

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Hadj Lakhdar - Batna

Faculté des sciences

Département d'informatique

Mémoire

Présenté en vue de l'obtention du diplôme de
Magistère en informatique

Option : Ingénierie des systèmes informatiques

THÈME

**Gestion des ressources radio
Dans un Réseau hétérogène**

Par

Gourdache Samir

Soutenance le : 16 Mars 2011

Devant le jury composé de :

Pr. Benmohammed Mohamed, Professeur à l'université de Constantine : Président du jury,
Dr. Zidani Abdelmadjid, Maître de Conférences-A- à l'université de Batna : Examineur,
Dr. Kazar Okba, Maître de Conférences-A- à l'université de Biskra : Examineur,
Dr, Bilami Azeddine, Maître de Conférences-A- à l'université de Batna : Rapporteur.

A tous ceux pour qui mes joies sont les leurs, et mes tristesses sont les leurs aussi, je dédie cet humble travail en signe de reconnaissance et d'affection;

sachant que

ce que je leur offre est en deçà de ce que je leur doit,

et peu, en face de ce qu'ils méritent.

Je tiens à remercier Mr Bilami azzeddine pour avoir accepté de diriger ce travail et pour avoir eu confiance en mes choix. Mes remerciements vont aussi à messieurs les membres du jury qui ont bien voulu examiner et évaluer ce travail. Mr Belahcen Moussa voudra bien recevoir ma reconnaissance pour son amabilité et sa serviabilité. Je saisi cette occasion aussi pour remercier messieurs Benabes Nacer eddine, Benafia ali, et Derradji salah, tous enseignants à l'université de Batna, pour les sincères sentiments de respect et d'estime qu'ils ont manifesté à mon égard.

Je ne saurais, enfin, oublier tous ceux qui ont marqué de l'intérêt à l'idée de me voir achever ce travail.

ABSTRACT

Nous proposons une approche pour la gestion des ressources radio dans un réseau sans fil hétérogène qui s'adapte mieux à la gestion de la mobilité, tout en renforçant les autres composants de cette fonction (contrôle d'accès, mobilité, handover, partage de charge,...etc). L'idée principale est d'exploiter, aux moments opportuns, les disparités tant temporelles que spatiales, qui caractérisent la répartition des états de charge en ressources radio des différentes zones des différents réseaux constituant le réseau hétérogène. Ces disparités sont dues, d'une part, aux fluctuations dans les degrés d'utilisation des ressources radio au cours du temps et, d'autre part, aux décalages temporels qui caractérisent ces fluctuations. Le volume de ressources radio "oisives" générées dans certaines "zones de couverture de base sans fil" ZBS (cellule, espace couvert par un point d'accès,...etc.), viennent "soulager" les zones ayant un besoin de ces ressources. L'approche utilise le handover vertical en tant que mécanisme de base de rééquilibrage. Nos apports, nous les sous-titrons "prévision" et "provision". Nous nous appuyons à cet effet sur deux particularités très intéressantes des réseaux hétérogènes. La première de ces particularités est constituée par l'existence, dans un même espace, d'un fort degré de chevauchement de couverture sans fil par les différentes technologies d'accès. La deuxième particularité vient de la nature multimodale des équipements des utilisateurs. Pour la prévision, ces deux caractéristiques nous permettent d'établir une sorte de carte dynamique de répartition des différents équipements des utilisateurs à travers tout le réseau hétérogène. Elle met en évidence dans chaque zone de base, une variété de sous zones différentes, entraînant un découpage plus fin de l'espace de couverture sans fil et rendant, par là même, la tâche des "capteurs de contextes" plus aisée. Un "capteur de contexte" est une entité "logicielle" attachée à chaque ZBS dont le rôle est de tenir celle-ci au fait du contexte à tout moment. En termes de provision en ressources radio, nous proposons un schéma pour la redistribution des ressources radio dans tout le réseau hétérogène. Sa raison d'être est de rétablir les déséquilibres de charges qui se créent du fait de l'activité qui règne dans le réseau. Plusieurs critères structurant la conception de ce schéma ont été observés. Nous avons, à dessein, choisi un schéma simple, fonctionnant de manière cyclique et distribuée, où les entités impliquées peuvent se contenter d'un minimum de connaissances mutuelles grâce une "symétrie de comportement". Le schéma est flexible, efficace, équitable, et pareto optimal. Cette dernière propriété est rendue effective par l'introduction d'un "comportement de port virtuel". Le schéma nous permet aussi de déduire certaines formules relatives aux nombres maximums de cycles nécessaires pour réaliser la distribution, aux pourcentages des ressources mobilisées durant les divers cycles, ainsi qu'aux cas particuliers où les distributions nécessitent les nombres maximums de cycles. Enfin nous pensons que notre schéma doit être analysé en tant point d'articulation des divers composants impliqués dans la GRR.

Table des matières

INTRODUCTION	8
CHAPITRE 1	10
LES TECHNOLOGIES D'ACCES RADIO	
1.1 Introduction	10
1.2 TAR mise en œuvre dans les réseaux GSM	10
1.2.1 Présentation	10
1.2.2 Interface Air du réseau GSM	10
1.2.2.1 Canaux logiques	12
1.2.3 Principe de réutilisation des ressources	16
1.2.4 Code de couleur BSIC	17
1.2.5 L'entrelacement et construction de salves (bursts)	17
1.2.6 Contrôle par la base de la puissance d'émission	18
1.3 TAR mise en œuvre dans les réseaux GPRS	18
1.3.1 Présentation	18
1.3.2 Interface air du GPR	19
1.3.2 Allocation des ressources	20
1.3.3 Canaux logiques GPRS	20
1.3.4 Association des canaux logiques aux canaux physiques	21
1.4 TAR utilisant la communication par étalement de spectre	22
1.4.1 Présentation	22
1.4.1.1 Justification théorique de l'étalement de spectre	22
1.4.2 Techniques de l'étalement de spectre	23
1.4.2.1 Etalement par séquence directe	23
1.4.2.2 L'étalement par saut de fréquences	24
1.4.2.3 Synchronisation	24
1.4.3 IS-95 (cdmaOne)	24
1.4.3.1 Les codes pseudo bruit Courts, Longs et de Walsh	25
1.4.3.2 Uplink (reverse direction)	26
1.4.3.3 Downlink (forward direction)	27
1.4.3.4 Les canaux logiques	28
1.4.3.4.1 Canaux de la voie "Forward"	28
1.4.3.4.2 Canaux de la voie "reverse"	30
1.4.3.5 Control de puissance	30
1.4.4 CMDA2000	31
1.4.4.1 Présentation	31
1.4.4.2 Les canaux logiques	33
1.4.4.3 Quelques caractéristiques du cdma2000	34
1.4.5 WCDMA (Wideband CDMA)	35
1.4.5.1 Aspects de la couche physique du WCDMA	35
1.4.5.2 Caractéristiques de la couche physique de l'UMTS	37
1.4.5.3 Protocoles UMTS	38
1.4.5.4 Architecture radio	39
1.4.5.5 Plan de contrôle et plan utilisateur	41
1.4.5.6 Canaux physiques	41
1.4.5.7 Canaux de transport	43
1.5 TAR mise en oeuvre dans le réseau WiMAX	45

1.5.1 OFDM (Orthogonal Frequency Multiplexing)	46
1.5.1.1 Orthogonalité des porteuses des sous canaux	46
1.5.1.2 Implémentation de L'émetteur/récepteur (Génération du signale)	47
1.5.2 Avantages de l'OFDM	48
1.5.3 Technique d'accès OFDMA	50
1.5.4 Allocation des sous-porteuse	52
CHAPITRE 2	55
GESTION DES RESSOURCES RADIO (GRR)	
2.1 Introduction	55
2.2 Les composants de la GRR	56
2.3 Le Contrôle d'Admission	56
2.3.1 CA pour contrôler la Qualité du signale	57
2.3.1.1 Interférences et SIR (Signal to Inteference Ratio)	57
2.3.1.2 Charge de la cellule (ou le réseau).	58
2.3.1.3 Faisabilité de l'allocation de puissance (Power Allocation Feasibility).	58
2.3.1.4 Contrôle d'Admission Optimum avec des contraintes de qualité du signale	59
2.3.2 CA pour contrôler la probabilité d'échec du handover	59
2.3.2.1 Politique des Bandes de Garde (Guard Band policy)	59
2.3.2.2 L'état de charge dans la cellule hôte et les cellules voisines en tant que critère de CA	60
2.3.2.3 Ressources disponibles dans la cellule hôte et les cellules voisines	60
2.3.2.4 Estimation de la probabilité d'échec du handover / de surcharge.	60
2.3.3 CA pour contrôler les paramètres de QoS du niveau paquet	61
2.3.3.1 CA pour le contrôle de la vitesse de transmission.	61
2.3.4 Schémas de CA basés sur le revenu	61
2.4 Le Handover	62
2.4.1 Principes du Hard handover	62
2.4.1.1 Critères de déclenchement du hard handover	63
2.4.2 Principes du soft handover	64
2.4.2.1 Algorithme du soft handover	65
2.5 Contrôle de puissance	66
2.5.1 Contrôle de puissance en boucle ouverte	67
2.5.2 Contrôle de puissance en boucle fermée	68
2.5.3 Contrôle de puissance en boucle extérieure	68
2.6 Ordonnancement (Scheduling)	68
2.6.1 Aspects majeurs de l'ordonnancement sans fil	69
2.6.2 Quelques approches pour l'ordonnancement sans fil	70
CHAPITRE 3	73
ETAT DE L'ART DE LA PRISE DE GESTION	
DES RESSOURCES RADIO DANS LES RESEAUX SAN FIL HETEROGENES	
3.1 Introduction	73
3.2 La Prise de Décision	74
3.2.1 Mécanismes basés sur une fonction	74
3.2.2 Mécanismes basés sur un modèle mathématique	77
3.2.3 Mécanismes basés sur la Théorie des jeux	83
3.2.4 Mécanismes basés sur AHP et GRA	86
CHAPITRE 4	88
SCHEMA POUR LA GESTION DES RESSOURCES RADIO	
DANSUN RESEAU SANS FIL ET MOBILE HETEROGENE	
4.1 Introduction	88
4.2 Modélisation d'un réseau sans fil hétérogène	88
4.3 Schéma proposé pour la redistribution des RR	91

4.3.1 Principe	91
4.3.2 Problématique de la mise en œuvre du schéma	92
4.3.3 Schéma de base pour la GRR	96
4.3.3.1 Version Préliminaire	96
4.3.3.2 Commentaires sur la version préliminaire du schéma	102
4.3.3.3 Version "finale"	104
4.3.4 Résultats des Simulations	107
4.3.4.1 Cadre des simulations	108
4.3.4.2 Résultats des Simulations	109
CONCLUSION	112
REFERENCES	113

INTRODUCTION

De l'avis des spécialistes, les réseaux sans fil du futur seront des réseaux hétérogènes [1]. C'est une vision qui écarte l'alternative d'une seule technologie dominante, fournissant un service ubiquiste. En revanche, elle plaide pour une coexistence des différentes technologies d'accès radio (*TAR*), au sein de réseaux sans fil hétérogènes. Ces *TARs* devront être combinées de manière à permettre aux utilisateurs mobiles de tirer profit des caractéristiques propres de chacune d'entre elles, en termes de couverture, de débits, de *QoS*, de coûts, d'implémentation, ...etc, et d'être toujours connectés selon le label "*Always Best Connected anywhere and at anytime*" (une des clés importantes du succès de ces réseaux). Ce concept de réseau sans fil et de mobiles hétérogène (*RSMH*) ne va sans poser de réels défis à ses concepteurs, et tout particulièrement d'un point de vue Gestion des Ressources Radio (*GRR*). La rareté des ressources radio, les différents facteurs de dégradation qui affectent la propagation des ondes radio, et le fait que les communications s'opèrent à travers l'air ambiant, font de l'interface air dans les réseaux sans fil un vrai goulot d'étranglement. Ces faits réunis exigent, en conséquence, des politiques et techniques de *GRR* adéquates.

En fait, la *GRR* est un problème à facettes multiples, impliquant plusieurs aspects interdépendants. Dans les nombreux travaux dédiés à la *GRR* dans les réseaux sans fil hétérogènes (Heterogeneous Wireless Networks), trois fonctions clés récurrentes, centrées sur l'assurance d'un certain niveau de *QoS* (Quality of Service), peuvent être identifiées. Il s'agit du Control d'Admission d'Appel (*C.A.P*), de la Gestion de la Mobilité (*G.M*) et du Partage de Charge (*P.G*).

Le *C.A.P*, prend en charge le trafic entrant des utilisateurs, aussi bien celui des appels nouveaux que celui ayant pour origine des handovers (dans le cas qui nous concerne il s'agit des handovers inter technologies), et détermine s'il y a lieu ou non d'autoriser ces demandes de connexion, en générant le cas échéant les paramètres nécessaires. Une fois la connexion au réseau autorisée, la qualité du service qui lui est fournie (*QoS*) devrait être assurée pendant toute la durée de la session de communication. De même, les décisions d'admissions de nouvelles connexions dans le réseau ne devraient pas résulter en une altération de la *QoS* des connexions en cours. Pour jouer son rôle de régulateur et améliorer la justesse de ses prises de décisions, le *C.A.P* gagnerait beaucoup à être au fait du contexte (informations sur la mobilité des utilisateurs, les opérations de handover potentielles ainsi que les mesures de *QoS* actuelles).

La *G.M* est un autre volet impliquant la *GRR*. En effet, la nature aléatoire de la mobilité de utilisateurs, ajoutée au manque de modèles précis et modérément complexes (de ce phénomène), introduisent beaucoup d'incertitudes quant à la direction, l'importance, et la localisation des flux des utilisateurs mobiles à travers le réseau. A moins d'être pris en considération, ce fait entraînerait une inefficacité de la *GRR* et une détérioration de la *QoS*. Anticiper sur l'évolution des flux des utilisateurs, et par conséquent, sur les demandes de ressources qui en découlent, et des risques qu'encours la *QoS*, est capital pour l'entité chargée de la *G.M*. En effet, bien qu'on puisse avoir un *C.A.P* efficace qui régulerait l'affluence des mobiles, il y aura toujours le phénomène (incontrôlable) de la mobilité qui viendra provoquer des disparités entre les charges des zones de services, et créer de la sorte, de fortes demandes de ressources radio dans certaines d'entre elles. Il est donc important d'avoir un mécanisme de partage de charges (*P.G*) qui vienne renforcer le *C.A.P*, en réparant les déséquilibres que cette dernière fonction ne réussirait pas à éviter, d'où l'intérêt de la fonction *P.G*. La *QoS* constitue une préoccupation commune aux trois composants de la *GRR* dont il vient d'être question. Ceci implique que la mise en œuvre de toute approche touchant la *GRR*, doit considérer la coopération très étroite entre ces composants (*C.A.P*, *G.M* et *P.G*). Le travail que nous présentons est en grande partie orienté vers la recherche d'une telle mise en oeuvre.

L'idée que nous voulons explorer dans notre approche part du constat que le fonctionnement des réseaux sans fil est marqué par la fluctuation dans le degré d'utilisation des ressources radio (degrés qui dépendent, notamment, de la phase du cycle de vie du réseau, de la période de la journée, et de la localisation des zones de services). Dans le cas d'un réseau sans fil et de mobiles hétérogène, cela engendre des situations dans lesquelles certaines zones d'un réseau sont surchargées alors que dans leurs

régions de couverture communes avec d'autres réseaux, se trouve des zones ayant un volume de ressources "oisives". Si nous supposons que ces fluctuations sont asynchrones, que les ages des réseaux sont différents et que les zones de couvertures sont nombreuses mais non identiques, alors on peut espérer, pouvoir tirer profit de ce recouvrement des réseaux dans la gestion des ressources radio, en créant une sorte de réserve quasi permanente et partageable de ressources radio (d'une certaine quantité), mobilisée périodiquement à l'effet de réparer les déséquilibres en ressources radio qui viendraient à apparaître ca et la dans le réseau hétérogène . Les raisons qui poussent à croire qu'un tel scénario est possible sont que:

- Les réseaux sont déployés dans des zones où une activité et un potentiel d'utilisateurs élevés sont prévus (donc les régions de couvertures sont celles où l'activité est censée être intense),
- Les réseaux sont dès leurs conceptions surdimensionnés, en prévision de la montée en charge progressive, future.
- Un réseau jeune, est une opportunité pour les autres réseaux (bien que pour l'opérateur l'utilisation des ressources se trouve accélérée), alors qu'un réseau en fin de cycle voit sa durée de vie prolongée car offrant des ressources laissées libres par la migration des anciens utilisateurs (ne serait ce que pour des applications ne nécessitant pas un débit ou un QoS élevés).

Partant de ces constats notre travail adopte une approche, intuitive, pour la *GRR* dans un réseau sans fil et de mobiles hétérogène (*RSMH*), proposée dans certains travaux de recherche [69,78], qui consiste à utiliser le handover vertical (*HOV*) en tant qu'outil de base pour cette gestion. Cette approche s'articule autour d'un schéma de redistribution périodique des ressources radio, dont le but est de mettre en œuvre une politique agressive de réutilisation des ressources radio "oisives", dans tout le réseau hétérogène. Cela se traduira par une quête permanente des dites ressources radio "oisives" en vue de leur utilisation là où le besoin se fait sentir. Pour jouer le rôle qui lui est dévolu au sein de toute l'architecture, ce mécanisme doit répondre à certains critères de conception tels que l'efficacité, la flexibilité, l'équité, la simplicité, l'extensibilité .. etc.

La suite de ce travail sera organisée comme suit : dans le premier chapitre seront passés en revue les principales technologies d'accès sans fil à l'œuvre dans les réseaux actuellement déployés. A travers cette revue, nous cherchons à montrer en quoi consiste cette hétérogénéité, des réseaux sans fil et de mobiles, dont il est question dans le présent travail. Le deuxième chapitre, sera consacré à la gestion des ressources radio dans les réseaux sans fil "homogènes", c.a.d. à technologie d'accès unique. Il sera question des fonctions les plus importantes que la *GRR* doit assurer dans ces types de réseaux. Au troisième chapitre, un état de l'art en matière de *GRR* dans les réseaux sans fil hétérogènes est présenté. Nous insistons plus particulièrement sur les mécanismes de prise décisions adoptés dans les divers travaux de recherches dédiés à la question. La partie qui constitue le cœur de notre travail fera l'objet du quatrième chapitre. Après une présentation du modèle de *RSMH* que nous avons adopté, nous détaillerons les deux apports majeurs que nous estimons avoir réalisés. Le premier de ces apports a trait à la prévision, l'autre à la provision des ressources radio. Nous discuterons, bien entendu, de divers points se rapportant au schéma que nous proposons et détaillerons le schéma proprement dit. Nous cloterons le chapitre par les commentaires des résultats que les simulations nous ont fournis.

