorithmes sont toujours meilleurs par rapport à AODV en termes de nombre de nœuds décédés.

Nous constatons que pour AODV, et pour une vitesse des nœuds égale à 4 mètres/seconde par exemple, le premier nœud meurt environ 793 secondes plus tôt que dans LEAR-AODV, 1125 secondes plus tôt que dans PAR-AODV, et 1182 secondes plus tôt que dans LPR-AODV. Cette diminution de performances, par rapport au cas où les nœuds sont fixes, est tout à fait légitime. À mesure que la vitesse de déplacement des nœuds augmente, le taux de

consommation de l'énergie dans le réseau augmente aussi. Ceci est normal puisqu'une vitesse de déplacement plus élevée implique plus de découvertes de route, et par conséquent une consommation d'énergie plus élevée dans le réseau. En outre, à mesure que la mobilité des nœuds augmente, la différence entre AODV et nos algorithmes devient moins importante.

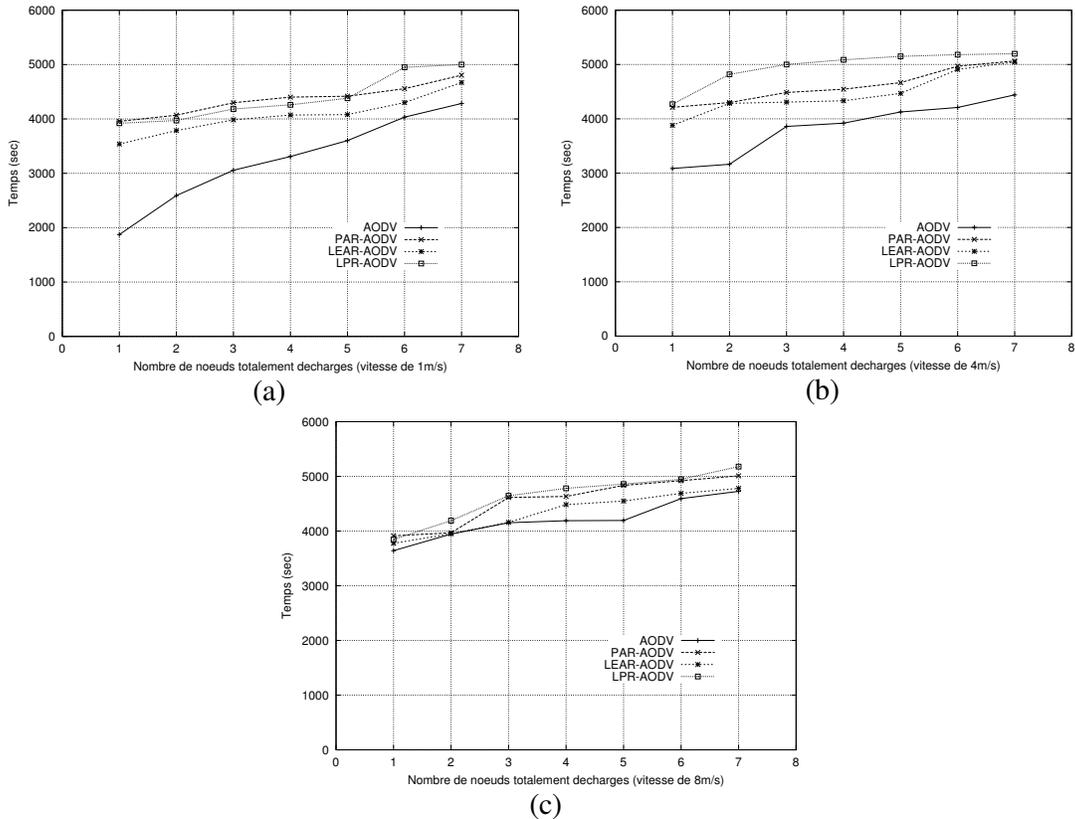


Figure 8.5 Nombre de nœuds morts avec une vitesse (a) de 1m/s (b) 4m/s (c) 8m/s.

Afin de trouver la meilleure route disponible, les algorithmes que nous avons proposés ont besoin, lors des processus de découverte/maintien de route, de propager davantage de paquets de contrôle dans le réseau. Pour mesurer cette surcharge de signalisation, nous avons calculé le rapport entre la quantité de paquets de contrôle (en octets) et la quantité de paquets données (en octets) transmis dans le réseau pendant un temps de simulation de 6000 secondes. Nous pouvons voir, dans la Figure 8.6, les valeurs de la surcharge en fonction de la vitesse de déplacement des nœuds dans le réseau. La différence entre AODV et nos algorithmes (LEAR-AODV, PAR-AODV et LPR-AODV) croît avec la vitesse des nœuds. Ceci est dû au fait qu'en plus des mécanismes introduits dans nos algorithmes pour maintenir les routes, s'ajoute le fait que les routes ne deviennent plus valides avec des vitesses de déplacement plus grandes.

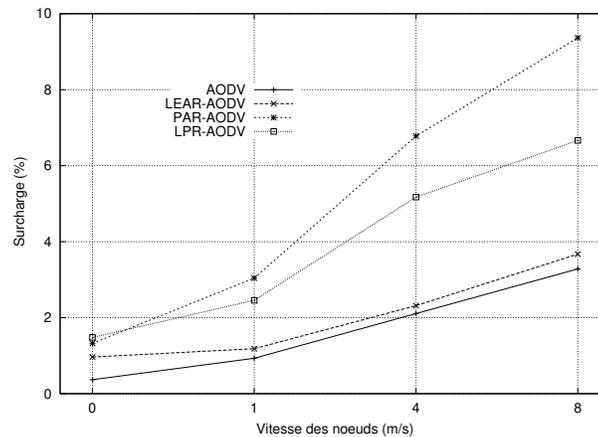


Figure 8.6 Surcharge de signalisation en fonction de la vitesse des nœuds.