Chapitre 1

LES TECHNOLOGIES D'ACCES RADIO

1.1 Introduction

Le réseau radio est la partie du réseau sans fil qui inclut la station de base (ou le point d'accès), les stations mobiles, ainsi que l'interface entre elles. Cette partie du réseau revêt une importance considérable, vu que c'est celle qui est directement liée à l'utilisateur mobile. La station de base (ou le point d'accès) a une connexion radio avec le mobile, et elle doit être capable de communiquer avec ce dernier à l'intérieur d'une certaine zone de couverture, et de maintenir des standards de qualité d'appel. Le réseau radio doit être en mesure d'offrir une capacité et une couverture suffisantes. Les stations mobiles sont contraintes à la compétition pour l'accès sans fil au réseau, à travers l'interface air. Différentes technologies d'accès sans fil ont été mises en œuvre dans les divers réseaux déployés de part le monde. L'examen des implémentations de ces différentes technologies d'accès radio (*TAR*) nous donnera une idée assez précise de ce en quoi cette hétérogénéité consistera, dans les réseaux sans fil hétérogènes prédits pour le futur, et conditionnera certainement les approches et les choix qui présideront à la conception de ces réseaux.

1.2 TAR mise en œuvre dans les réseaux GSM

1.2.1 Présentation

Le GSM, (Global System for Mobile communications), est un système cellulaire et numérique de télécommunication mobile. Il a été rapidement accepté et a vite gagné des parts de marché telles, qu'aujourd'hui plus de 180 pays ont adopté cette norme et plus d'un milliard d'utilisateurs sont équipés d'une solution GSM [1]. Les objectifs fixés au système GSM sont [2] :

- Offrir un vaste éventail de services de télécommunications compatibles avec ceux des réseaux fixes.
- Offrir des services spécifiques dus à la mobilité des usagers.
- Assurer la compatibilité d'accès à n'importe quel utilisateur dans n'importe quel pays exploitant le système GSM.
- Assurer la localisation automatique des mobiles sous la couverture globale de l'ensemble des réseaux.
- Permettre une grande variété de terminaux mobiles.
- Obtenir une bonne efficacité spectrale.
- Obtenir des coûts permettant d'assurer le succès du service.

1.2.2 Interface Air du réseau GSM

Le spectre radio étant une ressource limitée, partagée par plusieurs utilisateurs, les défenseurs de la norme doivent l'utiliser à bon escient et avec parcimonie. Le premier choix architectural dans le GSM a été donc de partager la bande de fréquence allouée, en utilisant une combinaison des deux techniques de multiplexage *TDMA/FDMA* (Time/Frequency Division Multiple Access). La partie FDMA divise la bande de fréquences allouée en plusieurs sous bandes (124 dans le cas du GSM) d'égales largeurs, espacées de 200 kHz. Chacune de ces sous bandes de fréquences est ensuite découpée, à son tour, en intervalles de temps égaux, utilisant un schéma TDMA.

Dans le système GSM/DCS (PCS aux E.U), deux bandes de fréquences sont utilisées, l'une autour des 900 MHz et l'autre autour de 1,8 GHz (1,9 aux E.U et 298 porteuses). Chaque bande est divisée en deux sous bandes servant, l'une pour le transfert d'informations entre le mobile et la station de base (**voie montante**), et l'autre pour la liaison entre la station de base et le mobile (**voie descendante**) [3]. La figure 1.1 illustre les bandes de fréquences utilisées pour la liaison entre mobile et station de base dans le réseau GSM.

- Bande EGSM étendue (bande de largeur totale 35 MHz)
 - de 880 à 915 MHz du mobile vers la base,
 - de 925 à 960 MHz de la base vers le mobile,
 - écart entre les deux fréquences 45 MHz,
 - 174 canaux espacés de 200 kHz.

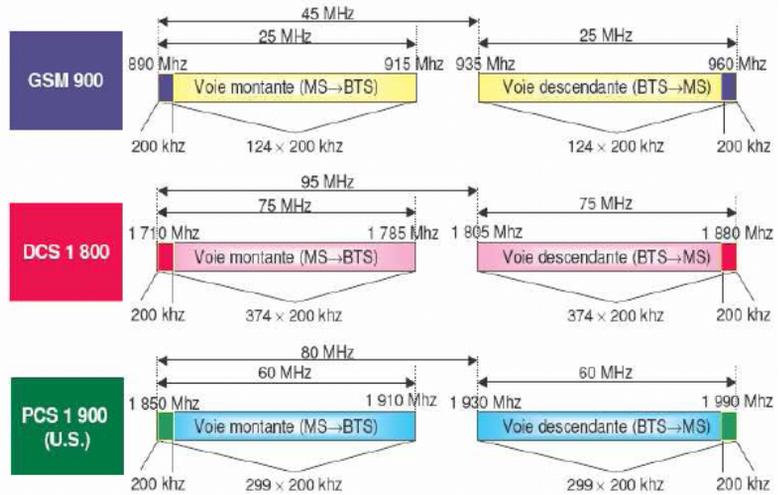
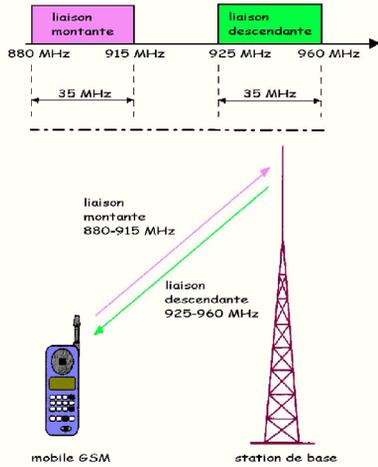


Figure 1.1. Bandes de fréquences pour la liaison entre mobile et station de base dans le GSM.

- Bande DCS (bande de largeur totale 75 MHz)
 - de 1710 à 1785 MHz du mobile vers la base,
 - de 1805 à 1880 MHz de la base vers le mobile,
 - écart entre les deux fréquences 95 MHz,
 - 374 canaux espacés de 200 kHz.

A l'intérieur d'une cellule, on dispose donc d'un certain nombre de fréquences ou porteuses qu'il faut répartir entre les différents utilisateurs. Lors d'une conversation courante, un téléphone mobile n'a pas besoin du canal de transmission en permanence, grâce aux techniques de compression de débit. Pour

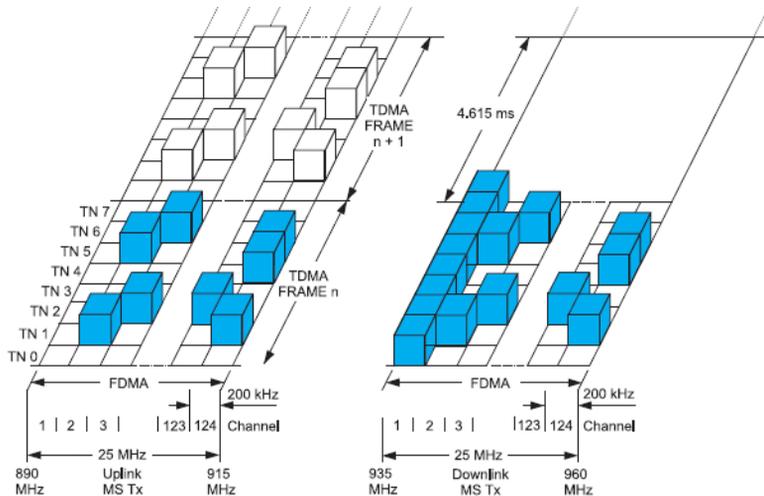


Figure 1.2 interface air du GSM : trame TDMA [3].

exploiter les instants libres, les choix de conception suivants ont été adoptés (voir figure 1.2) :

- chaque porteuse est divisée en 8 intervalles de temps appelés *time-slots*. La durée d'un slot a été fixée pour le GSM à 7500 périodes du signal de référence fourni par un quartz à 13 MHz qui rythme tous les mobiles GSM :

$T_{slot} = 7500/13 \text{ MHz} = 0,5769 \text{ ms}$ soit environ **577 μs** ,

- sur une même porteuse, les slots sont regroupés par paquets de 8 qui constituent une **trame (frame) TDMA**. la durée de la trame est donc : $TTDMA = 8 T_{slot} = \mathbf{4,6152 \text{ ms}}$.

Chaque porteuse GSM (ou DCS) est identifiée de manière unique par un numéro n , désigné par le sigle ARFCN, codé sur 10 bits conformément au plan suivant, où la fréquence de la voie descendante est exprimée en MHz:

- pour $1 \leq n \leq 124$ $f = 935 + (0,2 \times n)$ (GSM)
- pour $975 \leq n \leq 1024$ $f = 935 + (0,2 \times (n-1024))$ (GSM étendu EGSM)
- pour $512 \leq n \leq 885$ $f = 1805,2 + (0,2 \times (n-512))$ (DCS 1800)

Exemple : pour $n=10$, voie descendante à $935 + (0,2 \cdot 10) = 937 \text{ MHz}$ et voie montante à $f_d - 45 = 892 \text{ MHz}$.

Remarque : Ce numéro de canal peut changer durant une communication lorsque la qualité devient insuffisante (**saut de fréquence** ou **frequency hopping**).

Un mobile GSM en communication n'utilisera qu'un **time-slot**, ce qui permet de faire travailler jusqu'à 8 mobiles différents sur la même fréquence de porteuse. Le signal radio émis dans un **time-slot** est souvent appelé **burst**. Les slots sont numérotés par un indice TN qui varie de 0 à 7.

Durant une communication téléphonique, le mobile GSM reçoit des informations de la station de base et émet des informations vers celle-ci:

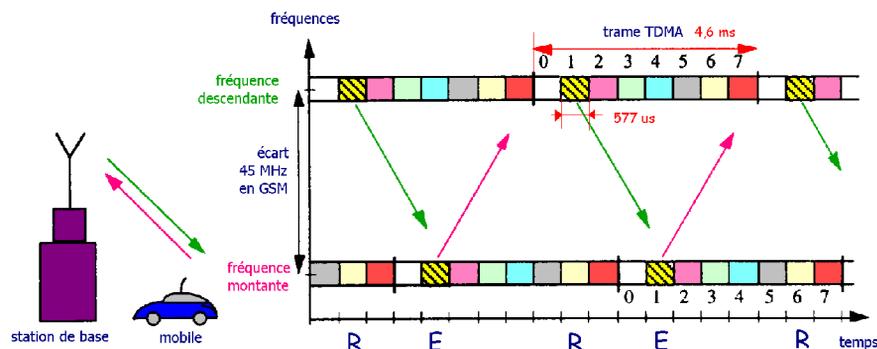


Figure 1.3 Mobile en conversation sur le time-slot 1 [3].

- ces échanges se font sur deux fréquences différentes et n'ont pas lieu au même moment.
- au niveau du mobile, l'émission et la réception sont décalées dans le temps de 3 time-slots.
- pour conserver la même numérotation des slots, le début de la trame TDMA du mobile est décalée de 3 time-slots / début de la trame de la base.

Le mobile reçoit donc le signal émis par la base sur la fréquence descendante f durant un time slot soit $577 \mu\text{s}$, puis 3 time-slots soit $1,7 \text{ ms}$ plus tard (figure 1.3), émet son signal vers la station de base sur la fréquence montante plus basse ($f-45 \text{ MHz}$ pour le GSM).

1.2.2.1 Canaux logiques

Pour renforcer l'interface radio, qui est le maillon faible de la chaîne de transmission, un certain nombre de fonctions de contrôle ont été mises au point pour rattacher le mobile à une BTS favorable, pour établir une communication, surveiller son déroulement et assurer les handovers. Ces fonctions de contrôle engendrent des transferts de données additionnels : remontées des mesures, messages de contrôle...etc. Plusieurs **canaux logiques** ont été ainsi définis pour les différents types de fonction (veille, scrutation, mesures, contrôle...) ; ils forment une architecture complexe qu'il est nécessaire de connaître pour comprendre le fonctionnement d'un mobile pendant les différentes phases de communication ou pendant

sa veille. Ils n'existent que sur l'interface radio et perdent ensuite toute leur signification sur les autres interfaces du système [5] : Abis, Ater, A, ...etc. Sur l'interface radio il faut :

- diffuser des informations système : **Broadcast Channels**,
- prévenir les mobiles des appels entrants et faciliter leur accès au système : **Common Control Channel**
- contrôler les paramètres physiques avant et pendant les phases actives de transmission : **FACCH** et **SACCH**,
- fournir des supports pour la transmission de la signalisation : **SDCCH**.

On n'utilise pas un canal physique plein pour chacune de ces tâches : ce serait gâcher de la ressource radio car elles ne nécessitent pas, en général, un débit comparable à celui de la voix codée (**TCH**).

Pour introduire plus de souplesse et allouer moins d'un slot par trame, on définit des structures de **multitrames** (figure 1.4). La structure de multi-trames est définie comme une succession d'un slot donné sur des trames TDMA successives, c'est-à-dire sur un canal physique. Entre deux slots d'une multi-trames, il s'écoule donc 4,615 ms.

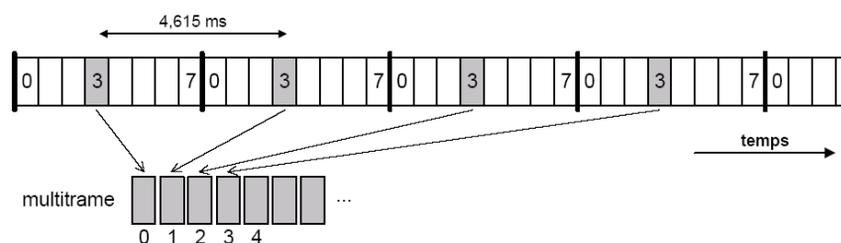


Figure 1.4 structure d'une multi-trames GSM [5].

Chaque multi-trame transporte, avec une périodicité bien définie, un certain type d'informations de contrôle ou de signalisation. Cet ensemble de time-slots forme un canal logique. Certaines multi-trames sont définies à 26 trames, d'autres à 51 trames.

Il y a deux types de canaux dans l'interface air: les canaux physiques et les canaux logiques. Un "canal physique" est constitué par la répétition périodique d'un slot dans la trame TDMA d'une fréquence particulière. Un "canal logique", renvoie au type d'information spécifique qui est transporté par le canal physique. En effet, l'interface radio représente la partie délicate de la chaîne de transmission et le système doit faire face aux différents problèmes du lien mobile-réseau au niveau de la propagation (atténuation, évanouissements, interférences...), mais aussi au niveau de la gestion du réseau : il est nécessaire d'avoir des fonctions de contrôle pour que le mobile se rattache à la station de base la plus favorable, pour établir et surveiller le déroulement d'une communication ou encore assurer le *handover*. L'utilisation de canaux logiques va permettre une utilisation efficace des ressources radio et une qualité de service satisfaisante. Parmi ces canaux on distingue les canaux dédiés (TCH et SDCCH), c'est à dire alloués à un mobile, et les autres canaux qui sont partageables entre mobiles et servent au contrôle et à la signalisation. Alors que les canaux physiques ne font que transporter des informations, quelles qu'elles soient, les canaux logiques permettent de distinguer les différents types d'informations circulant dans le système. Il y a deux grandes classes de canaux logiques, les canaux dédiés et les canaux communs (figure 1.5):

- Un canal logique dédié fournit une ressource réservée à un seul mobile (point à point). On lui attribue une paire de slots (montant et descendant) où il est le seul à recevoir et à émettre dans la structure de multitrame. Les canaux dédiés sont duplex.
- Un canal logique commun est simplex (attribué sur une voie seulement suivant les cas) et partagé par un ensemble de mobiles. Dans le sens descendant, les informations y sont diffusées à plusieurs mobiles à l'écoute. Ces données peuvent concerner le système dans son ensemble ou un mobile en particulier qui, par exemple, peut recevoir un appel entrant. Dans le sens montant, la fonction d'accès multiple est remplie. Chacun peut émettre et les collisions sont résolues par les méthodes classiques de résolution de contention (Aloha slotté).

Tous les canaux logiques n'ont pas les mêmes besoins, certains se contentent de faibles débits alors que d'autres sont beaucoup plus gourmands en ressources. Afin de gérer les débits et de définir une périodicité sur les canaux logiques, on a créé deux structures de Multi-trames, rythmant les périodes d'allocation des ressources physiques requises pour les canaux logiques (figure 1.6). La Multi-trame 26, composée de 26 trames TDMA, d'une durée de 120 ms, et la Multi-trame 51, composée de 51 trames TDMA, d'une durée de

Broadcast Channel non dédié diffusion ↓	Frequency Correction Channel (FCCH)	calage sur fréquence
	Synchronization Channel (SCH)	synchronisation en temps & identification de la BTS
	Broadcast Control Channel (BCCH)	information système
Common Control Channel non dédié diffusion ↓ et accès multiple ↑	Paging Channel (PCH) ↓	recherche du mobile en cas d'appel entrant
	Random Access Channel (RACH) ↑	accès aléatoire du mobile
	Access Grant Channel (AGCH) ↓	allocation de ressources
	Cell Broadcast Channel (CBCH) ↓	diffusion de messages courts
Dedicated Control Channel dédié ↑↓	Stand-Alone Dedicated Control Channel (SDCCH)	signalisation
	Slow Associated Control Channel (SACCH)	supervision lente de la communication
	Fast Associated Control Channel (FACCH)	signalisation rapide (handover)
Traffic Channel (TCH) dédié ↑↓	Full rate, Enhanced Full Rate & Half Rate	parole
	débit utilisateur < 14,4 kbits/s	données

Figure 1.5 Types de canaux logiques et leurs fonctions dans le GSM.[5]

235,8 ms. Pour gérer ces deux multi-trames, on a créé la Super-trame, structure rassemblant 26 Multi-

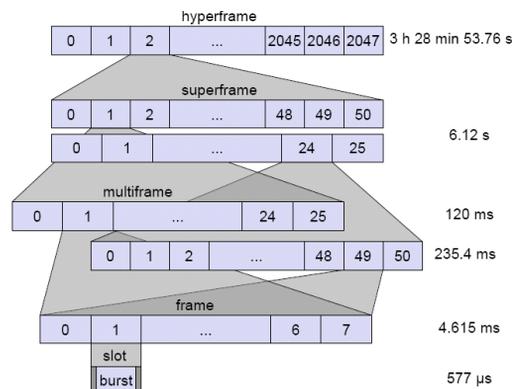


Figure 1.6 structure hiérarchique de trames

trame 51 ou indifféremment 51 Multi-trame 26, d'une durée de 6,12 s. La Super-trame n'a pas de fonction essentielle, son seul rôle est de contribuer à la définition de l'Hyper-trame, qui est composée de 2048 super-trames, avec une période de 3h 28mn 53s 760 ms. Elle va servir de base à la création d'un code de temps : "Time code", chaque trame TDMA composant l'Hyper-trame se voit attribuer un numéro : "Frame Number". Ce compteur va permettre au mobile de se synchroniser finement avec la cellule en se verrouillant dessus, le "Time code" démarre au même instant pour toutes les fréquences d'une même BTS. Le compteur est aussi utilisé pour le chiffrement dans l'algorithme A5.

Chaque BTS est équipée pour travailler sur un certain nombre de canaux, en général 5 ou 6, qui sont autant de paires de fréquences émission-réception. Toute BTS émet en permanence des informations sur son canal **BCH** (Broadcast Channel) appelé aussi **voie balise**. Ce signal constitue le lien permanent

reliant mobile et station de base de la mise en route du mobile jusqu'à sa mise hors service, qu'il soit en communication ou non.

Le fonctionnement du mobile se décompose en 2 phases [3] :

⇒ **mobile en veille** : le mobile échange avec sa base des signaux de contrôle **sur la voie balise** (émission en slot 0 à f_1 , réception en slot 0 à $f_1 + 45$ MHz)

Le niveau de la voie balise (BCH) est connu et sert pour un certain nombre de fonctions de contrôle (figure 1.7) :

- à la mise en route du mobile, son récepteur scrute la bande GSM pour chercher le signal BCH de niveau le plus élevé. C'est avec la station de base correspondante que le mobile se mettra en communication.
- ce signal contient des informations concernant les opérateurs et les fréquences balises des cellules voisines.
- ce signal véhicule les messages qui seront affichés sur l'écran du mobile.
- toutes les 15 secondes si le signal reçu est fort et toutes les 5 secondes s'il est faible, le récepteur écoute les balises des cellules voisines pour détecter un changement de cellule.
- l'émission balise n'occupe le canal de transmission que dans le sens base - mobile. La liaison montante pourra donc être utilisée par le mobile pour signaler son désir de se connecter au réseau pour une communication (RACH : random access channel).

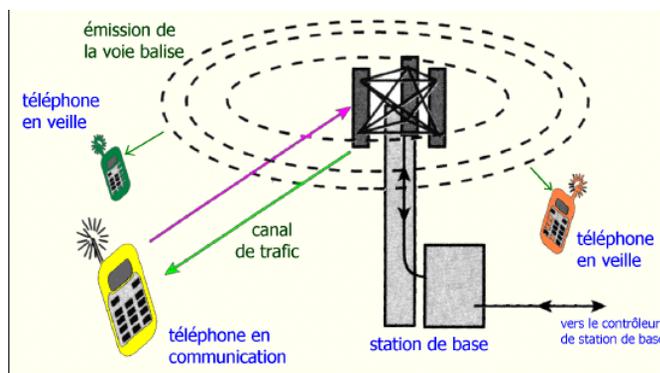


Figure 1.7 Par la voie balise la SB diffuse ses informations vers tous les mobiles de la cellule.[3]

⇒ **mobile en communication** : le mobile échange avec la base des signaux de parole et de contrôle **sur la voie de trafic** (émission en slot i à f_2 , réception en slot i à $f_2 + 45$ MHz)

- il émet et reçoit maintenant sur une nouvelle paire de fréquences allouées par la base pour la durée de la communication : c'est le TCH (Traffic Channel)
- parallèlement à cette activité principale, il écoute périodiquement les voies balises de la cellule et des cellules voisines pour détecter une variation de niveau lui indiquant un changement de cellule.

Les différents utilisateurs d'un système cellulaire sont à des distances variables de leur station de base et endurent des délais de propagation variables. Or l'onde électromagnétique se propage à la vitesse de la lumière soit $c = 300\,000$ km/s. Cette vitesse est très élevée, mais pas infinie, et les retards engendrés par la distance se font sentir sur le timing puisqu'une distance de 30 km cause un retard de **100** μ s. Prenons l'exemple de deux mobiles MS1 et MS2 appartenant à la même cellule ; le premier, MS1 est en limite de cellule alors que le second, MS2 est situé près de la station de base (figure 1.8).

On suppose que ces deux mobiles utilisent des slots consécutifs sur la même porteuse : MS1 émet dans le slot 1, MS2 dans le slot 2 :

- en l'absence de compensation du temps de propagation, les bursts émis par chacun des mobiles se chevaucheraient au niveau du récepteur de la BS.
- pour pallier à cette difficulté, la station de base va compenser ce retard en gérant un paramètre TA (Time Advance) correspondant au temps de propagation aller-retour.

- le mobile éloigné doit avancer l'émission de chacun de ses bursts par rapport à l'instant nominal de début de slot
- la distance entre mobile et station de base étant susceptible de varier en permanence, ce paramètre TA est réajusté à chaque trame et pourra prendre une valeur comprise entre 0 et 63.

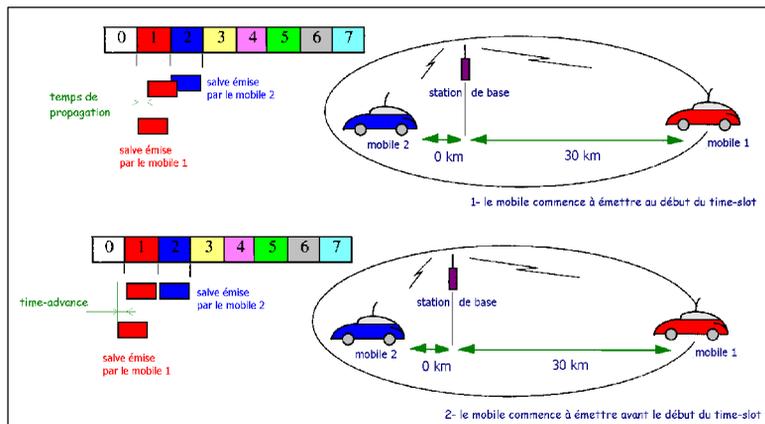


Figure 1.8 Importance du paramètre de Time Advance.

Remarque : la détermination du paramètre TA permet à la base de connaître la distance à laquelle se trouve le mobile. Par triangulation avec une deuxième station de base, on pourra donc déterminer la position exacte d'un mobile.

Pendant un échange de données vocales, le mobile continue l'écoute des balises des cellules voisines pour détecter un éventuel changement de cellule. Cette écoute se fait entre l'émission et la réception du burst suivant. Vu le faible temps disponible, le mobile ne pourra faire qu'une mesure de niveau. Pour décoder les informations provenant de la balise d'une cellule voisine, il lui faut davantage de temps, surtout qu'il faut « attraper » le time-slot 0 qui contient les informations recherchées. C'est la raison pour laquelle le mobile s'arrête d'émettre et de recevoir toutes les 26 trames (slot idle) ce qui lui permet d'écouter et de décoder le canal de contrôle d'une cellule voisine.

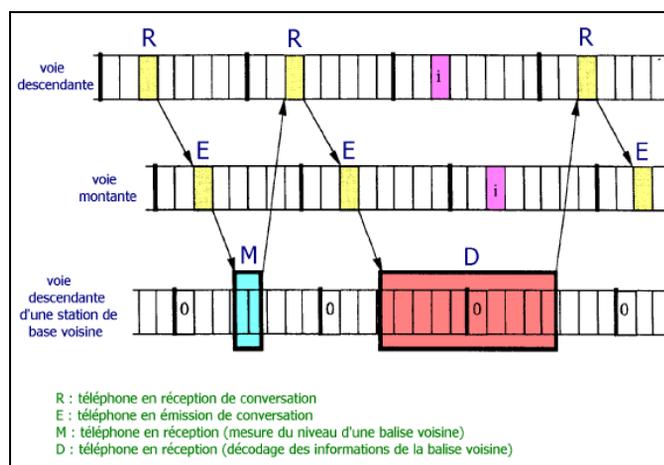


Figure 1.9 Mesure de niveau et décodage des BCCH des cellules voisines

Lors d'une communication, le mobile peut effectuer sur demande du réseau, une augmentation ou une diminution de son émission par pas de 2 dBm toutes les 60 ms. Ce changement sur l'initiative du réseau, résulte d'une lecture des remontées de mesures effectuées par le mobile toute 480 ms sur le canal SACCH (soit toutes les 4 multi_trame26 = 104 trames TDMA).

1.2.3 Principe de réutilisation des ressources

Le principe de la réutilisation des fréquences repose sur l'atténuation que subissent les ondes radio lorsqu'elles se propagent dans l'atmosphère. Lorsqu'on se trouve assez loin d'un émetteur, le signal envoyé

par celui-ci est très faible. On peut alors utiliser la même fréquence que l'émetteur lointain sans crainte d'interférences, le signal local étant beaucoup plus puissant que le signal lointain. Le concept de motif cellulaire a donc été introduit pour permettre la réutilisation d'une même fréquence dans des endroits différents. Un "motif", aussi appelés cluster, est le plus petit groupe de cellules contenant une et une seule fois l'ensemble des canaux radio (fréquences).

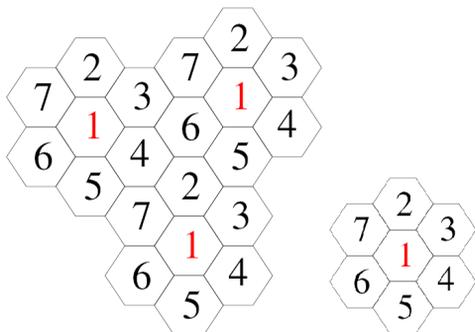


Figure 1.10 un motif élémentaire (à droite) et un ensemble de motifs dans un réseau (à gauche).[7]

Le GSM utilise donc un réseau maillé, formé d'émetteurs disséminés sur la zone à couvrir. Deux émetteurs voisins utilisent des fréquences différentes, mais des émetteurs éloignés réutilisent les mêmes fréquences selon le principe énoncé plus haut. En pratique, le terrain est "découpé" en petites zones, appelées cellules, caractérisées chacune par une fréquence précise. On réunit un certain nombre de cellules utilisant des fréquences différentes pour former un motif. On répète alors ce motif pour couvrir tout le territoire, permettant ainsi à partir d'un nombre de fréquences limité de mettre en place un grand nombre d'émetteurs.

Chaque émetteur nécessite une infrastructure complexe pour fonctionner et dialoguer tant avec l'ensemble du réseau qu'avec les téléphones mobiles présents sur sa zone de service. On appelle cette infrastructure "station de base". Un réseau de radiotéléphonie cellulaire se compose donc d'un ensemble de stations de base réparties sur la zone géographique à couvrir.

1.2.4 Code de couleur BSIC

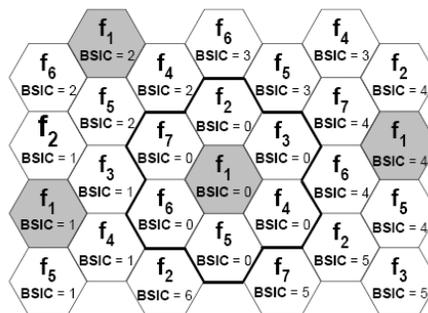


Figure 1.11 Utilisation des codes BSIC dans un motif à 7

La même fréquence peut être utilisée pour supporter la voie balise de deux stations suffisamment éloignées. Les deux stations ne se brouillent pas sur leurs zones de service respectives mais un mobile situé à mi-distance peut recevoir alternativement l'une ou l'autre station avec un niveau de champ suffisant. Afin de différencier les deux stations, on utilise le code de couleur BSIC. Le couple (fréquence, BSIC) permet sur une zone donnée de déterminer parfaitement une cellule. A l'intérieur d'un motif, on utilise le même BSIC. Ainsi, les cellules voisines (cellules de fréquences de voie balise identique) ne font pas partie du même motif.

1.2.5 L'entrelacement et construction de salves (bursts)

Lors de la transmission, les erreurs arrivent sur un groupe de bits consécutifs (par salves ou *bursts*). Le but de l'entrelacement est de séparer les erreurs pour pouvoir les corriger plus facilement. Pour se faire, les bits sont introduits ligne par ligne dans une matrice à L lignes et n colonnes pour être lus, ensuite,

par colonnes. La première colonne contiendra, ainsi, les numéros de bits (0,n,2n,...Ln). Le paramètre n est appelé le facteur ou degré d'entrelacement. Dans le cas de l'entrelacement des canaux de la voix, les 456 bits obtenus après le codage (des canaux) sont divisés en 8 blocs de 57 bits. Le premier bloc contient les bits de numéros: 0,8,16,24,...448. La figure 1.12 représente la chaîne de recombinaison que subissent les 8 blocs issus du codage de la parole, et illustre le schéma de construction des *bursts*. On peut voir, sur

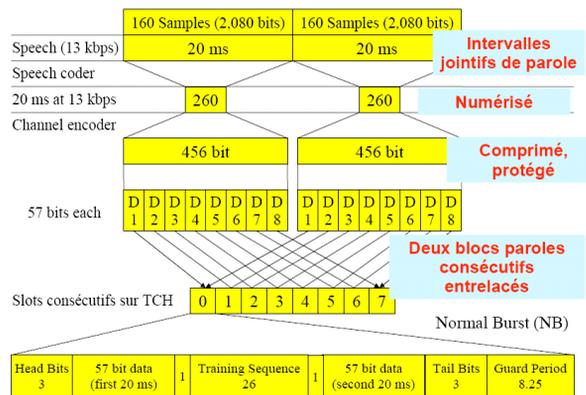


Figure 1.12 Entrelacement et construction de salve

le schéma, qu'il y a 2 séquences de 57 bits dans un burst, et que pour transmettre notre échantillon, il faut utiliser 4 bursts. Donc notre message sera envoyé sur 4 trames.

1.2.6 Contrôle par la base de la puissance d'émission

La station de base contrôle de nombreux paramètres du mobile, dont la puissance d'émission particulièrement. A partir des estimations des niveaux de puissance et des taux d'erreur montants et descendants, qui sont reportés par le mobile dans des messages de mesure, le BSS (ou BTS) détermine les niveaux de puissance adéquats. L'ajustement du niveau émis est fait de façon à minimiser la puissance requise par l'émetteur tout en conservant la qualité de la communication. Les deux objectifs majeurs recherchés, sont la **diminution du niveau d'interférence** due aux canaux adjacents et **l'augmentation de l'autonomie** des mobiles. En conséquence, l'amplificateur de puissance RF de tout mobile GSM doit être équipé :

- d'une entrée commandant la puissance de sortie
- d'un dispositif de mesure de la puissance émise

Dans les mobiles actuels, la mesure de la puissance est faite soit par le contrôle du courant absorbé par l'amplificateur de puissance, soit à l'aide d'un ensemble coupleur directif-détecteur Schottky.

1.3 TAR mise en œuvre dans les réseaux GPRS

1.3.1 Présentation

Le GPRS ne constitue pas à lui tout seul un réseau mobile à part entière, mais une couche supplémentaire rajoutée à un réseau GSM existant. Il peut donc être installé sans aucune licence supplémentaire. Ceci signifie que tous les opérateurs qui disposent d'une licence GSM peuvent faire évoluer leur réseau vers le GPRS. De plus, le GPRS utilise les bandes de fréquences attribuées au GSM, c'est à dire une bande dans les 900 MHz, une autre dans les 1800 MHz et enfin une troisième, aux USA, dans les 1900 MHz. Les opérateurs GSM actuels ont de fait un quasi monopole sur le GPRS, ce qui n'est pas le cas pour l'UMTS.

Le GPRS, appelé aussi GSM 2+, repose sur la transmission en mode paquet, c.a.d qu'il n'y a pas de circuit physique dédié à la station mobile GPRS. Un canal physique est établi, dynamiquement, seulement lorsque les données ont besoin d'être transférées. Une fois les données transmises la ressource peut être réallouée à une autre station, pour une utilisation plus efficace du réseau.

Conçu pour réutiliser au maximum les infrastructures GSM existantes, le déploiement du GPRS nécessite la mise en place d'une infrastructure réseau basée sur la commutation de paquets et l'introduction de passerelles pour s'adosser aux réseaux GSM existants(voir figure 1.13).

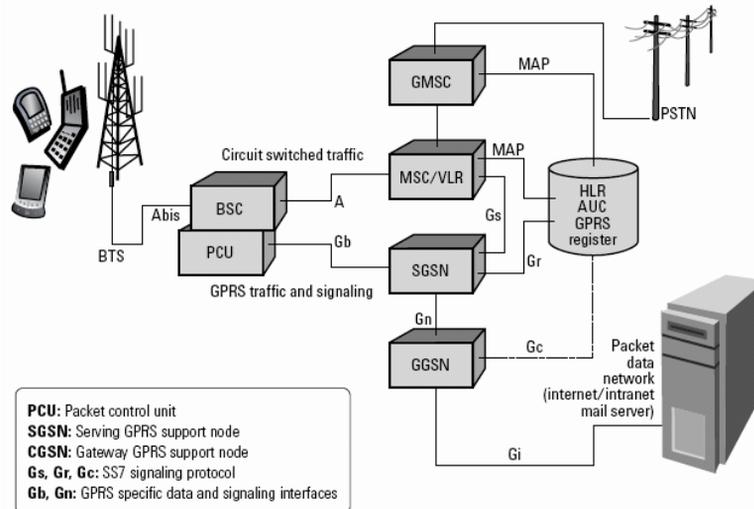


Figure 1.13 Architecture du réseau GPRS.

Cette technologie, capable de fournir des débits par utilisateur allant jusqu'à 115 kb/s (contre 9,6 kb/s pour le GSM), offre les fonctionnalités intéressantes suivantes :

- plusieurs canaux peuvent être alloués à un même utilisateur ;
- ces mêmes utilisateurs peuvent partager un même canal ;
- le débit est indépendant des liens montants et descendants.

Avec le GPRS, on dispose d'un débit compris entre 40 et 115 kbit/s, selon le nombre de canaux virtuels ou "time slots" utilisés, et le schéma de codage choisi (CS1 à CS4). Ce dernier agit sur la compression des données comme un multiplicateur de débit. En mode multislots 3+1 (trois slots pour la transmission dans le sens réseau vers portable, et un slot pour le sens portable vers réseau), on atteint un débit de 40 kbit/s avec un schéma de codage CS2. En mode multislots (8+1) avec le codage CS4, on atteint en pratique 115 kbit/s (en théorie 175 kbit/s). Si, comme cela semble être la volonté actuelle des opérateurs, on réutilise l'infrastructure GSM existante, en conservant notamment le réseau des stations de base (BTS) actuelles mises à niveau par logiciel, seules les versions CS1 et CS2 seront implantées. Pour les versions CS3 et CS4, des transformations importantes de l'infrastructure devront être opérées.