L'algorithme Q-AOMDV n'a pas été testé pour le cas où les nœuds sont mobiles. Parce qu'à chaque déplacement d'un nœud, l'algorithme re-initialise sa table de routage et par conséquent, il re-initialise la Q-fonction (fonction des Q-valeurs) correspondante déjà apprise. Il recommence alors de nouveau, à apprendre une nouvelle table qui risque d'être re-initialisée si un autre nœud se déplace aussi.

Rappelons que pour une mobilité nulle l'algorithme Q-AOMDV donne les meilleurs performances, en terme de survivabilité et de surcharge, que tous les autres algorithmes. Malheureusement, à cause des éléments cités ci-dessus, l'algorithme Q-AOMDV reste inadéquat pour des réseaux ad hoc de forte mobilité.

## 7. Conclusion

Dans ce chapitre, nous proposons quatre algorithmes de routage basés sur le protocole AODV. Les algorithmes proposés ont pour but de rallonger la vie des batteries de chaque nœud, et par conséquent la survivabilité d'un réseau ad hoc. Il est à noter que les protocoles de routage actuels normalisés à l'IETF, tels que AODV, ne s'intéressent pas à l'impact du routage sur la consommation de l'énergie dans le réseau. Ils s'intéressent, en revanche, à trouver le plus court chemin en terme de nombres de sauts.

Le premier algorithme proposé, LEAR-AODV, équilibre la consommation de l'énergie dans un réseau en se basant seulement sur des informations locales. Effectivement, la prise de décision dans LEAR-AODV est distribuée sur tous les nœuds, et ne nécessite pas

d'informations globales sur le réseau (elle n'a besoin que du niveau d'énergie du nœud). Un avantage de LEAR-AODV réside dans son interopérabilité avec AODV (un réseau peut contenir un ensemble de nœuds exécutant LEAR-AODV, et un autre ensemble de nœuds exécutant AODV). Un autre avantage réside dans sa simplicité et dans sa facilité à être intégré dans tous les algorithmes de routage ad hoc existants (réactifs ou proactifs). Nous l'avons, par exemple, aisément intégré au protocole PAR-AODV. L'objectif pour le second algorithme, PAR-AODV, est de rechercher de le meilleur chemin lors d'une découverte de route. Pour cela, PAR-AODV utilise uniquement la capacité résiduelle de chaque nœud. En revanche, pour le troisième algorithme LPR-AODV, il considère également la vitesse de décharge des nœuds, lors d'une découverte de route. Le dernier algorithme proposé, Q-AOMDV, calcule plusieurs routes allant de la source de trafic jusqu'à la destination. Il équilibre la consommation de l'énergie à travers tous les nœuds mobiles du réseau en envoyant les paquets de données sur des routes différentes. Il réduit ainsi, au minimum la variance dans les énergies résiduelles de tous les nœuds. Cet algorithme est malheureusement inadapté aux réseaux ad hoc où les nœuds sont mobiles. En revanche, il donne des résultats remarquables pour un réseau ad hoc avec des nœuds fixes.

Les résultats de simulations prouvent bien que les mécanismes, rajoutés au protocole de routage AODV, améliorent considérablement ses performances en terme de survivabilité du réseau. Ces mécanismes réalisent ceci avec des surcharges minimums, et sans avoir à affecter d'autres couches protocolaires.

Comme perspectives à ces travaux, il nous semble intéressant de tester l'algorithme LEAR-AODV sur des protocoles proactifs (OLSR, par exemple), ou même hybrides (ZRP, par exemple). Il serait également intéressant de trouver comment minimiser la consommation de l'énergie dans les réseaux ad hoc utilisant ces deux dernières approches (proactif et hybride).

# Chapitre 9

## Conclusion

Les premiers réseaux cellulaires ont été conçus pour offrir principalement un service de téléphonie. Les systèmes cellulaires actuels promettent, quant à eux, une diversification des services proposés avec un débit nettement supérieur. La diversification des services (voix, messages courts, services multimédias, accès au réseau Internet, etc.) impose plusieurs niveaux de qualité de service (QoS) à garantir. Cependant, la disponibilité des ressources radios durant toute la durée de l'appel n'est pas nécessairement garantie, et les utilisateurs mobiles peuvent ainsi subir une dégradation/coupure du service. Partant du fait que la coupure/dégradation d'un appel en cours, appartenant à une classe de service de haute priorité est généralement moins désirée que la coupure/échec de connexion d'un appel appartenant à une classe de priorité inférieure, de nouveaux mécanismes de contrôle d'admission d'appels (CAC) sont fortement nécessaires. En effet, un contrôle d'admission d'appel efficace est exigé pour remédier à cette limitation du nombre de ressources radio disponibles sur l'interface radio des réseaux cellulaires. Une gestion efficace des ressources radio, par l'intermédiaire de politiques d'allocation dynamiques, s'avère également indispensable pour remédier à ce genre de problèmes. Ces nouveaux mécanismes consistent à définir des règles de gestion des ressources pour chaque classe de trafic, dans le but d'optimiser le taux d'utilisation et de satisfaire de multiples contraintes de qualité de service.

Afin de remédier à ce genre de problèmes, de nombreuses propositions existent dans la littérature ; leur objectif principal est d'éviter aux usagers le désagrément causé par des coupures de communication. Nous avons cependant constaté que ces solutions ignorent souvent

l'expérience et la connaissance qui pourraient être acquises pendant l'exécution en temps réel du système.

Un des principaux objectifs de cette thèse, a été donc de tirer profit de cette expérience et de cette connaissance afin d'optimiser certains problèmes rencontrés dans les réseaux cellulaires. Ainsi, la première contribution a donc été de trouver une nouvelle approche pour résoudre le problème de contrôle d'admission d'appel CAC dans un réseau cellulaire où les canaux sont alloués de façon permanente aux cellules (FCA). Ce réseau est supposé multi-services, et donc supporte plusieurs classes de trafic possédant différents niveaux de QoS. Cette approche utilise un système autonome basé sur l'algorithme d'apprentissage par renforcement : 'Q-Learning'. Le Q-learning est une méthode de résolution permettant d'apprendre par l'expérience, la meilleure action à effectuer en fonction de l'état actuel du réseau.

Constatant les bienfaits de cette nouvelle approche de CAC pour un système cellulaire de type FCA, nous avons utilisé la même idée pour un système à allocation dynamique des ressources (DCA). Dans DCA, tous les canaux sont maintenus dans un pool commun, tandis que les stations de base tentent d'allouer les canaux à la demande des utilisateurs. De nombreuses propositions ont été réalisées, dans la littérature, pour allouer cet ensemble de canaux. Toutefois, tout comme pour le CAC, aucune n'exploite l'expérience du fonctionnement passé du système. De ce fait, nous avons proposé une solution au problème d'allocation dynamique des ressources dans un réseau cellulaire multi-services. Cette solution, appelée Q-DCA, est également basée sur l'apprentissage par renforcement. La solution Q-DCA consiste à choisir le meilleur canal parmi tous les canaux disponibles dans le pool commun, tout en respectant un taux d'interférences tolérable. L'objectif étant de maximiser le taux d'utilisation du canal afin de maximiser le service.

Ces mécanismes proposés afin de résoudre des problématiques aussi complexes, que celles liées aux réseaux cellulaires, sont des solutions originales et de nature intelligente. Outre l'originalité de ces mécanismes, les avantages acquis, en utilisant de telles approches, peuvent être récapitulés comme suit. Contrairement à d'autres approches, elles sont plus adaptables aux variations de l'état du réseau (exp. variations des conditions de trafic, pannes d'équipements, etc.). Les tâches de contrôle d'admission et d'allocation dynamique des canaux sont déterminées rapidement et avec peu d'efforts de calcul. Elles sont obtenues grâce à une simple spécification de préférences entre les classes de trafic. Une originalité de notre proposition Q-DCA, est le fait qu'elle combine à la fois la recherche de la politique de CAC optimale, ainsi

que la recherche de la meilleure stratégie d'allocation dynamique de canaux. Nous avons démontré également, par un large ensemble d'expériences, que ces mécanismes donnent de meilleurs résultats comparés à d'autres heuristiques. Ce sont des algorithmes distribués, et les informations de signalisation échangées entre les stations de base sont quasi nulles. Ces mécanismes sont donc plus attrayants, en raison de la simplicité de leur mise en œuvre.

Vu le grand succès commercial de ces réseaux cellulaires, le développement d'équipements mobiles n'a pas cessé de prendre de l'ampleur. Grâce aux assistants personnels (PDA - Personal Digital Assistant) et aux ordinateurs portables, l'utilisateur devient de plus en plus nomade. Dans ce contexte, de nouvelles solutions de communication sans fil, de plus en plus performantes, sont en train de naître. Parmi ces principales solutions, et qui éliminent le besoin d'infrastructure fixe pour communiquer, on trouve les réseaux ad hoc.

Les réseaux ad hoc sont des architectures un peu particulières de réseaux locaux sans fil, basées sur des technologies comme le Wi-Fi qui permet de se connecter à Internet à haut débit dans un rayon de quelques centaines de mètres autour d'une borne radio (Station de Base – BS), elle même connectée au réseau filaire. Mais, alors que chaque utilisateur d'un réseau Wi-Fi se connecte via une borne radio, dans un réseau ad hoc, les terminaux (ordinateurs ou téléphones mobiles, etc.) peuvent aussi communiquer entre eux, sans intermédiaire, donc sans infrastructure. Ils peuvent même servir de relais les uns aux autres. C'est une sorte "d'architecture molle", évolutive et automatique. À la clef : souplesse et autonomie, puisque le réseau évolue en fonction des accès et des utilisateurs du moment. Il peut même être totalement indépendant de toute infrastructure. En outre, la portée du signal est démultipliée par le nombre d'utilisateurs, et les débits peuvent être préservés, alors qu'ils sont obligatoirement partagés avec une borne radio Wi-Fi.

Ces réseaux reposent avant tout sur les développements logiciels de nouveaux protocoles de routage "intelligents", qui prennent en compte les spécificités de ces réseaux (mobilité des terminaux, capacité limitée des batteries, etc.). Effectivement, en raison de la capacité limitée des batteries des terminaux, la consommation de l'énergie devrait être un critère fondamental lors de la conception de tels algorithmes de routage. Nous avons démontré, grâce à un ensemble d'expériences, que le routage utilisé est l'un des principaux facteurs agissant sur le taux de consommation d'énergie dans ces réseaux. Il est à noter que les protocoles de routage actuels, normalisés dans le groupe MANET de l'IETF, ne prennent pas en compte la métrique consommation de l'énergie pendant le processus de découverte de route. Ils s'intéressent, en

revanche, à découvrir le plus court chemin. La métrique consommation de l'énergie, peut s'avérer plus efficace. D'autant que cette nouvelle métrique n'est pas forcément un frein à la rapidité du processus de routage. Autrement dit, une métrique basée sur la consommation de l'énergie permettrait d'augmenter la durée de vie du réseau ; c'est probablement le facteur principal pour pouvoir communiquer.