Les données utilisateurs sont transférées de manière transparente entre le terminal mobile et les réseaux de données externes par une technique de « **tunneling** », avec le protocole **GTP** (*GPRS Tunneling Protocol*).

Il faut savoir que ce protocole est utilisé dans GPRS pour la transmission des paquets de données utilisateur et de signalisation entre entités GPRS, via l'établissement de tunnels de communication. Le GTP utilise soit TCP, soit UDP, selon qu'une connexion fiable est nécessaire ou non.

On peut donc résumer les intérêts principaux du GPRS dans les temps d'accès réduits, de l'ordre d'une seconde pour commencer un transfert de données, un débit plus élevé qu'en GSM, un mode de commutation par paquets permettant d'utiliser les ressources radios, et enfin la possibilité de facturation en fonction du volume de données transféré plutôt qu'en fonction du temps de connexion.

1.3.2 Interface air du GPRS

La principale nouveauté, coté interface air, est l'allocation dynamique de canal; ce qui autorise des transferts à débits variables. Cette nouveauté est caractérisée par les points suivants:

- Le GPRS est capable d'allouer :
 - le même time slot à plusieurs utilisateurs
 - partage des ressources lors des périodes chargées

- plusieurs time slots à un seul utilisateur (maximum 8)
 - atteinte de hauts débits (maximum théorique 171,2 kbits/s)
- Les canaux UL et DL sont réservés séparément et il est possible que les services GPRS et GSM utilisent alternativement le même time slot.

Les utilisateurs GPRS partagent avec ceux du GSM la même trame TDMA, augmentant ainsi les exigences en capacité. Dans une certaine mesure, le GPRS tient compte de cette augmentation de demande de capacité en multiplexant plusieurs utilisateurs sur les mêmes canaux physiques (time slots). De plus, l'interface air du GPRS alloue les ressources (time slots) pour la voix et les PDCH (packet data channels), de manière dynamique. Certains canaux physiques sont configurés pour les données, mais peuvent être reconfigurés pour la voix au besoin.

Le GPRS doit aussi allouer des ressources pour la signalisation et le contrôle du trafic. Vu que le GPRS a ses propres paramètres pour l'accès au réseau et le contrôle des appels, il a besoin de canaux séparés pour les fonctions de contrôle commun de diffusion (tels que le paging, random access, et access grant) ainsi que des canaux de trafic associés (similaire à SACCH).

1.3.2 Allocation des ressources

Avant de décrire les options pour l'allocation des ressources, revoyons d'abord un peu de terminologie. Les trames GSM ou GPRS consistent en huit timeslots. Le timeslot 0 est réservé pour le canal GSM BCH (Broadcast CHannel). Les timeslots CS (Circuit Switched) sont utilisés pour les appels en mode commutation de circuits GSM (voix ou données). Le PBCH (Packet Broadcast Channel) et le P-Data (Packet-Data) sont utilisés pour les canaux GPRS.

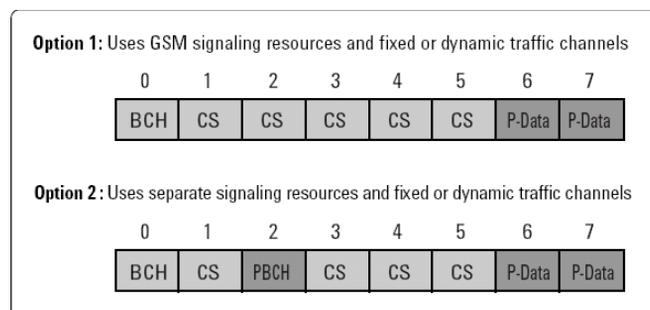


Figure 1.14 Canalisation de l'interface air du

Deux options sont disponibles pour l'établissement des canaux de l'air interface du GPRS, (illustrées dans la figure 1.14) :

- Option 1: utilise les ressources de signalisation du GSM, mais établit pour le control du trafic des canaux packet-data séparés. Les canaux de trafic peuvent être fixes ou dynamiques.
- Option 2: sépare entièrement les ressources du GPRS de celles du GSM. Il y a plusieurs configurations possibles avec cette option. Un PBCH peut être utilisé pour transporter l'information du BCH-GPRS, des canaux de control communs, des canaux GPRS packet-data, et les canaux de trafic associés. Si les canaux packet-data ne sont pas transportés, le PBCH ou, si des ressources PDCH additionnelles sont nécessaires, des timeslots séparés peuvent être configurés

1.3.3 Canaux logiques GPRS

Les canaux logiques sont de fonctions prédéfinies supportées par des trames dans un canal physique. Les canaux physiques GPRS transportent en général deux types d'informations: la signalisation de control pour l'établissement et le maintient d'un service GPRS, et le trafic de données de l'utilisateur. Ainsi, les canaux logiques GPRS peuvent être classés comme des canaux de control paquet et des canaux de trafic paquet. Les canaux GPRS de control paquet sont de plus divisés en des fonctionnalités de control

commun, et des fonctions de control de diffusion, d'une manière similaire au GSM, comme résumé ci-après :

PCCCH (Packet Common Control Channel) est un ensemble de canaux logiques utilisés pour la signalisation commune entre la station mobile et la station de base.

- PRACH (Packet Random Access Channel),
- PPCH (Packet Paging Channel),
- PAGCH (Packet Access Grant Channel),
- PNCH (Packet Notification Channel).

PBCCH (packet broadcast control channel) est utilisé pour diffuser des packet-data d'information système à tous les mobiles GPRS dans une cellule. Le PBCCH pourrait ne pas être présent pour certaines combinaisons de canaux, cas où le BCCH sera utilisé pour diffuser l'information système paquet.

PTCH (packet traffic channel) consiste en les sous canaux, associés à du trafic, suivants:

- PDTCH (Packet Data Traffic Channel),
- PACCH (Packet Associated Control Channel),
- PTCCH (Packet Timing Advance Control Channel).

1.3.4 Association des canaux logiques aux canaux physiques

Les canaux logiques que nous venons de définir pour le GPRS ne demandent pas nécessairement des ressources physiques séparées. Les canaux logiques sont associés à des canaux physiques en utilisant la technique du "multi tramage"(multiframeing). La multiframe de base du GPRS[13] est définie par l'occurrence d'un même canal physique dans 52 trame TDMA successives(numérotées de 0 à 51), et non 26 ou 51 comme dans le GSM. La multiframe est organisée comme suit :

- 12 X 4 = 48 timeslots radio pour le transport des données et de la signalisation;
- 2 timeslots de contrôle de l'avance en temps PTCCH;
- 2 timeslots *idle*.

Les 48 timeslots radio sont divisés en 12 blocs radio. Chaque bloc contient 4 timeslots, qui sont pris dans 4 trames TDMS successives. Contrairement au GSM, l'unité élémentaire allouée en GPRS est un bloc, soit 4 timeslots GSM. Cette unité correspond à la taille des blocs RLC-MAC. Un bloc RLC-MAC se transmet donc exactement dans un bloc de la multiframe GPRS, soit 4 PDCH sur 4 trames successives (figure 1.15).

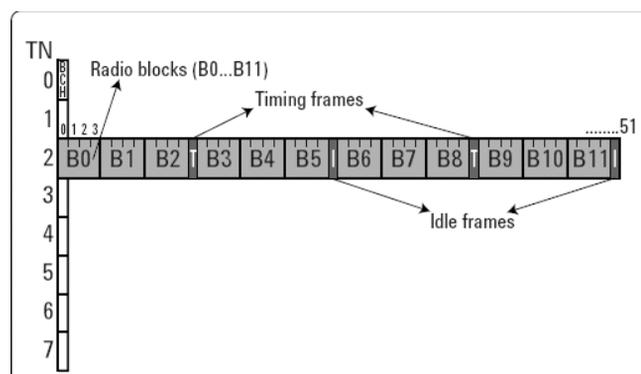


Figure 1.15 Association des canaux canaux logiques aux canaux physiques.

Un bloc radio peu transporter un bloc RLC-MAC de données ou de signalisation, indépendamment du bloc RLC-MAC transporté par le bloc précédent. Les messages RLC-MAC de contrôle de l'avance en temps, ou PTCCH, subissent un traitement particulier. Ils sont transmis dans 4 slots répartis sur deux multitrames (slots des trames 12 et 38).

L'ETSI définit trois combinaisons de canaux logiques qui peuvent être associés à un seul canal physique en utilisant la multi trame à 52 trames.

1.4 TAR utilisant la communication par étalement de spectre

1.4.1 Présentation

1.4.1.1 Justification théorique de l'étalement de spectre

L'étalement de spectre tire sa justification théorique du théorème de la capacité d'un canal de communication, de Shannon, représenté par la formule :

$$\text{Capacity} = \text{Bandwidth} \times \log_2 (1 + S/N) \quad (1.1)$$

Cette équation exprime la relation qui lie :

- la capacité (Capacity) du canal, en bits par seconde (bps),
- la bande passante (Bandwidth) nécessaire, en Hz,
- et le rapport (S/N) de la puissance du signal à celle du bruit.

Plus explicitement[10], cela signifie que plus on a de bande passante, meilleur est le rapport S/N, plus nous pouvons injecter de débit dans le canal (cela coule de source). Cependant, considérons une situation où le signal est plus faible que le bruit. Sous ces conditions la relation (1.1) devient plus simple, et peut être approximée par :

$$\text{Capacity} / \text{Bandwidth} = 1.44 * S/N \quad (1.2)$$

Ce que dit cette relation est que pour préserver la capacité du canal de communication on peut échanger une bande passante plus large contre un meilleur rapport signal à bruit, ou vice versa. C'est à dire, qu'on

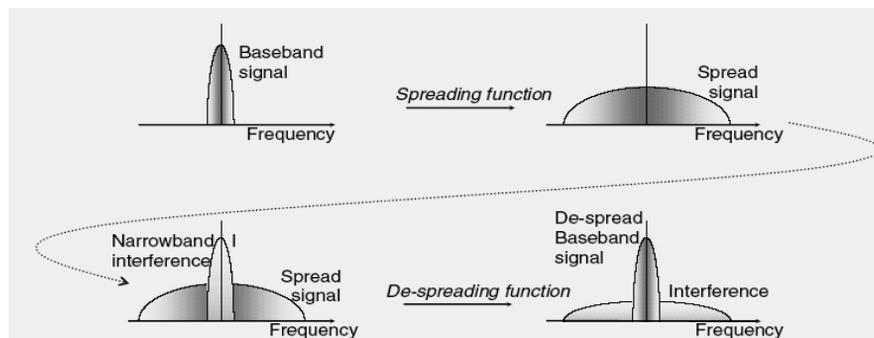


Figure 1.16 Une explication de l'étalement de spectre.

source: Steve Rackley. *Wireless Networking Technology From Principles to Successful Implementation*. Newnes 2007.

peut maintenir ou même augmenter la performance de la communication (Capacité élevée) en autorisant ou injectant plus de bande passante (Bandwidth élevée), même lorsque la puissance du signal est en dessous du niveau du bruit.

Si nous pouvons donc trouver une manière d'encoder nos données dans un signal de large bande passante, nous pourrions alors les transmettre, sans erreurs, dans des conditions où le bruit serait beaucoup plus puissant que le signal utilisé. Par conséquent l'énergie utilisée pour transmettre le signal serait étalée sur la bande passante, et ressemblerait à du bruit. C'est cette idée, simple, qui est derrière les techniques de l'étalement de spectre (figure 1.16).

Il y a plusieurs raisons à l'étalement du spectre, et si c'est fait proprement on pourrait en tirer multiples bénéfices, dont:

- La possibilité d'un adressage sélectif. Si l'étalement du signal est fait de manière correcte, utilisant une méthode d'encodage propre, alors le signal ne peut être décodé que par un récepteur qui connaît le code de l'émetteur. Donc, grâce au code de l'émetteur, nous pouvons adresser un

récepteur spécifique et vice versa. C'est là le principe du Code Division Multiple Access (CMDA).

- Le partage de la bande passante. Avec un choix judicieux des codes de modulation, il est tout à fait faisable d'avoir plusieurs paires d'émetteurs/récepteurs occupants la même fréquence. Cette aptitude à partager la bande passante est très appréciée pour les besoins des utilisateurs de diffusion commerciale, d'autant plus que le spectre radio est fragmenté.
- Protection contre les mises à l'écoute. Si un écouteur indiscret ne connaît pas le code de modulation d'une transmission par étalement de spectre, tout ce qu'il entendra sera plus de bruit, que quelque chose à espionner. D'où, une immunité contre les interceptions des communications.
- Immunité contre les interférences. Si un signal radio externe interfère avec une transmission par étalement de spectre, il sera rejeté par le mécanisme de démodulation d'une façon similaire au bruit. Nous retournons ainsi au point de départ de cette discussion, qui est que les méthodes d'étalement de spectre peuvent fournir d'excellent taux d'erreurs même avec des signaux qui ne sont pas trop forts.
- Difficulté de détection. Puisque la liaison de l'étalement de spectre met beaucoup moins de puissance par bande passante qu'une liaison radio conventionnelle, l'étalant sur une plus large bande passante, et nécessitant la connaissance du code de liaison pour la démodulation, les signaux d'étalement de spectre sont difficiles à détecter. Cela signifie qu'ils peuvent coexister avec d'autres signaux plus conventionnels sans causer d'interférences catastrophiques aux liaisons à bandes étroites.

1.4.2 Techniques de l'étalement de spectre [11]

Nous allons présenter deux techniques de base, d'étalement de spectre, abondamment décrites dans la littérature; il s'agit de l'étalement par séquence directe et de l'étalement par saut de fréquences.

1.4.2.1 Etalement par séquence directe

L'étalement de spectre par séquence directe suppose, pour transmettre un signal en bande de base, de multiplier (avant modulation) chaque bit à transmettre par un code pseudo aléatoire de débit supérieur à celui du signal à transmettre. Le résultat binaire de cette opération modulera ensuite l'onde porteuse servant à la transmission effective de l'information. Le signal de sortie s'étalera donc sur une bande passante plus large que celle nécessaire à la transmission du signal d'origine, et ressemblera à du bruit mais contenant l'information.

Cette séquence de codes est en fait un code binaire pseudo aléatoire (souvent appelé "*pseudo-noise*" PN), spécifiquement choisi pour ces propriétés statistiques appréciables. La durée d'un seul bit dans le PN est appelée "*chip*", et le débit binaire du PN appelé "*chip rate*". Le processus de modulation de la fréquence porteuse avec le *code PN* est appelé étalement (*spreading*).

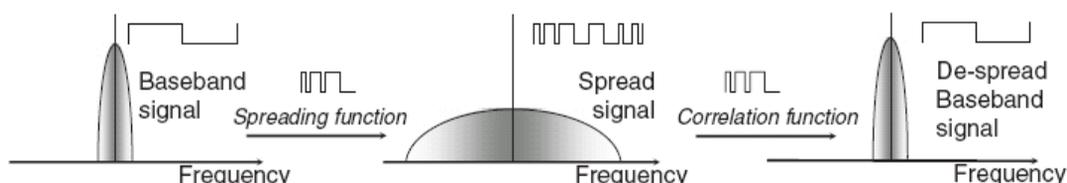


Figure 1.17 Etalement de spectre par séquence directe.

Source :Steve Rackley.*Wireless Networking Technology*
From *Principles to Successful Implementation*. Newnes

Un large éventail de pseudo codes existent et peuvent être utilisés à cette fin. Ces codes doivent, idéalement, être balancés, avec un nombre égal de 0 et de 1 sur toute la longueur de la séquence, comme ils doivent être sûrs en terme de cryptographie. Ces séquences de code sont produites par des générateurs de PN.

Coté réception, l'idée est d'utiliser une opération de corrélation. Cette opération mesure la similitude de la séquence de code PN reçue avec celle d'un PN connu du récepteur. Idéalement, si ces deux séquences sont les mêmes une forte corrélation est détectée, si ces codes sont différents c'est une corrélation faible qui est détectée (figure 1.17).

1.4.2.2 L'étalement par saut de fréquences

Avec cette technique, le spectre de fréquences radio disponible est partagé en un nombre discret de canaux fréquentiels. La fréquence porteuse de l'émetteur est réglée, au rythme du signal à transmettre, sur une valeur de fréquence déterminée d'une manière pseudo aléatoire. L'ordre dans lequel les différentes fréquences sont occupées est fonction d'une séquence de codes pseudo aléatoire PN, tandis que la vitesse

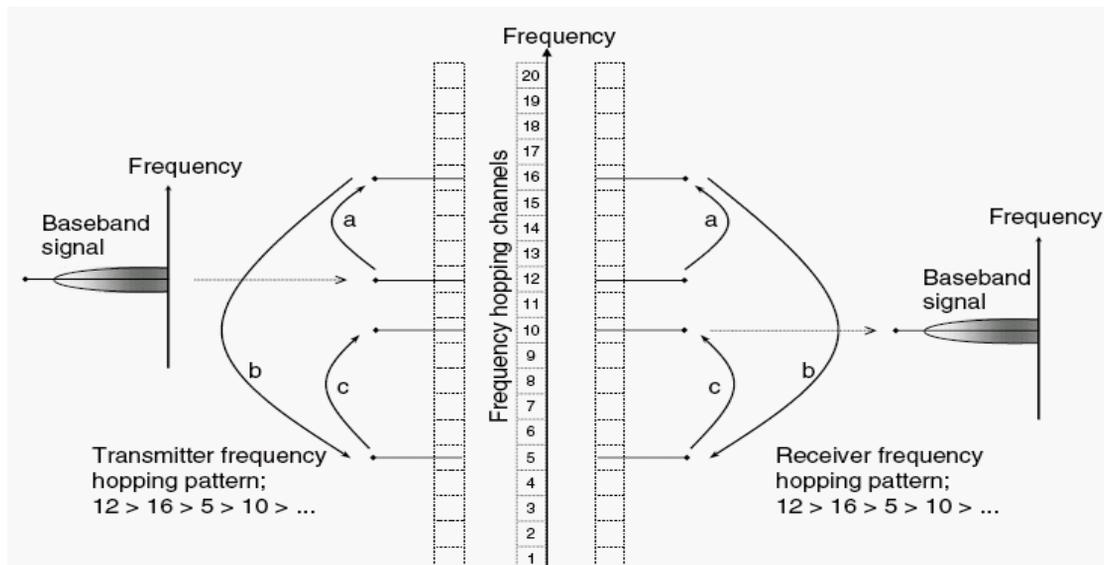


Figure 1.18 Etalement de spectre par saut de fréquence.