Dans la dernière partie de cette thèse, nous nous sommes intéressé à ces réseaux et, en particulier, à la manière d'étendre les protocoles de routage actuels en prenant comme objectif de rallonger la durée de vie des batteries, et par conséquent la survivabilité du réseau. Pour ce faire, nous avons proposé quatre nouveaux algorithmes de routage (LEAR-AODV, PAR-AODV, LPR-AODV et Q-AOMDV) basés sur l'un des plus importants protocoles de routage actuels qui est AODV. Ce dernier, ne considère pas la contrainte de l'énergie pour optimiser le routage, mais cherche plutôt, le chemin le plus court en terme de sauts.

Ces solutions d'extensions 'énergétiques' de AODV, sont des solutions simples, et l'ensemble des résultats de simulations démontre clairement, qu'elles améliorent considérablement ses performances en améliorant la survivabilité du réseau. Elles réalisent cet objectif avec des surcharges minimales, et sans avoir à affecter les autres couches protocolaires. Ces algorithmes équilibrent la consommation de l'énergie sur la totalité du réseau. Pour l'algorithme Q-AOMDV, l'objectif a été de bénéficier de l'expérience et de la connaissance pouvant être acquises pendant l'exécution en temps réel du système. Q-AOMDV équilibre la consommation de l'énergie à travers tout le réseau, en envoyant le trafic de données sur des routes différentes. Ceci est réalisé en utilisant un algorithme d'apprentissage par renforcement appelé Q-routing. Il réduit ainsi, au minimum la variance dans les énergies résiduelles de tous les nœuds.

En conclusion, nous pouvons dire que notre principale contribution, à travers ce travail, a été de proposer et d'éprouver des mécanismes permettant de résoudre un certain nombre de problèmes rencontrés dans les réseaux mobiles (CAC, allocation dynamique de ressources, routage ad hoc). Nous avons montré, tout au long de cette étude, qu'il est possible d'utiliser des techniques venant du monde de l'Intelligence Artificielle (IA), et plus spécialement des techniques d'apprentissage afin d'élaborer des mécanismes très efficaces, et très simples à mettre en œuvre.

## 1. Perspectives

Les travaux entrepris au cours de cette thèse nous inspirent plusieurs voies de recherche.

Tout d'abord, il serait intéressant de poursuivre notre travail sur le CAC et l'allocation dynamique des ressources, en effectuant des simulations sur des systèmes de taille plus importante, ou des systèmes cellulaires à capacité variable tels que les réseaux de la 3G (comme UMTS). En effet, la formulation sous forme de SMDP nécessite la caractérisation des probabilités de transition. Cela est difficile lorsque la capacité totale du système est inconnue<sup>25</sup>. Grâce à l'apprentissage par renforcement, il n'est plus nécessaire de spécifier ces probabilités de transitions. Il suffit, tout simplement, de considérer une capacité maximale et de laisser le système apprendre la politique optimale au travers d'un ensemble d'expériences.

Constatant qu'à la fin des simulations, plusieurs états n'ont pas été visités, il serait intéressant de voir comment limiter l'espace des états aux états les plus intéressants (par exemple, les états où la capacité restante du système dépasse les 70% de la capacité totale) afin d'améliorer le temps de convergence des mécanismes proposés.

Il serait également intéressant de tester nos algorithmes en utilisant d'autres méthodes d'apprentissage par renforcement, autres que le Q-learning. On pourrait, par exemple, considérer l'autre technique d'apprentissage par différences temporelles, *Sarsa* [48].

Une autre idée d'extension, est de considérer la mise en attente (queuing) des appels bloqués. Ceci devrait améliorer les performances de nos propositions. Une implémentation prenant en compte les contraintes réelles de l'environnement, serait également intéressante afin de tester la mise en oeuvre de ces méthodes.

Une autre suite à ce travail, consiste à intégrer ces mécanismes à la technique d'allocation hybride HCA. Nous avons évoqué dans le Chapitre 5 une manière de s'y prendre.

Enfin, étant donné le succès phénoménal des réseaux de capteurs (sensor networks), il nous paraît important de tester les algorithmes de routage ad hoc, proposées dans cette thèse, sur ce genre de réseaux.

---

25. Par exemple, si elle varie dans le temps, comme dans le cas des réseaux cellulaires utilisant CDMA.

Une dernière idée consiste à tester l'algorithme LEAR-AODV sur des protocoles proactifs (OLSR, par exemple), ou même hybrides (ZRP, par exemple). Il nous semble également important d'étudier la minimisation de la consommation de l'énergie dans les réseaux ad hoc utilisant des protocoles de routage proactifs ou hybrides.

# Liste des publications

## 1. Revues nationales et internationales

- [1] S. Senouci, and G. Pujolle, "New Channel Assignments in Cellular Networks: A reinforcement Learning Solution", in *Asian Journal of Information Technology (AJIT'2003)*, pp. 135-149, vol. 2, N° 3, Juillet-Septembre, 2003, Grace Publications Network.
- [2] S. Senouci, A.-L. Beylot, G. Pujolle, "Call Admission Control in Cellular Networks: A Reinforcement Learning Solution", to appear in *ACM/Wiley International Journal of Network Management*, 2003.

## 2. Conférences nationales et internationales avec actes

- [3] S. Lohier, S. Senouci, Y. M. Ghamri Doudane, and G. Pujolle, "A reactive QoS Routing Protocol for Ad Hoc Networks", *European Symposium on Ambient Intelligence (EUSAI'2003)*, Eindhoven, Netherlands, Novembre 2003, Lecture Notes in Computer Science, Springer Verlag.
- [4] P. Simons, Y. M. Ghamri Doudane, S. Senouci, F. Bachiri, A. Munaretto, J. Kaasila, G. Le Grand, and I. Demeure, "Ambience Project - Find a Meeting", *European Symposium on Ambient Intelligence (EUSAI'2003)*, Eindhoven, Netherlands, Novembre 2003, Lecture Notes in Computer Science, Springer Verlag.

- [5] S. Senouci, and G. Pujolle, "Dynamic Channel Assignment in Cellular Networks: A Reinforcement Learning Solution", *10th International Conference on Telecommunications ICT'2003*, Papeete, Tahiti, Février/Mars 2003.
- [6] L. Ouakil, S. Senouci, and G. Pujolle, "Performance Comparison of Ad Hoc Routing Protocols Based on Energy Consumption", *Ambience Workshop 2002*, Torino, Italy, Septembre 2002.
- [7] S. Lohier, S. Senouci, Y. M. Ghamri Doudane, and G. Pujolle, "QoS routing in ad hoc networks", *Med-Hoc-Net'2002*, Sardegna, Italy, Septembre 2002.
- [8] S. Senouci, A.-L. Beylot, Guy Pujolle, "Call Admission Control for Multimedia Cellular Networks Using Neuro-Dynamic Programming", *IFIP Networking, Networking'2002*, Pisa, Italy, Mai 2002, Lecture Notes in Computer Science, 2345, pp. 1208-1213, Springer Verlag.
- [9] S. Senouci et G. Pujolle, "Nouvelle approche pour le contrôle d'admission d'appels dans un réseau cellulaire", *CFIP'2002*, Montréal, pp. 305-318, Canada, Mai 2002.
- [10] Y. M. Ghamri-Doudane, S. Senouci et G. Pujolle, "Contrôle des réseaux ad hoc à base de politiques", *CFIP'2002*, Montréal, Canada, Mai 2002.
- [11] S. Senouci, A.-L. Beylot, G. Pujolle, "A dynamic Q-learning-based call admission control for multimedia cellular networks", *IEEE International Conference on Mobile and Wireless Communications Networks (MWCN'2001)*, pp. 37-43, Recife, Brazil, Août 2001 (prix du meilleur papier).
- [12] S. Senouci, D. Gaïti et G. Pujolle, "L'approche Multi-Agents pour le contrôle adaptatif d'une constellation de satellites", *JDIR'2000*, pp. 225-233, Paris-France, Novembre 2000.

### 3. Présentations dans des conférences

- [13] Y. Ghamri Doudane, S. Senouci, A. Munaretto et M. Dias-de-Amorim, "Le projet Ambience : les perspectives au niveau QoS", *DNAC'2002*, Paris-France, Décembre 2002.
- [14] S. Senouci, Y. Ghamri Doudane, A. Munaretto et M. Dias-de-Amorim, "Le projet Ambience", *Ecole de printemps DNAC'2003*, Tozeur, Tunisie, Avril/Mai 2003.

## Bibliographie

- [1] IETF MANET WG (Mobile Ad hoc NETwork), [www.ietf.org/html.charters/manet-charter.html](http://www.ietf.org/html.charters/manet-charter.html).
- [2] G. Calhoun, "Radio Cellulaire numérique", TEC and Doc, Paris, 1992.
- [3] J. D. Gibson, "The telecommunications Handbook", *IEEE press*, 1996.
- [4] K. Al Agha, G.Pujolle, et G.Vivier, "Réseaux de mobiles et réseaux sans fil", *Editions Eyrolles*, 2001.
- [5] X. Lagrange, P. Godlewski, S. Tabbane, "Réseaux GSM-DCS", *Editions Hermès*, Paris, France, 1995.
- [6] S. Tabbane, "Réseaux Mobiles", *Editions Hermès*, 1997.
- [7] W. C. Y. Lee, "Mobile Cellular telecommunications", *McGraw-Hill, Inc.*, New York, USA, 1995.
- [8] B. Jabbari, G. Colombo, "Network issues for wireless communications", *IEEE Communications Magazine*, Janvier 1995.
- [9] S. Oh, D. Tcha, "Prioritized channel assignment in a cellular radio network", *IEEE Transactions on Communications*, 40(7), pp.1259-1269, Juillet 1992.
- [10] Z. Ming, Y. Tak-Shing, "The non-uniform compact pattern allocation algorithm for cellular mobiles systems", *IEEE Transactions on Vehicular Technomogy*, 40, pp. 387-391, 1991.

- [11] I. Katzela, M. Naghshineh, "Channel assignment schemes for cellular mobile telecommunications systems," *IEEE Personal Communications Magazine*, Juin 1996.
- [12] S. Tekinay, B. Jabbari, "Handover and channel assignment in mobile cellular networks," *IEEE Communications Magazine*, (29), Novembre 1991.
- [13] J. Hua, S. Rappaport, "CBWL: A new channel assignment and sharing method for cellular communication systems", *IEEE Transaction on Vehicular Technology*, 43(2), pp. 313-322, Mai 1994.
- [14] D. Everitt, D. Manfield, "Performance analysis of cellular communication systems with dynamic channel assignment", *IEEE Journal on Selected Areas in Communications*, 7(8), pp. 1172-1179, Octobre 1989.
- [15] C.I. Justin Chuang, "Performance issues and algorithms for dynamic channel assignment", *IEEE Journal on Selected Areas in Communications*, 11(6), pp. 955-963, Août 1993.
- [16] D. Xuefeng, Ten H. Lai, "An efficient priority-based dynamic channel allocation strategy for mobiles cellular networks", 1997.
- [17] M. Zhang and T.S. Yum, "The nonuniform compact pattern allocation algorithm for cellular mobile systems," *IEEE Trans. Vehicular Technology*, Vol 40, pp. 387-391, 1991.
- [18] D. C. Cox and D. O Reudink, "Dynamic channel assignment in two dimensional large mobile radio systems", *Bell Syst. Tech. J.*, Vol. 51, pp. 1611-1627, 1972.
- [19] E. Del Re, R. Fantacci, and L. Ronga, "A dynamic channel allocation technique based on Hopfield neural networks", *IEEE Trans. Vehicular Technology*, Vol. 45, pp. 26-32, 1996.
- [20] D. D. Dimitrijevic and J. Vucetic, "Design and performance analysis of the algorithms for channel allocation in cellular networks", *IEEE Trans. Vehicular Technology*, Vol. 42, pp. 526-534, 1993.
- [21] K. N. Sivarajan, R.J McEliece, and J.W.Ketchum, "Dynamic channel assignment in cellular radio", *Proc. IEEE 40th Vehicular Technology Conf.*, pp. 631-637, 1990.
- [22] R. Beck and H. Panzer, "Strategies for Handover and Dynamic Channel Allocation in Micro-Cellular Mobile Radio Telephone Systems", *IEEE Transactions on Vehicular Technomogy*, vol. 1, pp. 178-85, 1989.