Source : Steve Rackley. *Wireless Networking Technology From Principles to Successful Implementation*. Newnes 2007.

des sauts d'une fréquence à une autre est fonction de la vitesse du signal d'origine. La bande passante d'un signal étalé par cette technique est simplement w fois le nombre de slots de fréquence, où w est la bande passante de chaque canal de saut (figure 1.18).

1.4.2.3 Synchronisation

Le succès des systèmes CDMA dépend d'une synchronisation précise. Pour rétablir un signal étalé, on a besoin d'utiliser le même code utilisé pour l'étalement. L'opération se déroule en deux étapes, l'acquisition et la recherche. D'abord l'acquisition (ou synchronisation grossière) est effectuée, où il est question d'aligner le code sous forme de *chips*. Ensuite, la recherche (ou synchronisation fine), où la corrélation est mesurée et le code du récepteur est mis en synchronisme avec celui transmis.

Maintenant que le principe de l'étalement de spectre est présenté nous allons passer en revue quelques standards de communications sans fil qui adoptent ce principe.

1.4.3 IS-95 (cdmaOne)

L'IS-95 (Interim Standard 95) incarne, avec l'IS136, la deuxième génération des communications sans fil aux Etats-Unis, destinés à concurrencer le système européen GSM.

L'IS-95 utilise la méthode d'accès CDMA, qui permet de grouper les utilisateurs dans une même bande de fréquences. La norme IS-95 est à séquence directe, et occupe une largeur de bande de 1,25 MHz par porteuses et par sens de communication [13]: station de base vers terminal et terminal vers station de base ou, dans le vocabulaire américain, *forward direction* et *reverse direction*. Une différence majeure entre

cdmaOne et les autres standards cellulaires vient du fait que dans cdmaOne, la même fréquence est réutilisée dans toutes les cellules du système, ce qui correspond à un facteur de réutilisation de 1, dû au fait que cdmaOne identifie les transmissions des différents mobiles grâce aux différents codes d'étalement qui identifient chaque mobile.

Pour combattre le phénomène d'évanouissement des canaux radiomobiles via la diversité d'antenne (*space diversity*), aussi bien les stations de base (SB) que les stations mobiles (SM), utilisent des antennes qui ont plus d'un élément (RAKE receivers).

L'utilisation du CDMA pour la séparation des utilisateurs nécessite une synchronisation précise entre les SBs pour éviter beaucoup trop d'interférences. Cette synchronisation est obtenue via l'utilisation dans chaque SB d'un récepteur GPS (Global Positioning system). Ces récepteurs GPS fourniront un timing du système, très précis. Une fois les SBs synchronisées, elles ont la responsabilité de fournir les informations de timing aux SMs. Ceci est accompli en acheminant de la SB aux SMs un paramètre identifiant le temps système, décalé du délai d'un aller ou d'un aller retour de la transmission. De cette manière, les SBs et les SMs sont assurées de rester synchronisées.

1.4.3.1 Les codes pseudo bruit Courts, Longs et de Walsh

Les codes d'étalement CDMA utilisés sont de trois types [14]. Les deux premiers types de codes sont des séquences pseudo aléatoires(ou pseudo noise PN) portant des noms particuliers: codes courts et codes longs(short PN and long PN). Les codes de Walsh sont ceux générés en utilisant les matrices de Walsh-Hadamard.

Code court

Le code court utilisé dans cdmaOne est une m-séquence de longueur $(2^{15} - 1)$ bits, soit 32 767 codes (générée par un "Linear Feed Shift Register" de 15 registres). Ces codes sont utilisés pour la synchronisation dans les voies montante(uplink) et descendante(downlink), ainsi que pour l'identification de la cellule/SB dans la voie descendante. Le code court se répète toutes les 26,666 ms. Durant 2 secondes la séquences se répète exactement 75 fois. La petitesse de cette séquence est recherchée, parce que durant l'établissement d'un appel, le mobile recherchera justement un code court et aura besoin de le trouver assez rapidement. Deux secondes est le maximum de temps que le mobile mettra pour trouver une station de base (SB), s'il en existe une.

Chaque SB se voit attribuée l'un de ces codes. Mais, puisque un code court est seulement une séquence, comment fait-on pour l'attribuer à toutes les stations? En le décalant cycliquement. D'après les propriétés des m-séquences, chaque code obtenu par décalage est un code indépendant, puisque la version décalée, d'une m-séquence, a une très petite corrélation. Dans cdmaOne ce décalage est de 512 chips pour chaque station adjacente. Donc, les différentes cellules (ou les différents secteurs d'une cellule) utilisent le même code court, mais avec des décalages différents(PN Offsets), et c'est ainsi que le mobile différencie une station des autres. Le mobile ajoute le code court à ses messages, avec le décalage spécifié, pour que la SB destinataire sache que ces messages lui sont destinés.

Code Long

Les codes longs sont de 2^{42} bits (générés par un LFSR de 42 registres), et défilent à 1,2288 Mb/s. Le temps que ça prend pour recycler cette longueur de code est de 41,2 jours. Ces codes sont utilisés à la fois pour étaler le signal et pour l'encrypter. Une version décalée du long code est générée par le téléphone cellulaire durant l'établissement d'un appel. Le décalage est appelé Long Code Mask, et est unique à chaque appel téléphonique. CdmaOne a un protocole de sécurité appelé CAVE qui requiert une clé d'authentification de 64 bits(A-Key), et un numéro série électronique(ESN) attribué à un mobile sur la base du numéro du téléphone. Le réseau cdmaOne utilise ces deux informations pour créer un nombre aléatoire, nécessaire à la génération d'un masque pour le code long, servant au cryptage et à l'étalement de chaque appel téléphonique. Ce masque n'est pas fixe mais change à chaque fois qu'une connexion est établie.

Il y a un code long public, et un autre privé. Le code long public est utilisé par le mobile pour communiquer avec la SB durant la phase d'attachement. Le code long privé est généré ensuite pour chaque appel, et abandonné à la fin de la communication.

Codes de Walsh

En plus des deux types de codes précédents, un autre type de code spécial, appelé Walsh, est aussi utilisé dans cdmaOne. Les codes Walsh n'ont pas les propriétés des m-séquences en termes de corrélation croisée. Le but principal des codes Walsh dans cdmaOne est réaliser une orthogonalité entre tous les utilisateurs dans une cellule. Chaque canal de trafic utilisateur est associé à un code Walsh différent par la SB. IS-95 a la capacité d'utiliser 64 codes, mais tous ne sont pas dédiés aux trafics utilisateurs. En d'autres termes, chaque station de base a la possibilité de communiquer simultanément sur 64 canaux différents.

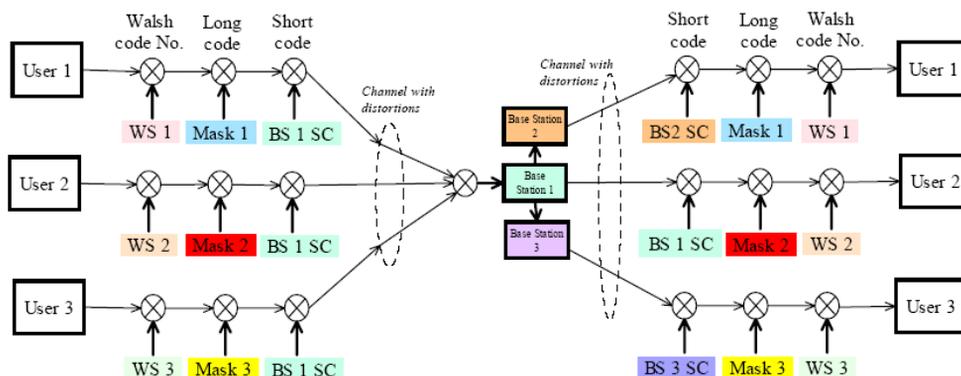


Figure 1.19 interrelations entre les codes utilisés dans le CDMA

Source: Charan Lagton. *Intuitive Guide to Principles of Communications*. www.complextoreal.com/

Ces codes sont utilisés pour l'étalement dans la voie "forward". Dans la voie "reverse", ils ne sont pas utilisés pour différencier les utilisateurs, mais pour la modulation 64-ary. La figure 1.19, représente l'utilisation des différents codes et les interrelations entre eux.

1.4.3.2 Uplink (reverse direction)

La couche physique de l'IS95 assure une double fonction: définir un canal physique et y fiabiliser le transport de l'information. Dans le sens montant un canal physique est matérialisé par l'information portée par un canal correspondant à un code spécifié. Pour chaque canal physique, cdmaOne génère une séquence aléatoire, appelée *long mask*, qui permet de moduler(étalement) l'information utile selon la méthode CDMA, permettant ainsi de multiplexer les utilisateurs sur une même bande passante[14].

Le générateur des séquences est constitué d'un registre de décalage de 42 cases. A débit de séquences de 1,228 8 Mchip/s, la durée qui sépare deux codes identiques est de :

$$(2^{42} - 1) / 1,228\ 8\ \text{Mchip/s} = 41,1\ \text{jours.}$$

La figure 1.20 retrace les étapes de modulation de la voie montante. Les données entrantes sont produites à un débit variable : 1200, 2400, 4800 et 9600 bit/s. Pour maintenir les trames à 20 ms, une répétition des informations est effectuée pour les débits inférieurs à 9600 bit/s. Ainsi, l'information à 1200 bit/s est répétée huit fois, tandis que celles à 2400 et 4800 le sont respectivement à quatre et deux fois. Ces données brutes doivent être protégées lors de leur transmission dans l'air. Cette protection est assurée par un code convolutionnel de taux 1/3 et de longueur 9, qui ajoute deux bits de redondance à chaque bit d'information, multipliant ainsi le débit des données par trois, soit 28,8 Kbit/s(9600 x 3).

Un entrelacement intervient sur chaque trame pour répartir les erreurs sur plusieurs blocs d'information. Cette opération permet de réduire l'effet d'évanouissement du signal lors de sa traversée de l'interface radio.

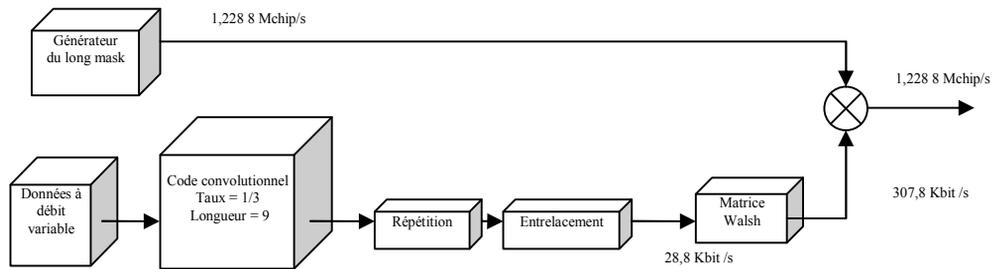


Figure 1.20. Modulation de la voie montante dans l'IS95.

source: Al Agha Pujolle Vivier. Réseaux de Mobiles & Réseaux Sans Fil. Editions Eyrolles.

S'ensuit alors une opération supplémentaire visant à renforcer la protection de l'information contre les erreurs. Elle est équivalente à l'application d'un code correcteur C(64,6). Chaque bloc de 6 bits est remplacé par la fonction de Walsh correspondante de 64 bits. Ceci amplifie le débit à 307,2 Kbit/s.

L'étalement intervient en tant que phase ultime de la modulation binaire. Le flot de données est étalé par le long et court code, respectivement. Un flux de 1,228 8 Mchip/s est envoyé vers le modulateur de phase QPSK (Quadrature Phase Shift Keying) pour une transmission vers la station de base par l'interface air.

1.4.3.3 Downlink (forward direction)

Les données sources à traiter arrivent avec l'un des quatre débits : 1200, 2400, 4800 ou 9600 bit/s. De la sorte, le débit du codeur de la voix change proportionnellement à l'activité vocale en entrée. Quoi qu'il en soit, le flot de données en sortie est passé à un codeur convolusionnel de taux 1/2, qui ajoute deux bits de redondance à chaque bit d'information, multipliant ainsi le débit des données par deux, soit 19,2 Kbit/s (9600 x 2). Le résultat du codeur convolusionnel est l'entrée de la fonction de répétition, qui répète les échantillons des données de débits réduits (1200, 2400 et 4800 bit/s) pour former un débit de sortie constant de 19,2 Kbit/s. Le flot binaire est ensuite entrelacé aléatoirement dans des trames (à un intervalle de 20 ms). Le but de l'entrelacement est, comme déjà signalé, de lutter contre les environnements d'évanouissement multi chemins, qui provoquent des erreurs, par salves, dans le canal radio. Le résultat de cet entrelacement est additionné (modulo 2) à un code brouilleur de débit 19,2 Kbit/s, obtenu d'un décimateur de type 64:1. Le décimateur choisit chaque 64^{ième} bit à partir d'un générateur de "long code" débitant à 1,2288 Mchip/s. Le générateur de "long code" crée de très longs codes ($2^{42} - 1$ bits) basés sur les informations spécifiques de l'utilisateur, telles que l'identifiant du mobile ou le numéro de série électronique de l'utilisateur. Les "long code" procurent un haut niveau de sécurité, à cause de leur longueur. Cette information est fournie au réseau lorsque la SM envoie ces informations de confirmation à la SB. Notons qu'un second décimateur de type 24:1 participe l'insertion de bits assurant le contrôle de puissance d'un débit de 800 bit/s. Notre flot de 19,2 Kbit/s sera étalé dans un premier temps par une

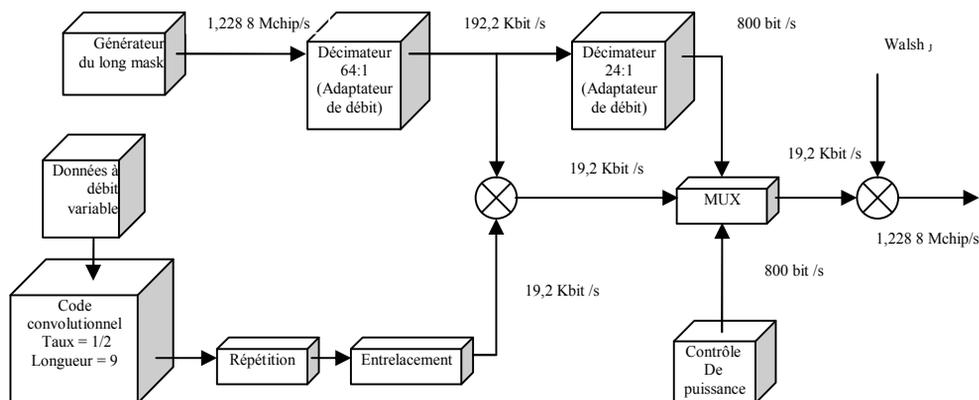


Figure 1.21 Modulation de la voie descendante dans l'IS95.

source: Al Agha Pujolle Vivier. Réseaux de Mobiles & Réseaux Sans Fil. Editions EYROLLES.

fonction Walsh fonctionnant à 1,2288 Mchip/s. L'étalement par matrice de Walsh dote chaque canal d'un identifiant unique. Enfin, le signal étalé (de 1,2288 Mchip/s) est étalé dans un deuxième temps par un "code court" de même débit(1,2288 Mchip/s). Un "code court" est un code d'un pseudo bruit de $2^{15} - 1$ bits de long. Toutes les SBs utilisent le même "code court", mais avec un décalage différent. Il existe 512 décalages différents qui, ainsi, ne permettent d'identifier de manière unique que 512 SBs cdmaOne. Un mobile peut facilement distinguer les transmissions de deux SBs différentes à travers les décalages de leur "code court". Enfin, le signal résultant est transmis sur le medium sans fil via une modulation QPSK. La figure 1.21 retrace les étapes de modulation de la voie descendante.

1.4.3.4 Les canaux logiques [14]

1.4.3.4.1 Canaux de la voie "Forward"

Une station de base peut avoir, comme nous l'avons vue, 64 canaux. Elle a un signal *pilot*, un canal *synch*, et 8 canaux *paging*. Le reste des canaux sont dédiés au trafic avec les mobiles. Ces canaux sont représentés sur la figure 1.22 et sont décrits dans ce qui suit.

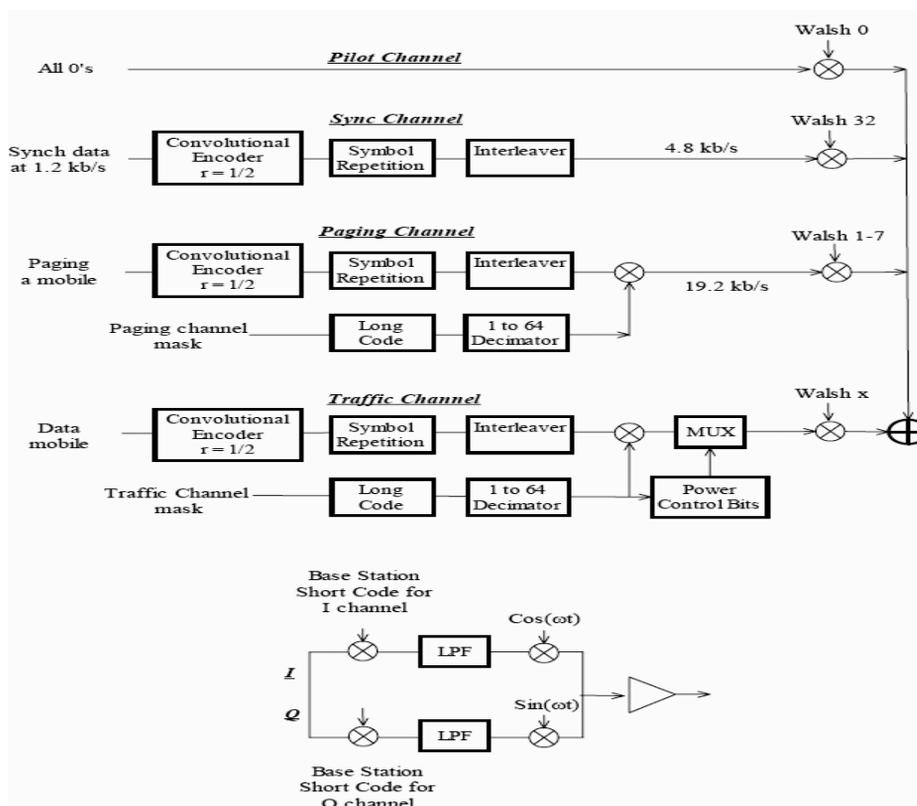


Figure 1.22 Les canaux de la voie descendante ou forward.
 Source Charan Lagton. *Intuitive Guide to Principles of Communications*. www.complextoreal.com/

Canal pilot

Commençons par montrer comment une station établie un contact avec les mobiles dans sa cellule. Elle transmet continuellement un signal représentant une série de zéros, qui est couvert par le code Walsh 0. Ce signal est multiplié par le code court de la station de base, qui est rappelons le, le même code court que toutes les SBs utilisent, mais avec PN offset différent. Les PN offsets du canal pilot sont toujours attribués aux SBs en des multiples de 64 chips, donnant ainsi 512 "Pilot Offset" possibles. Enfin, le signal résultant est converti sur la fréquence porteuse du canal de transmission.