- [23] A. Gamst, "Some Lower Bounds for a Class of Frequency Assignment Problems", *IEEE Transactions on Vehicular Technomogy*, vol. 35, 1986.
- [24] K. Sallberg et al., "Hybrid Channel Assignment and Reuse Partitioning in a Cellular Mobile Telephone System", *IEEE Transactions on Vehicular Technomogy*, pp. 405–11, 1987.
- [25] R. W. Nettleton, "A High Capacity Assignment Method for Cellular Mobile Telephone Systems," *IEEE Transactions on Vehicular Technomogy*, pp. 359–67, 1989.
- [26] C. L. I and P.-H. Chao, "Distributed Dynamic Channel Allocation Algorithms with Adjacent Channel Constraints", *PIMRC'94*, pp. 169–75, 1994.
- [27] K. Okada and F. Kubota, "A Proposal of a Dynamic Channel Assignment Strategy with Information of Moving Directions," *IEICE Transactions Fundamentals*, vol. E75-a, pp. 1667–73, 1992.
- [28] M. Serizawa and D. Goodman, "Instability and Deadlock of Distributed Dynamic Channel Allocation", *IEEE Transactions on Vehicular Technomogy*, pp. 528–31, 1993.
- [29] Y. Furuya and Y. Akaiwa, "Channel Segregation- A Distributed Channel Allocation Scheme for Mobile Communication Systems," *IEICE Transactions*, vol. 74, pp. 1531–37, 1991.
- [30] Y. Akaiwa and H. Andoh, "Channel Segregation-A self Organized Dynamic Allocation Method: Application to TDMA/FDMA Microcellular System", *IEEE Journal on Selected Areas in Communications (JSAC'1993)*, vol. 11, pp. 949–54, 1993.
- [31] M. Zhang and T.S. Yum, "Comparisons of channel assignment strategies in cellular mobile systems," *IEEE Trans. Vehicular Technology*, Vol. 38, pp. 211-215, 1989.
- [32] D. Hong, S. Rappaport, "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and non prioritized handoff procedure", *IEEE Transaction on Vehicular Technology*, 35(Mars), pp. 77-92, 08 1986.
- [33] O. T. W. Yu and V. C. M. Leung, "Adaptive resource allocation for priotized call admission over an ATM-based wireless PCN", *IEEE Journal on Selected Areas in Communications*, 15(7), pp. 1208-1225, Septembre 1997.

- [34] L. Yi Bing, M. Seshardi, "Queueing priority channel assignment strategies for PCS handoff and initial access", *IEEE Transactions on Vehicular Technology*, 43(03), pp. 704-712, Août 1994.
- [35] S. Tekinay, B. Jabbari, "A measurement-based prioritization scheme for handovers in mobile cellular networks", *IEEE Journal on Selected Areas in Communications*, 10(8), pp. 1343-1350, Octobre 1992.
- [36] S. Boumerdassi, "Mécanismes prédictifs d'allocation de ressources dans les réseaux cellulaires", *Thèse de doctorat d'état, Université de Versailles Saint-Quentin*, France, 1998.
- [37] A. S. Acompara, M. Naghshineh, "Control and quality of service provisioning in high speed micro cellular networks", *IEEE Personal Communications Magazine*, pp. 36-43, 01 1994.
- [38] M. Naghshineh, O. Schwartz, "Distributed call admission control in mobile/wireless networks", *PIMRS, Proceedings of Personal Indoor and mobile radio communications*, 1995.
- [39] B. Eklunch, "Channel utilization and blocking probability in a cellular mobile telephone with directed retry", *IEEE Transactions on communications*, 34(4), pp. 329-337, Avril 1986.
- [40] C. Tai-Po, S. Rappaport, "Overlapping coverage with reuse partitioning in cellular communications systems", *IEEE Transactions on Vehicular Technology*, 46(1), pp. 41-51, Février 1997.
- [41] P. Y. Tak-Shing, L. Y. Kwan, "Blocking and handoff performance analysis of directed retry in cellular mobile systems", *IEEE Transactions on Vehicular Technology*, 44(03), pp. 645-650, Août 1995.
- [42] R. Guérin, "Queueing-blocking system with two arrival streams and guard channel", *IEEE Transactions on Communications*, 36(2), pp. 153-163, Février 1988.
- [43] D. Everitt, "Traffic capacity of cellular mobile communications systems", *Computer Networks and ISDN Systems 20*, pp. 447-454, 1990.
- [44] S. Boucheron, "Théorie de l'Apprentissage", *Editions Hermès*, France 1992.
- [45] T. M. Mitchell, "Machine Learning", *McGraw-Hill companies, Inc.*, 1997.

- [46] L. P. Kaelbling, M. Littman, and A. Moore, "Reinforcement learning : a survey", *JAIR*, 4-285, Mai 1996.
- [47] R. S. Sutton and G. Barto, Andrew, "Reinforcement Learning: An Introduction", *MIT Press*, 1998.
- [48] M.-M. Corsini, cours sur l'apprentissage par renforcement, <http://www.sm.u-bordeaux2.fr/~corsini/Cours/HeVeA/rl.html>.
- [49] C. J. C. H. Watkins and P. Dayan, "Q-learning", *Machine Learning*, vol. 8, pp. 279-292, 1992.
- [50] C. J. C. H. Watkins, "Learning from delayed rewards," *PhD. thesis*, University of Cambridge, *Psychology Department*, 1989.
- [51] S. Thrun, "Efficient exploration in reinforcement learning", *Technical Report CMU-CS-92-102*, Carnegie Mellon, 1992.
- [52] M. Sridharan, G. Tesauro, "Multi-agent Q-learning and Regression Trees for Automated Pricing Decisions", in Proceedings of the *Seventeenth International Conference on Machine Learning (ICML'00)*, Stanford, CA, 29 Juin-2 Juillet, 2000.
- [53] J. A. Boyan and M. L. Littman, "Packet routing in dynamically changing networks: A reinforcement approach", In *Advances in Neural Information Processing Systems (NIPS'94)*, volume 6, pages 671-678, San Mateo, CA, 1994.
- [54] P. Marbach, O. Mihatsch and J. N. Tsitsiklis, "Call admission control and routing in integrated services networks using neuro-dynamic programming", *IEEE Journal on Selected Areas in Communications (JSAC'2000)*, vol. 18, N°. 2, pp. 197 –208, Février 2000.
- [55] P. Marbach, O. Mihatsch, M. Schulte and J. N. Tsitsiklis, "Reinforcement learning for call admission control and routing in integrated service networks," in *Jordan, M., et al., ed. Advances in NIPS 10*, MIT Press, 1998.
- [56] J. Nie and S. Haykin, "A Q-Learning based dynamic channel assignment technique for mobile communication systems", *IEEE Transactions on Vehicular Technology*, vol. 48, N°. 5, Septembre 1999.
- [57] P. Marbach, J. N. Tsitsiklis, "A Neuro-Dynamic Approach to Admission Control in ATM Networks: The Single Link Case", *ICASSP'97*, 1997.

- [58] R. Ramjee, R. Nagarajan and D. Towsley, "On Optimal Call Admission Control in Cellular Networks", *IEEE INFOCOM*, pp. 43-50, San Francisco, CA, Mars 1996.
- [59] Mitra, M. I. Reiman and J. Wang, "Robust dynamic admission control for unified cell and call QoS in statistical multiplexers," *IEEE Journal on Selected Areas in Communications (JSAC'1998)*, vol. 16, no. 5, pp. 692-707, 1998.
- [60] L. Breiman, J.H. Friedman, R.A. Olsen, and C.J. Stone, "Classification and Regression Trees", *Chapman & Hall*, 1984.
- [61] W.S. McCulloch et W. Pitts, "A logical calculus of the ideas Imminent in Nervous Activity", *Bulletin of math. Biophysics*, Vol. 5, 1943.
- [62] H. Tong and T. X. Brown, "Adaptive Call Admission Control under Quality of Service Constraint: a Reinforcement Learning Solution", *IEEE Journal on Selected Areas in Communications (JSAC'2000)*, vol. 18, N° 2, pp. 209-221, Février 2000.
- [63] COST242, "Multi-rate models for dimensioning and performance evaluation of multiservice networks", *M. Ritter, P. Tran-Gia Eds*, Juin 1994.
- [64] D.P. Bertsekas and J. N. Tsitsiklis, "Neuro-Dynamic Programming", *Athena Scientific*, 1996.
- [65] H. Tong, "Adaptive Admission Control for Broadband Communications", *Ph.D. thesis*, University of Colorado, Boulder, Eté 1999.
- [66] D. A. Levine, I. F. Akyildiz & M. Naghshineh, "A Resource Estimation and Call Adaptation Algorithm for Wireless Multimedia Networks Using the Shadow Cluster Concept", *IEEE/ACM Transactions on Networking*, Volume 5, N° 1, pp.1-12, Février 1997.
- [67] J. Tajima and K. Imamura, "A strategy for exible channel assignment in mobile communication systems," *IEEE Transaction on Vehicular Technology*, vol. 37, pp. 92-103, Mai 1988.
- [68] W.-B. Yang and E. Geraniotis, "Admission policies for integrated voice and data traffic in CDMA packet radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 12, pp. 654-664, Mai 1994.
- [69] Projet ITEA Ambience : <http://www.extra.research.philips.com/euprojects/ambience/>.