Côté réception, le mobile se saisit du signal et note la SB qui le transmet. Mais, si le code court est répété cycliquement, comme on l'a vu, comment le récepteur connaît-il le PN offset ? Toutes les autres SBs du

voisinage ne paraîtraient t-elles pas les mêmes? Oui, et le mobile à ce point ne connaît pas quelle SB il communique avec, il sait seulement qu'il a trouvé le réseau. Pour déterminer toutes les SBs possibles (il peut y en avoir 256), utilisant chacune un code court décalé de 512 chip, le réseau le signal et le timing du GPS.

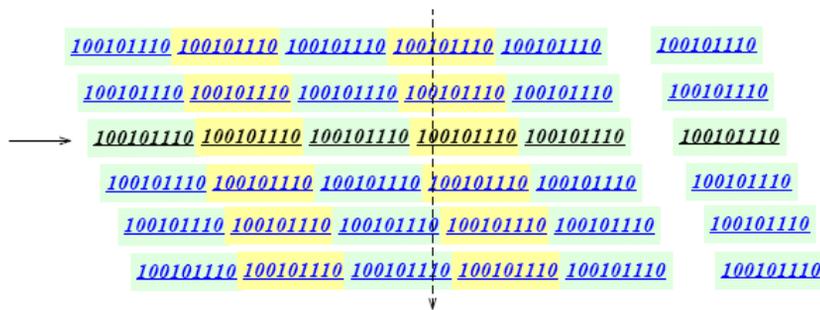


Figure 1.23 Le mobile cherche le code qui s'aligne avec le timing GPS.

Il saisi le code reçu à cet instant, fait la corrélation avec la donnée stockée et connaît quelle station de base il a trouvé.

source: Charan I.agton. *Intuitive Guide to Principles of Communications.*

La SB avec un offset zéro aligne la transmission de son pilot avec chaque top de temps du GPS correspondant à un nombre pair de secondes. Donc, supposons que notre mobile est dans une cellule appartenant à une SB avec un PN offset de 10. Cela signifie qu'il commencera sa transmission 10×512 chip = 5120 chips après chaque top de temps d'un nombre pair de secondes (voire figure 1.23). Lorsque le mobile se réveille et consulte son temps, il sait exactement où chaque code court de SBs devrait être. Tout ce qu'il aura à faire c'est de corréler les bits qu'il reçoit avec chacune des 256 séquences possibles. A la fin de ce processus il saura exactement quelle SB il entend.

Canal Synch

Une fois la sélection de la SB terminée, le mobile doit se synchroniser avec cette SB à travers les informations qu'il trouvera sur le canal Synch(de débit 1200 bps). L'information fournie par ce canal inclue le "pilot offset" du pilot que le mobile à acquis. Cette information permettra au mobile de savoir où chercher les "pilots" dans liste du voisinage. Elle inclu aussi le temps système, le temps du jour, basé sur le temps GPS. Le temps système sert à la synchronisation des fonctions du réseau. Par exemple, les générateurs de code PN sur la voie "reverse" utilise un "offset" zéro par rapport au nombre pair de secondes du temps GPS. Cependant, les mobiles connaissent seulement le temps système de la SB plus une incertitude due au délai de propagation de sa SB à son endroit. L'état du générateur de code long au temps système est aussi envoyé au mobile dans le message de synchronisation. Ceci permet au mobile d'initialiser et démarrer son générateur de code long en synchronisme très rapproché avec les générateurs de code long dans la SB. Le message de synchronisation notifie aussi le débit de données du canal "paging", qui peut être 4800 ou 9600 bits/s.

Canal Paging

Maintenant le mobile est prêt à recevoir et initier des appels. Les données sur le canal "paging" envoyées par la SB, incluent le ESIN(Electronic Serial Identification Number) du mobile, et est couvert par un code long. Comment le mobile pourrait t-il deviner ce qu'est ce code long? Au niveau "paging", le système utilise un code long public, car c'est un canal commun. Lorsque le bon mobile répond, un nouveau code long privé sera attribué à ce moment là, avant que l'appel ne soit établi. En parcourant le canal "paging" le mobile reconnaît son numéro de téléphone et répond par une sonnerie. Lorsque l'utilisateur accepte l'appel un message d'accès retourne à la SB. Il s'ensuit alors un protocole débouchant sur la détermination d'un même code long pour la couverture des messages, aussi bien chez le mobile que chez la SB.

Canal trafic

La station de base peut transmettre un trafic de données à 54 mobiles en même temps. Elle garde ces canaux séparés en utilisant des codes Walsh. Ceci est un multiplexage par division de code plutôt qu'une canalisation basée fréquence. Les codes Walsh sont utilisés par la SB seulement, de cette façon nous avons un CDMA synchrone sur la voie "forward", alors que sur la voie "reverse" c'est un CDMA asynchrone. La construction du trafic sur ce canal se fait comme déjà présenté (section voie "forward"). La SB combine tous ses canaux de trafic (chacun couvert par un code Walsh différent), tous les canaux "paging" (juste 8), le canal "pilot", et le canal "synch". Elle les additionne, et fait la conversion série parallèle vers les canaux I et Q. Chacun est couvert par un code court I et Q, modulé par QPSK sur une fréquence porteuse, et envoyés dans la cellule.

1.4.3.4.2 Canaux de la voie "reverse"

Dans cdmaOne, un mobile transmet seulement sur deux canaux, et jamais simultanément. C'est soit sur un canal "Access" soit sur un canal "Traffic".

Canal Access

Pour transmettre de l'information vers la SB, l'utilisateur se sert du canal Access. Ce canal lui permet tout à la fois, d'enregistrer sa localisation, de répondre à un message de *paging*, et d'envoyer un segment de données.

Canal Traffic

Le canal de trafic sert à envoyer de l'information ou de la signalisation dédiée. La transmission sur ce canal étant déjà présentée, nous n'y reviendrons pas.

1.4.3.5 Control de puissance

CdmaOne est un système à accès multiple où plusieurs utilisateurs ont accès à la même bande de fréquences. Le succès du système réside dans un control total de la puissance des signaux émis. Dans un environnement cdmaOne, chaque station mobile (SM) est une source de bruit pour les autres. Une SM voit l'environnement autour d'elle comme une addition cumulative d'information qui lui est destinée, et d'interférences. Les interférences, à leur tour, sont constituées d'informations pour d'autres SMs, plus du bruit provenant de sources diverses. Il est clair alors que, si les interférences sont trop importantes, l'information ne pourra pas être retrouvée [16].

Généralement, la station de base (SB) est bombardée par des signaux provenant de plusieurs SMs. Certaines de ces SMs sont plus proches de la SB que d'autres, et leurs signaux par conséquent, plus forts que ceux des mobiles lointains (l'évanouissement et les chemins multiples atténuent les niveaux des puissances). C'est le problème "Near-Far", inhérent aux communications CDMA. Ces systèmes sont "Interference limited", c.a.d, que leur capacité est étroitement dépendante de la puissance des signaux qui règne dans le réseau.

Pour prendre en charge tous ces aspects, les systèmes CDMA implémentent une fonction de contrôle de puissance dédiée, qui gèrent les niveaux de puissance de tous les mobiles pour, les maintenir en dessous d'un certain seuil requis, et les rendre très proches, que le mobile soit proche ou loin. A cet effet, le système maintient ce qu'on appelle une boucle de control de puissance.

Selon le standard IS-95, le control de puissance CDMA est un processus à trois étapes [19]:

1. Reverse Link Open-Loop Power Control (lien mobile à base): ce control est accompli en ajustant la puissance de transmission du mobile de sorte que le signal reçu à la SB soit constant indépendamment de la distance du mobile à la SB.
2. Reverse Link Closed-Loop Power Control: ce control est accompli au moyen de deux commandes ("power-up" et "power-down") provenant du site de la cellule. Un seul bit de puissance (1 pour 0,5 dB de puissance en plus et 0 pour 0,5 db de puissance en moins) est inséré dans le flot de données de la voie "forward", toutes les 1,25 ms. En recevant cette commande de la SB, le mobile répond en ajustant sa puissance de transmission comme instruit.

- Forward Link Power Control (base to mobile): c'est un processus à une étape. La SB contrôle de puissance de transmission de sorte qu'un mobile donné reçoive un plus de puissance pour venir à bout du phénomène d'évanouissement, des interférences, erreurs, ...etc. Dans ce mécanisme, le site de la cellule réduit sa puissance de transmission pendant que le mobile calcul le "frame error rate"(FER). Une fois que le mobile détecte un FER de 1%, il envoi une requête pour stopper la réduction de la puissance.

1.4.4 CMDA2000 [16]

1.4.4.1 Présentation

cdma2000 incarne le prolongement des réseaux de deuxième génération fondés sur le multiplexage par code. Il descend directement de cdmaOne (IS-95), avec lequel il est entièrement compatible. Bien que cdma2000 puisse être utilisé comme une interface air de réseaux 3G purs (utilisation du spectre IMT-2000), son principal avantage est son aptitude à recouvrir (overlying) aussi, sur une même bande de fréquences, des systèmes (2G) IS-95. Cet aspect est très important pour les opérateurs d'Amérique du nord du fait que la bande de fréquence spécifiée par l'UIT pour IMT-2000 est déjà en utilisation dans ces régions. Il peut servir des clients dans des cellules de différentes tailles, des pico et micro cellules aux macro cellules, voire des méga cellules.

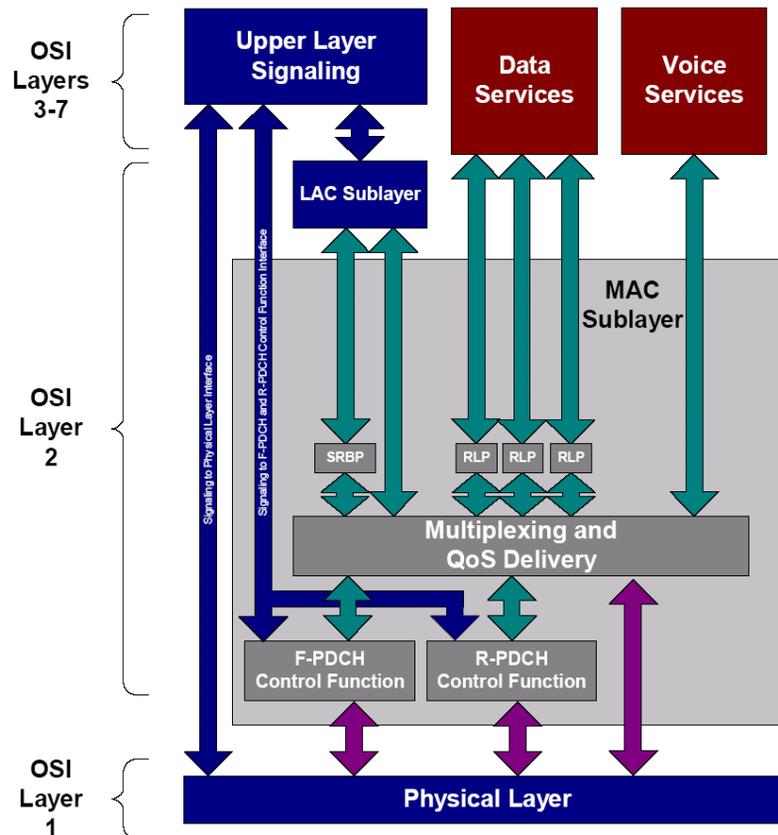


Figure 1.24 Architecture du cdma2000

source: Introduction to cdma2000 Standards for Spread 3GPP2 V.1.0

Cdma2000, dans sa forme d'origine, a surtout augmenté l'efficacité du système IS-95 par de nouveaux apports, incluant (i) des apports en termes de diversité en transmission et d'antennes intelligentes, et (ii) un pilot séparé pour chaque uplink, permettant ainsi une réception cohérente.

Cdma2000 utilise plusieurs fonctionnalités qui n'étaient pas présentes dans cmdaOne(IS-95) :

- Les canaux SCH (Supplemental CHannel) sont améliorés, pour qu'ils supportent non seulement les informations de control (avec des débits de données inférieurs à 9.6 et 14.4 Kbit/s), mais aussi les transmissions de données à hautes vitesses (jusqu'à 1 Mbit/s). Plus précisément, les canaux SCH opèrent dans deux modes. En mode 1x, ils sont complètement analogues à ceux de IS-95. En mode 3x, les débits de données sont supérieurs à 14.4 Kbit/s, et sont négociés entre la SB et la SM.
- Des canaux dédiés et des canaux communs sont utilisés pour les services de données. Le canal de control commun transporte les courts messages et la signalisation MAC pour les données paquet. Les canaux de control dédiés peuvent, soit transporter les messages courts de l'utilisateur, soit l'information sur le control de puissance et la continuité du lien.
- Les codes Walsh ont maintenant une longueur variable. Ces codes de longueur variable peuvent être utilisés pour la transmission de données sur les canaux SCH.
- Control de puissance rapide dans la voie descendante (downlink).
- Un canal pilot pour chaque voie montante (uplink), permettant une réception cohérente.
- Support d'antennes intelligentes, par l'introduction de canaux pilot auxiliaires dédiés et commun.

Cdma2000 a aussi inclus deux modes de fonctionnements possibles : le mode appelé 1x, qui utilise une

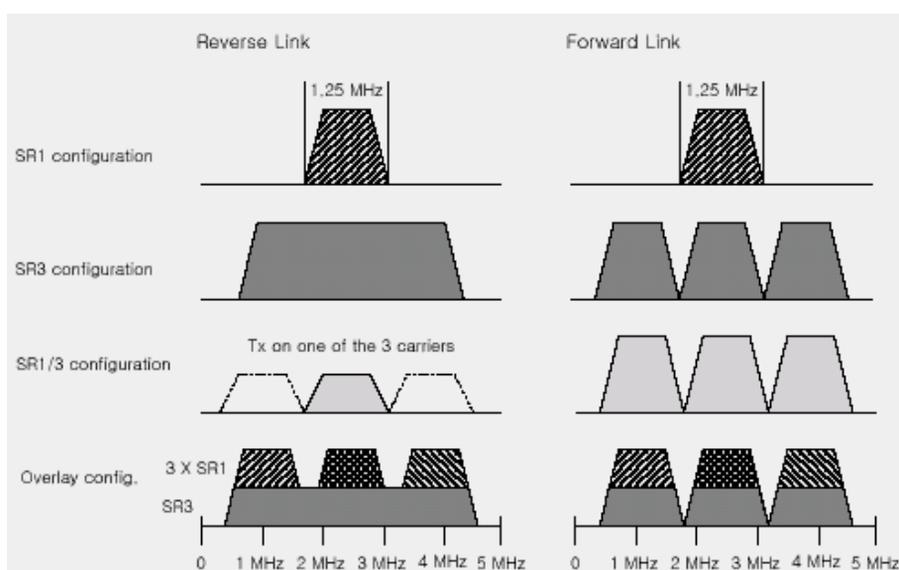


Figure 1.24 Bandes passantes de fonctionnement du cdma2000.

source: Serge Willenegger. cdma2000 Physical Layer: An Overview. QUALCOMM International, Switzerland 2000.

largeur de bande de 1,25 MHz, et le mode appelé cdma multi-porteuses (ou mode 3x), qui utilise une largeur de bande de 5 MHz en combinant trois porteuses IS-95, en une seule bande (figure 1.24).

Cdma2000 - mode 3x

Le mode à 5 MHz du cdma2000, est aussi connu comme le "mode 3x" (car il utilise trois porteuses), ou bien comme mode cdma multi porteuses. L'idée de base derrière le 3x est d'utiliser une bande passante plus large, autorisant, ainsi, une vitesse de données plus grande. Cette bande passante plus large, est réalisée de différentes manières pour la voie uplink et la voie downlink. Dans le "uplink", la SM adopte une vitesse de chip plus large, c.a.d $3 \times 1.2288 = 3\ 6864$ Mchip/s. Dans le "downlink", la SB transmet en parallèle sur trois porteuses adjacentes. Les données sont d'abord codées pareil à IS-95, et puis démultiplexées sur les trois porteuses. Ainsi, la diversité en fréquences à une plus large bande est encore disponible.

Cdma2000 utilise un duplexage par division de fréquences (Frequency Division Duplex), mais un duplexage par division de temps (Time Division Duplex) est aussi supporté. Dans le mode FDD, le trafic

uplink et downlink sont transportés sur des canaux fréquentiels différents. Le mode TDD est avantageux dans les environnements où seulement des bandes de fréquences amoindries sont disponibles.

Cdma2000 - mode 1x

Un des aspects clé du mode 1x est sa bonne compatibilité avec le système IS-95 d'origine. Quant elle opère dans le mode 1x, une SB cdma2000 supporte la communication avec des SMs IS-95 aussi bien que des SMs cdma2000. En plus des canaux existants dans IS-95, un certain nombre de nouveaux canaux, supportant des fonctionnalités additionnelles comme les données en paquet et les antennes intelligentes, sont aussi apparus.

1.4.4.2 Les canaux logiques

Canaux dédiés sur la voie "downlink " Cdma2000 a aussi bien, des canaux compatibles avec IS-95, que des canaux de nouveaux types. Le canal FCH (*Fundamental CHannel*) est similaire au canal Trafic, et autorise aussi quatre débits de données différents; selon le "rate set", le débit maximum est de 9.6 ou 14.4 Kbit/s. Le canal de control dédié, est soit transmit à un débit maximum, soit pas du tout. Il peut y avoir aussi un ou deux canaux F-SCHs (*Forward Supplemental CHannels*), fournissant de hauts débits de données. La raison à avoir des canaux additionnels, et non pas juste un canal avec un facteur d'étalement (spreading) variable, est qu'il est plus facile pour un "émetteur-récepteur" d'allumer et de fermer les canaux SCHs, que de constamment changer le débit du code d'un seul canal disponible. Noter, qu'un FCH existe toujours, pour qu'un SCH puisse seulement se produire, conjointement à un FCH. Dans la voie downlink, l'information de signalisation peut être transmise soit en utilisant le "blank-and-burst", où elle remplace quelques informations de la parole, soit par transmission sur un canal séparé, le F-DCCH (*Forward Dedicated Control CHannel*). La dernière option est spécialement utile quand une transmission VoIP est utilisée. Si des antennes directives sont utilisées, alors le F-APICH (*Forward dedicated Auxiliary Pilot CHannel*) permet d'obtenir l'estimation du canal pour ce lien spécifique. Enfin, le F-PDCH (*Forward Packet Data CHannel*) transporte les paquets de données, et est attribué à un utilisateur pour de très courts intervalles de temps (1.25ms, 2.5ms, ou 5ms).