- [70] Y. M. Ghamri-Doudane, S. Senouci, Guy Pujolle, "Contrôle des réseaux ad hoc à base de politiques", *CFIP 2002*, Montréal, Canada, Mai 2002.
- [71] R. Rajan, D. Verma, S. Kamat, E. Felstaine, et S. Herzog, "A policy framework for integrated and differentiated services in the Internet", *IEEE Network Magazine*, vol. 13, N°. 5, pp. 36-41, September/Octobre 1999.
- [72] C. Huitema, "Routing in the internet", *Prentice Hall*, 1995.
- [73] T. Clausen (Ed.), and P. Jacquet (Ed.), "Optimized Link State Routing Protocol", *IETF Internet Draft*, draft-ietf-manet-olsr-11.txt.
- [74] C. E. Perkins et P. Bhagwat, "Highly dynamic Destination-Sequenced Distance-Vector routing for mobile computers", *ACM SIGCOMM'94*, pp.234-244, October 1994.
- [75] D. B. Johnson, D. A. Maltz, Y.-C. Hu, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR)", *IETF Internet Draft*, draft-ietf-manet-dsr-09.txt.
- [76] C. E. Perkins, E. M. Belding-Royer, and S. R. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing", *IETF Internet Draft*, draft-ietf-manet-aodv-13.txt.
- [77] M. K. Marina, S. R. Das, "Ad hoc on-demand multipath distance vector routing", *ACM SIGMOBILE Mobile Computing and Communications Review*, Vol. 6, Issue 3, pp. 92-93, July 2002.
- [78] Z. J. Haas, M.R. Pearlman, P. Samar, "The Zone Routing Protocol (ZRP) for Ad Hoc Networks", *IETF Internet Draft*, draft-ietf-manet-zone-zrp-04.txt.
- [79] M. Stemm and R.H. Katz, "Measuring and reducing energy consumption of network interfaces in hand-held devices", *IEICE Transactions on Fundamentals of Electronics, Communications, and Computer Science*, 1997.
- [80] C. E. Jones, K. M. Sivalingam, P. Agrawal, and J. -C. Chen, "A survey of energy efficient network Protocols for wireless Networks", in *Wireless Networks 7*, pp. 343-358, 2001.
- [81] K. M. Sivalingam, J.-C. Chen, P. Agrawal, and M. Sivastava, "Design and analysis of lower-power access protocols for wireless and mobile ATM networks", *ACM/Blatzer Wierless Networks 6(1)*, pp. 73-87, 2000.

- [82] S. Singh and C.S. Raghavendra, "PAMAS - power aware multi-access protocol with signalling for ad hoc networks", *ACM Computer Communication Review* 28(3), pp. 5-26, July 1998.
- [83] P. Karn, "MACA - a New Channel Access Method for Packet Radio", *ARRL/CRRL Amateur Radio 9th Computer Networking Conference*, Ontario, Canada, pp. 134-140, September 1990.
- [84] M. Zorzi and R.R. Rao, "Error control and energy consumption in communications for nomadic computing", *IEEE Transactions on Computers*, 46(3), pp. 279-289, 1997.
- [85] P. Lettieri, C. Fragouli and M.B. Srivastava, "Low power error control for wireless links", *Proceedings ACM MobiCom*, Budapest, Hungary, 1997.
- [86] S. Ni, Y. Tseng, Y. Chen, et J. Chen, "The broadcast storm problem in a mobile ad hoc network", *Proceedings of the Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pp. 152-162, 1999.
- [87] M. Maleki, K. Dantu, and M. Pedram, "Power -aware Source Routing in mobile ad hoc networks", *Proceedings of ISLPED '02*, Monterey, CA, pp. 72-75, August 12-14, 2002.
- [88] K. Woo, C. Yu, D. Lee, H. Y. Youn, and Ben Lee, "Non-Blocking, Localized Routing Algorithm for Balanced Energy Consumption in Mobile Ad Hoc Networks", *MASCOTS'01*, Cincinnati, Ohio, August 15-18, pp. 117-124, 2001.
- [89] M. Maleki, K. Dantu, and M. Pedram, "Lifetime Prediction Routing in Mobile Ad Hoc Networks", *IEEE Wireless Communication and Networking Conf.*, Mars, 2003.
- [90] R. Wattenhofer, L. Li, P. Bahl, et Y.-M. Wang, "Distributed Topology Control for Power Efficient Operation in Multihop Wireless Ad Hoc Networks", *IEEE INFOCOM'2001*, April 2001.
- [91] R. Ramanathan, and R. Rosales-Hain, "Topology Control of Multihop Wireless Networks using Transmit Power Adjustment", *IEEE INFOCOM'2000*, March 2000.
- [92] M. Zorzi and R.R. Rao, "Energy efficiency of TCP in a local wireless environment", *Mobile Networks and Applications* 6, pp. 265-278, 2001.
- [93] Intel Corporation, Microsoft and Toshiba Corporation, "Advanced Configuration and Power Interface", <http://www.teleport.com/~api>.

- [94] P. Agrawal, J.-C. Chen, S. Kishore, P. Ramanathan and K.M. Siva-lingam, "Battery power sensitive video processing in wireless networking", *Proceedings of IEEE PIMRC'98*, 1998.
- [95] The Network Simulator ns-2, <http://www.isi.edu/nsnam/ns/>.
- [96] A. Kamerman, and L. Monteban, "WaveLAN-II: A High-Performance Wireless LAN for the Unlicensed Band", *Bell Labs Technical Journal*, pp. 118-133, Summer 1997.
- [97] X. Zeng, R. Bagrodia, and M. Gerla, "GloMoSim: a Library for Parallel Simulation of Large-scale Wireless Networks", *Proceedings of the 12th Workshop on Parallel and Distributed Simulations (PADS '98)*, Banff, Alberta, Canada, May 26-29, 1998.
- [98] R. Bagrodia, and al, "PARSEC: A Parallel Simulation Environment for Complex Systems," *IEEE Computer*, Oct. 1998.
- [99] Stojmenovic and X. Lin, "Power-aware localized routing in wireless networks", *Proc. IEEE IPDPS*, Cancun, Mexico, May, 2000.
- [100] Hagen Woesner, Jean-Pierre Ebert, Morten Schlager, and Adam Wolisz, "Power-Saving Mechanisms in Emerging Standards for Wireless LANs: The MAC Level Perspective", *IEEE Personal Communications*, Vol. 5, Issue 3, pp. 40-48, June 1998.
- [101] "Complete Bluetooth Tutorial", <http://infotooth.tripod.com/tutorial/complete.htm>.