Canaux communs sur la voie "downlink" Le premier group de canaux communs est celui des canaux pilot. En plus du canal pilot commun, connu, du IS-95, il peut y avoir aussi des canaux pilot additionnels dédiés à un lien spécifique (voir plus haut), de même que le F-TDPICH(*Forward Transmit Diversity Pilot CHannel*) et le ATDPICH(*Auxiliary forward Transmit Diversity Pilot CHannel*) utilisés pour la diversité en transmission. Juste comme dans IS-95, le F-SYNC(*Forward SYNChronization channel*) sert à la SM pour acquérir une information système initiale. Un changement important par rapport à IS-95 est observé dans le "paging". Alors que cdma2000 à encore un canal paging pour des raisons de compatibilité de retour arrière, les fonctionnalités sont habituellement assurées par tout un ensemble de nouveaux canaux, chacun d'eux étant optimisés pour une tâche spécifique. Le F-QPCH(*Forward QuickPaging CHannel*) contient un très court message qui dit à la SM si un message de paging est en attente ou non. Si aucun message n'est en attente, alors la SM peut immédiatement retourner en mode endormi, ce qui préserve l'énergie. C'est seulement dans les (statistiquement rares) cas où un message de paging est en attente, que la SM reste allumée, et écoute le message suivant sur le F-CCCH(*Forward Common Control CHannel*). La configuration système et d'autres paramètres sont transmis sur le F-BCCH(*Forward Broadcast Control CHannel*).

Un autre groupe de canaux de control comprend le F-PDCCH(*Forward packet Data Control CHannel*), qui contient la signalisation et l'information de control pour le canal de donnée paquet. Enfin, signalons aussi le F-CACH(*Forward Common Assignment CHannel*) et le F-CPCCH(*Forward Common Power Control CHannel*).

Canaux dédiés sur la voie "uplink" Les canaux dédiés sur la voie uplink sont du même type que ceux de la voie downlink, à savoir, le R-PICH (*Reverse Pilot CHannel*), R-FCH (*Reverse Fundamental CHannel*), le R-DCCH (*Reverse Dedicated Control CHannel*), et le R-SCH (*Reverse Supplemental CHannel*). En

plus, il y a le R-CQICH (*Reverse Channel Quality Indicator CHannel*) et le R-ACKCH (*Reverse ACKnowledgement CHannel*).

Canaux communs sur la voie "uplink" Les canaux communs de la voie uplink sont similaires à ceux du IS-95, et incluent le R-ACH (*Reverse Access CHannel*), qui existe aussi dans IS-95, ainsi que deux nouveaux canaux: R-EACH (*Reverse Enhanced Access CHannel*), and the R-CCCH (*Reverse Common ControlCHannel*). La raison derrière l'introduction de nouveaux canaux Access se trouve dans le fait que pour cdma2000 les messages d'accès aléatoire sont plus fréquents et plus longs. Ceci mène à une grande probabilités de perte de messages, une latence élevée, et une utilisation des ressources plus importante. Cdma2000 définit donc deux modes d'accès aléatoire: un *mode* dit *basic*, pour des messages courts, similaire à IS-95, où les messages sont transmis avec des puissances de plus en plus fortes, jusqu'à ce que la SB envoie un accusé de réception du message; et pour les messages plus longs, un *mode de réservation d'accès*. Dans ce dernier mode, les messages d'accès aléatoire sont scindés en deux parties. La SM, dans la première partie, envoie une très courte requête pour obtenir de la capacité sur le "reverse link". La SB accuse réception de cette requête sur le canal F-CACH. Une fois que la SM reçoit cet accusé de réception, elle pourra transmettre la partie donnée du message sur le canal R-CCCH. Ce canal a un control de puissance, et est ainsi plus fiable et plus efficace en terme de puissance que le R-ACH.

1.4.4.3 Quelques caractéristiques du cdma2000

Cdma2000 présente un nombre d'améliorations dans la modulation, le codage, et le control de puissance ce qui, comparé à IS-95, rehausse la flexibilité et augmente le débit. Dans ce qui suit nous discutons des plus importantes de ces améliorations.

Control de puissance rapide sur la voie "downlink" Dans IS-95, seule la voie uplink avait un control de puissance rapide, où la puissance de transmission de la SM pouvait changer chaque 1.25ms. Dans cdma2000, la voie downlink aussi a cette capacité. Donc, le control de puissance de la voie downlink est maintenant capable de suivre l'évanouissement rapide; ceci résulte en des améliorations de performance considérables.

Meilleur codes correcteurs d'erreur Cdma2000 utilise différents codes convolutifs, et inclue aussi les turbo codes en option. Les codes convolutifs ont des taux de 1/2 et 1/4.

Diversité en transmission (Transmit diversity): Une des améliorations de la conception de la voie descendante du système cdma2000 est l'utilisation de la diversité en transmission côté station de base. Bien sûr, cette approche ne sera effective que si la diversité est aussi employée côté récepteur. La SB utilise des antennes séparées, pour transmettre les multiples sous-porteuses. Les signaux sortant des différentes antennes vont subir le phénomène d'évanouissement de manière indépendante, augmentant ainsi la diversité de fréquence.

Modulation pour le downlink Cdma2000 utilise des codes OVVSF, c.a.d, des codes Walsh à taux d'étalement variable, où les taux peuvent varier entre 2 et 512. Ceci donne une flexibilité accrue dans le nombre de débits de données supportés. En plus, cdma2000 définit aussi un ensemble de QOFs (Quasi Orthogonal Functions) qui interviennent lorsque tout les codes OVVSF ont été utilisés, c.a.d, tout les codes avec un facteur d'étalement 256, ou ceux dont le code père, dans l'arbre des codes, est utilisé. Cette situation se produit quand les antennes intelligentes entre en jeu. La modulation elle-même a changé, permettant au facteur d'étalement des codes Walsh de doubler, et ainsi mettre plus d'utilisateurs dans une cellule.

Modulation pour le uplink La modulation sur la voie uplink présente des changements considérables par rapport à la voie uplink IS-95. Différents canaux logiques transmis à partir d'une SM sont séparés par des codes Walsh différents. Chaque SM transmet un canal (de trafic) fondamental, un canal pilot, et peu être d'autre canaux de control. L'existence du canal pilot dédié est une différence majeure par rapport à IS-95; elle a l'avantage de permettre une démodulation cohérente facile, même si en contre partie la consommation d'énergie est augmentée. La modulation utilisée est le BPSK.

HDR (High Data Rate) est un service améliorant celui de 1X pour les données. Au lieu d'une modulation QPSK utilisée en mode 1X et 3X, HDR utilise une modulation plus efficace, 16-QAM, qui code 4 bits par symbole transmit, offrant ainsi un débit de 621 Kbit/s. Cependant, en cas de fortes interférences, la modulation HDR régresse vers une modulation 8-PSK ou QPSK, ce qui diminue les débits offerts.

Durée des trames En plus de la trame 20ms de IS-95, cdma2000 définit aussi des trames de 5ms, 40ms et 80ms, ce qui donne une plus grande flexibilité. Les trames longues sont principalement réservées pour les services qui ne sont pas temps réel. Pour des transmission des données avec des débits entre 81.6 et 3091 Kbit/s, des tailles de trames de 1.25, 2.5, et 5s sont possibles.

1.4.5 WCDMA (Wideband CDMA)

WCDMA est un standard d'interface air 3G, basé sur la technologie CDMA. Contrairement à l'obligation pour les SBs dans cdma2000 d'opérer de manière synchrone, WCDMA utilise un schéma asynchrone. Ceci permet une installation/intégration plus facile des composants "indoor" avec les infrastructures "outdoor". Dans la spécification WCDMA, le terme "wideband" dénote l'utilisation d'une large porteuse. En effet, WCDMA fonctionne avec une porteuse de 5 MHz; c.a.d, quatre fois celle de cdmaOne et 25 fois celle du GSM. Les systèmes basés sur le WCDMA peuvent coexister avec des systèmes de génération plus ancienne si le spectre correspondant peut être dégagé. En effet, de larges porteuses requièrent plus de spectre disponible, ce qui pose des difficultés dans les cas de pénuries de spectre, comme cela se produit avec les opérateurs Nord Américains [20].

Les débits supportés sont de 144 Kbit/s en environnement rural extérieur, 383 Kbit/s en environnement urbain extérieur, et 2 Mbit/s en indoor (mobilité réduite). Quatre classes de services sont définies dans le cadre de l'UMTS, et peuvent être réparties en deux groupes :

- Les classes **A** (conversational) et **B** (streaming) pour les applications subissant des contraintes temps réel.
- Les classes **C** (interactive) et **D** (background) pour les applications avec des données sensibles aux erreurs de transmission.

Les principales contraintes retenues pour la définition des classes de services de l'UMTS sont:

- Le délai de transfert de l'information,
- La variation du délai de transfert des informations,
- La tolérance aux erreurs de transmission.

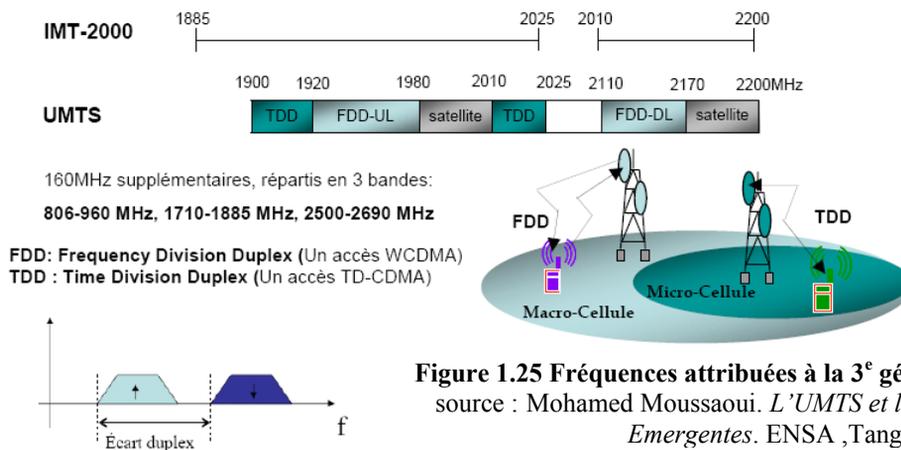
Dans les sections qui suivent, nous allons parcourir quelques aspects ayant trait aux couches physiques et de liaison, où nous décrirons brièvement les principaux canaux de chacune d'elles.

1.4.5.1 Aspects de la couche physique du WCDMA

La couche physique WCDMA implémente une interface air basée sur un étalement direct CDMA sur une bande passante de 5 MHz. Le débit chip fixé, de 3.84 Mchip/s, est très proche de celui du cdma2000 (3.84 Mchip/s). Le standard se décline en deux modes de duplexage principaux: le mode FDD et le mode TDD. Le mode FDD (*Frequency Division Duplex*) nécessite l'allocation de deux bandes de fréquences, une pour la voie montante(uplink) et l'autre pour la voie descendante(downlink), permettant ainsi d'émettre et de recevoir en même temps. FDD n'est pas très efficace dans l'allocation de la bande passante pour tous les types de services. L'accès à un service internet, par exemple, requière plus de débits sur le "downlink" que sur le "uplink". Bien sûr, par un ajustement du facteur d'étalement, on peut, dans le FDD, utiliser juste le débit de données requis, cependant, échanger une capacité uplink contre une autre downlink, est impossible.

Le mode TDD (*Time Division Duplex*), d'un autre côté, utilise la même bande de fréquences pour le uplink et le downlink en allouant des "timeslots" à chaque direction de voie. Par conséquent, TDD peut allouer efficacement la capacité entre la voie montante et la voie descendante, offrant ainsi un support aux demandes de trafic asymétrique. En contre partie, une meilleure synchronisation, dans le temps, que celle du FDD est requise pour assurer que les transmissions d'une SM et d'une SB ne se recouvrent jamais dans

le temps. La figure 1.25 représente les fréquences attribuées dans le cadre des réseaux sans fil de troisième génération, ainsi que les modes de duplexage utilisés.



Un certain nombre de canaux physiques sont supportés, par WCDMA, pour les voies uplink et downlink. Ces canaux physiques sont un moyen de transmission des données transportées sur les canaux logiques. Des trames de 10ms et deux méthodes de duplexage (FDD et TDD) sont définis. La structure de base des trames est la même dans les deux modes (FDD et TDD), cependant, les trames TDD contiennent des "switching points" pour la séparation du trafic uplink/downlink. Le taux de slots uplink/downlink dans une trame peut varier pour s'adapter à un trafic asymétrique, avec des taux variant de 15/1 jusqu'à 1/7. Des structures possibles de trames TDD WCDMA sont illustrées sur les figure 1.26 1- 4.

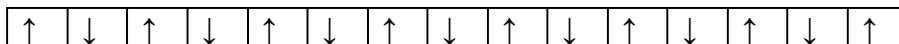


Figure 1.26-1 Configuration à points de commutation multiples (allocation DL/UL symétrique).

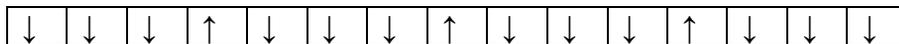


Figure 1.26-2 Configuration à points de commutation multiples (allocation DL/UL asymétrique).

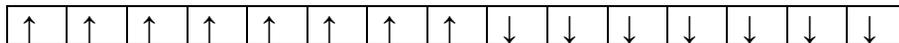


Figure 1.26-3 Configuration à un seul point de commutation (allocation DL/UL symétrique).

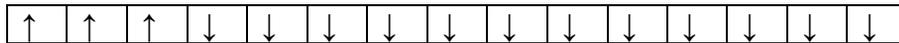


Figure 1.26-4 Configuration à un seul point de commutation (allocation DL/UL asymétrique).

La nature asynchrone du fonctionnement des SBs doit être prise en considération lors de la conception des algorithmes de soft handover pour le WCDMA. Dans un effort pour supporter des capacités accrues à travers le HCS(Hierarchical Cell Structure), WCDMA emploie aussi une méthode de handover, appelée handover inter fréquences. Dans le HCS plusieurs fréquences porteuses différentes sont utilisées simultanément dans une même cellule afin de servir des demandes plus nombreuses dans les "hot spots". Pour exécuter un handover dans les situations HCS, la SM a besoin de l'aptitude à mesurer la force du signal d'une fréquence porteuse alternative pendant qu'elle est toujours connectée sur la fréquence courante. Deux méthodes pour les mesures inter fréquences existent pour le WCDMA. La première, appelée "dual receiver mode", est utilisée quand la diversité d'antennes est employée. Elle se sert de différentes branches d'antennes pour estimer des fréquences porteuses différentes. La seconde, appelée "slotted mode", compresse les données transmises (en utilisant, peut être, un ratio d'étalement plus bas durant une période plus courte) pour épargner du temps pour les mesures sur des fréquences porteuses alternatives [20].

La couche physique fournit deux types d'accès en mode paquet à l'aide de l'accès aléatoire (random access) et des canaux (utilisateur) dédiés. Le "random access" est basé sur l'approche "slotted ALOH", sur la voie "uplink" seulement, et pour les courtes et peu fréquentes salves (bursts). Cette méthode d'accès est plus efficace en termes d"overhead", puisque le canal n'est pas maintenu entre les salves d'information. L'accès dédié sert les salves d'informations plus fréquentes sur les deux voies (uplink et downlink). Plus

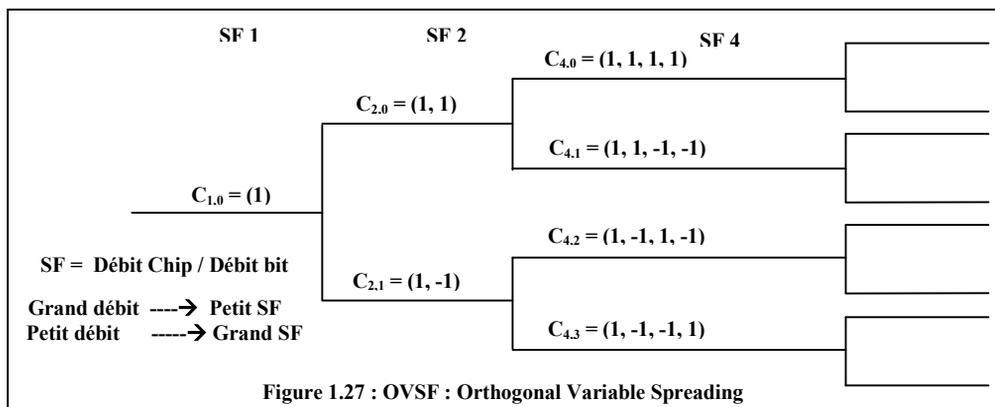
encore, la couche physique WCDMA assure des possibilités de diffusion et de paging pour les couches supérieures.

Dans les sections suivantes nous présentons quelques caractéristiques de la *TAR* implémentée dans l'UMTS, l'un des standards représentatifs des réseaux sans fil WCDMA.

1.4.5.2 Caractéristiques de la couche physique de l'UMTS

Sommairement décrite, la couche physique de l'UMTS présente les caractéristiques suivantes:

- *Une large bande.* Avec des canaux de 5 Mhz la capacité est accrue, et l'UMTS a le double de la capacité des systèmes CDMA de bande étroite dans les environnements urbains des zones périphériques.
- *Étalement.* Deux types de codes sont utilisés dans l'UTRAN : codes de canal (canalization codes), et codes de brouillage (scrambling codes).



Les codes OVSFs de type Walsh-Hadamard (orthogonal variable spreading factor), sont utilisés pour séparer les applications issues d'une même source (Figure 1.27). Par exemple, sur la voie uplink, les codes OVSFs servent à séparer le canal de données et de contrôle pour un utilisateur, tandis que sur la voie downlink les codes OVSFs servent à séparer les connexions des différents utilisateurs dans une cellule:

- 2 séquences situées sur le même niveau hiérarchique sont orthogonales.
 - 2 codes situés sur la même branche sont non-orthogonaux.
- ⇒ limite importante dans l'allocation des codes
 ⇒ ex : si $C_{2,1}$ utilisé, on ne peut plus utiliser $C_{1,1}$, $C_{4,1}$, $C_{4,2}$ et tous les descendants
 ⇒ débit utilisateur élevé ↔ SF faible ↔ on bloque une partie importante de l'arbre

Ces codes prennent leurs valeurs dans la plage 4 à 256 dans le uplink FDD, et de 4 à 512 dans le downlink (et de 1 à 16 dans les deux voies du TDD). Selon le facteur d'étalement (SF), il est possible de réaliser différents débits de données.

Les codes de brouillage. Les propriétés d'inter corrélation entre les séquences OVSF ne sont garanties que lorsqu'elles sont synchronisées, d'où la nécessité d'un 2ème traitement par des codes dits "de brouillage" alloués selon les principes suivants (La figure 1.28) :

Sens descendant :

- un code de canalisation par canal physique,
- la cellule dispose de la totalité de l'arbre OVSF,
- un code de brouillage par cellule,
 ⇒ les codes de brouillage distinguent les cellules entre elles.
- 512 codes disponibles

⇒ contrainte sur la planification du réseau (cellules proches donc codes différents).

Sens montant :

- un code de (channelization) par canal physique,
- le mobile dispose de la totalité de l'arbre OVSF,
- un code de brouillage par mobile
⇒ les codes de brouillage distinguent les mobiles entre eux
- 2^{24} codes disponibles,
⇒ pas de contrainte sur la planification entre les mobiles (aléatoire).

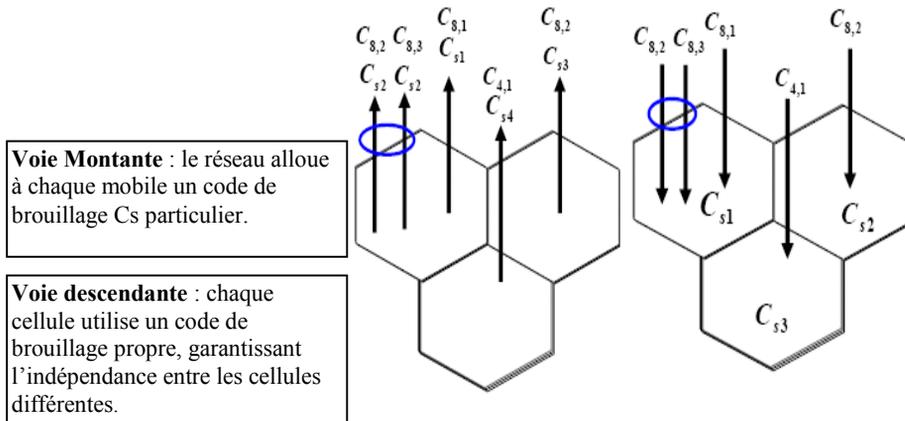


Figure 1.28 Utilisation des codes dans UTRAN.

source: M. Moussaoui. l'UMTS et les technologies émergentes. ENSA, Tanger, Maroc 2009.

- Support d'antennes adaptatives. Le support de rangés d'antennes adaptatives augmente l'efficacité et la capacité du spectre, en optimisant les performances des antennes pour chaque SM.
- Codage canal et entrelacement. Selon le BER et les délais requis, différents schémas de codage pourraient être appliqués. Codage convolutif, turbo codage, ou pas de codage du tout.
- Démodulation cohérente downlink/uplink et control de puissance rapide.
- Support pour la diversité en transmission sur le downlink et techniques de détection multi utilisateurs.

1.4.5.3 Protocoles UMTS

La figure 1.29, donne une vue simplifiée de l'architecture UMTS. Elle divise l'UMTS en trois parties significatives: le mobile (Equipement Utilisateur), le UTRAN (réseau d'accès), et le CN (réseau cœur). L'information qui transite par le UTRAN est logiquement divisée en deux parties :

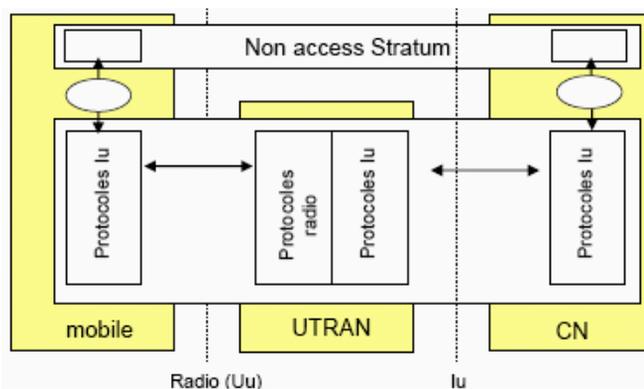


Figure 1.29 Le découpage en strates dans l'architecture protocolaire de l'UMTS.

source: M. Moussaoui. l'UMTS et les Technologies Emergentes. ENSA, Tanger. Maroc 2009.

- le "Access Stratum" (AS), renvoie aux informations nécessaires pour l'interaction entre le mobile et le UTRAN. Elle regroupe toutes les fonctions liées au réseau d'accès et les fonctions de gestion des ressources radio.
- le "Non Access Stratum" (NAS), renvoie aux informations transférées entre le CN et le mobile à travers le UTRAN. Elle regroupe toutes les autres fonctions du réseau UMTS comme: les fonctions d'établissement d'appel: **CC (call control)** pour les appels circuit, et le **SM (session mangement)** pour les appels paquet **Les fonctions de gestion de la mobilité en mode veille.**

La raison de cette division est le désir de rendre le transfert d'information entre le CN et le EU indépendant des technologies d'accès radio (figure1.30).

Iu: l'interface entre le CN et AN(Access Network) capable de connecter de technologies différentes au réseau coeur comme:
 ---- Le **BRAN**: Bradband Radio Access Network),
 réseau d'accès large bande WLAN
 ---- Le **SRAN**: Satellite Radio Access Network
 ---- L'**UTRAN**: le réseau d'accès de l'UMTS

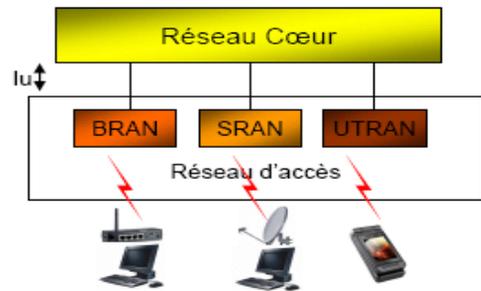


Figure 1.30 Indépendance de la couche d'accès.
 source: M.Moussaoui. l'UMTS et les technologies Emergentes. ENSA, Tanger. Maroc 2009.

La seule vision qu'a le NAS du canal de communication utilisé est le RAB (Radio Access Bearer). Dans l'AS, le RAB est décomposé en deux parties(figure 1.31):

- Le radio bearer, correspondant au segment " interface radio " du RAB;
- L'Iu bearer, correspondant au segment "interface Iu " du RAB.

Le NAS ne connaît pas les caractéristiques du RAB. Le RAB est caractérisé par des attributs de qualité de service, négociés entre l'utilisateur et le réseau cœur.

En fonction de la valeur des ces différents attributs, l'UTRAN doit être en mesure d'effectuer les opérations suivantes:

- le choix d'un codage canal,
- le dimensionnement des ressources radio associée au RAB,
- l'allocation du radio bearer et l'Iu bearer,
- la configuration des protocoles radio, en fonction des caractéristiques des SDU qui seront échangés sur le RAB.

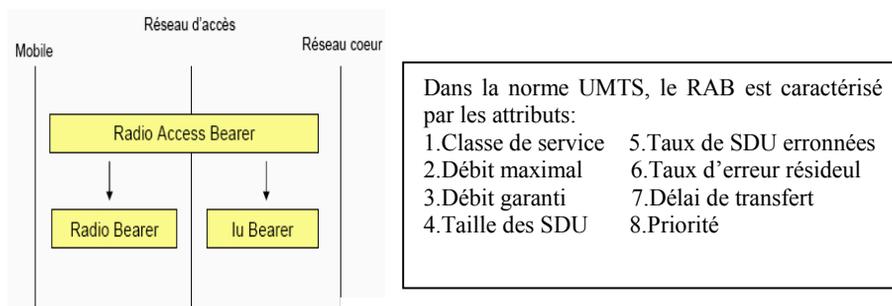


Figure 1.31 Le RAB et ses attributs.
 source:M.Moussaoui. l'UMTS et les Technologies Emergentes. ENSA, Tanger. Maroc 2009.

1.4.5.4 Architecture radio

La figure 1.32 montre la réalisation des parties basses de la pile de protocoles dans l'EU(Equipement Utilisateur). Comme indiqué, une couche physique(L1), une couche liaison(L2), et une couche réseau (L3 :la partie pour le RRC) ont été spécifiées[23].

La couche L1, est responsable de la transmission des données sur l'interface air. Les solutions FDD et TDD ont été présentées. La couche L2 contient quatre sous couches:

- **La sous-couche MAC (Medium Access Control).** La couche MAC est située au dessus de la couche physique avec la quelle elle échange des informations à travers des canaux de transport. Ces canaux décrivent comment les données sont transmises à travers l'interface air et avec quelles caractéristiques.

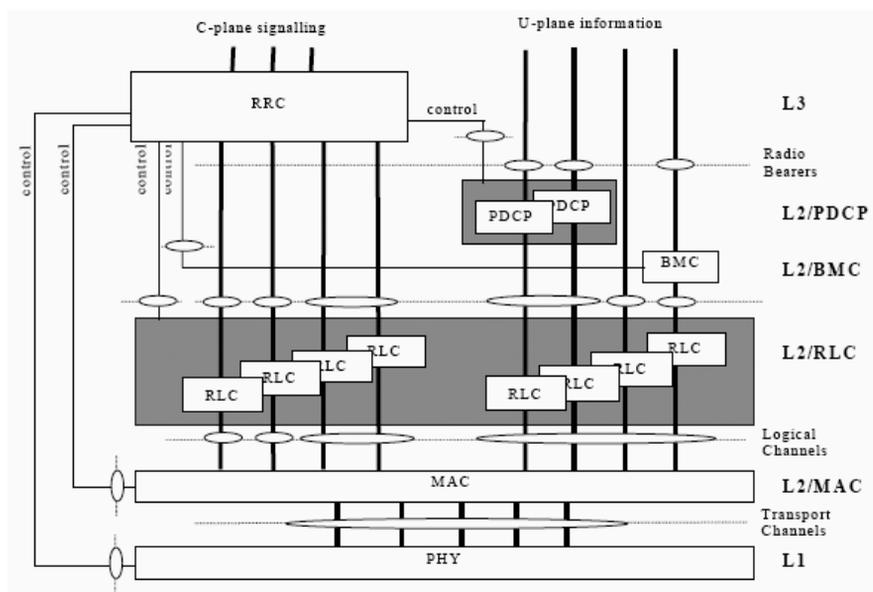


Figure 1.32 Architecture des protocoles de l'interface radio.
source : Matti Turunen. *UMTS Protocol Functionalities*. Nokia Mobile Phones, 1999.

Un ensemble de canaux logiques sont utilisés pour la communication avec les couches supérieures. La responsabilité de la couche MAC dépasse le seul fait de faire correspondre des canaux logiques à des canaux physiques. Elle intervient aussi pour traiter les priorités des EUs et des flots de données d'un EU particulier, surveillance du trafic, cryptage, multiplexage des données sur les canaux de transport, et plus encore.

- **La couche RLC (Radio Link Control).** Cette couche assure un transport fiable et transparent des données entre deux équipements(m.e.p de la Qos, des connexions RLC, notification des erreurs, etc). Il y a une connexion par radio bearer.

Les deux sous-couches de protocole(L2) restantes sont utilisées dans le plan utilisateur seulement:

- **PDCP (Packet Data Convergence Protocol).** Elle est responsable de l'émission et de la réception des unités de données du protocole (PDUs) radio de la couche réseau. Elle assure deux fonctions importantes:

- l'indépendance des protocoles radio de l'UTRAN par rapport aux couches de transport réseau,
- le support d'algorithmes de compression de données ou d'en-têtes de paquets de données.

- **BMC (Broadcast/Multicast Control).** assure les fonctions de diffusion de messages sur l'interface radio.

La couche L3 contient:

La couche RRC (Radio Resource Control). La fonction principale de la couche RRC est la gestion de la connexion de signalisation entre le mobile et l'UTRAN. Les fonctions remplies par cette couche incluent:

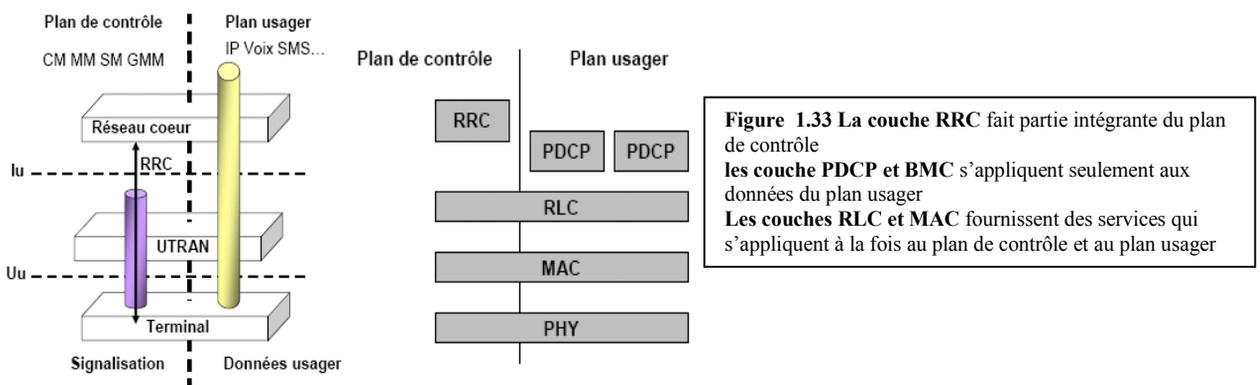
- la diffusion d'informations,
- la gestion des connexions entre l'EU et l'UTRAN, qui comprend leurs établissement, maintenance, et libération,
- la gestion des radio bearers, qui comprend leur établissement, maintenance, libération, et la mobilité correspondante, de la connexion,
- la boucle extérieure de contrôle de puissance,
- évaluation du rapport de mesurage de l'EU.

Deux modes de fonctionnement sont définis pour l'EU: le mode oisif "idle" et le mode dédié. Dans le mode *idle* l'entité pair de la RRC de l'EU est dans le Node B, tandis qu'en mode dédié elle est dans le SRNC.

1.4.5.5 Plan de contrôle et plan utilisateur

L'UMTS sépare en deux plans le flux de données qui transitent par l'interface radio (figure 1.33).

- **Le plan usager:** regroupe l'ensemble des données qui sont échangées au niveau NAS.
- **Le plan de contrôle:** utilisé pour véhiculer l'ensemble de la signalisation entre le mobile et le réseau,
 - **La signalisation AS:** fonctions de l'UTRAN d'établissement de connexion RRC
 - **La signalisation NAS:** qui correspond aux couches de protocoles MM, CM, GMM, SM, assurant les fonctions d'établissement et de gestion d'appel.



1.4.5.6 Canaux physiques

On appelle canal physique [24] l'association d'une fréquence porteuse, d'une paire de codes, l'un pour l'embrouillage et l'autre pour l'étalement (scrambling ou chanelization code), et d'une durée temporelle exprimée en multiple de chip. Des multiples particuliers sont prédéfinis: un slot représente 2560 chips, une trame 38400 chips (15 slots), soit respectivement 0.667 ms et 10 ms.

Canaux physiques de la voie downlink (Forward Link)

- PSCH (*Physical Synchronization Channel*). Le canal PSCH fournit les informations de timing, comme il est aussi utilisé par la SM pour les mesurages de handover.
- Downlink DPCH (*Downlink Dedicated Physical Channel*). A l'intérieur d'un downlink DPCH, données et informations de control (d'un utilisateur) générées dans la couche 2 et la couche 1, respectivement, sont transmises par un multiplexage dans le temps. Ce canal peut donc être vue comme un multiplexage temporel du canal DPDCH (*Downlink Dedicated Physical Data Channel*) et du canal DPCCH (*Downlink Dedicated Physical Control Channel*).
- CPICH (*Common Pilot Channel*). Utilisé comme canal de référence pour une détection cohérente dans la voie downlink et pour le support du control de puissance rapide.
- P-CCPCH, S-CCPCH (*Primary and Secondary Common Control Physical Channel*) et PDSCH (*Physical Downlink Shared Channel*). Diffusion d'information système (primary); paging et réponse des couches hautes aux accès initiaux (secondary).

Canaux physiques de la voie uplink (Reverse Link)

- Uplink DPDCH (*Uplink Dedicated Physical Data Channel*). Ce canal transporte les données générées dans la couche 2 et plus.
- Uplink DPCCH (*Uplink Dedicated Physical Control Channel*). Canal utilisé pour transporter les informations de control (du DPDCH), telles que les commandes de control de puissance, générées en couche 1. Les informations de control de la couche 1 consistent en les bits du pilot connu pour une détection cohérente, les commandes du control de puissance de transmission, etc.

- PRACH (*Physical Random Access Channel*) and PCPCH (*Physical Common Packet Channel*). Ce sont des canaux pour le transport du trafic de données utilisateur. Le schéma d'accès aléatoire WCDMA est basé sur la technique "slotted ALOHA". Plus d'un canal d'accès aléatoire peuvent être mobilisés si la demande dépasse la capacité.
- PUSCH (*Physical Uplink Shared Channel*) (mode TDD). Pour le transport du trafic des données utilisateur.

La notion de contrôle du DPCCH est limitée à la couche physique, même si certains éléments proviennent des couches supérieures. Ces éléments sont les suivants [24]:

- **Une séquence pilote**, connue à la fois du récepteur et de l'émetteur pour synchroniser le récepteur et estimer le canal.
- **TFCI (Transport Format Combinaison Indicator)**, des bits décrivant le traitement appliqué sur les données transmises pour aider le récepteur. En effet, le facteur d'étalement, le type de codage protecteur d'erreur, le facteur et la taille de l'entrelacement et d'autres paramètres, peuvent varier d'un slot à un autre. Pour en informer le récepteur, les TFCI pointent sur des formats préétablis entre l'émetteur et le récepteur.
- **FBI (FeedBack Information)**, des bits de rétroaction, pour les algorithmes bouclés nécessitant une interaction rapide, au rythme d'un slot, entre émetteur et récepteur. Le contrôle de puissance dispose de bits spécifiques.
- **TPC (Transmit Power Control)**, pour le contrôle de puissance en boucle fermée.

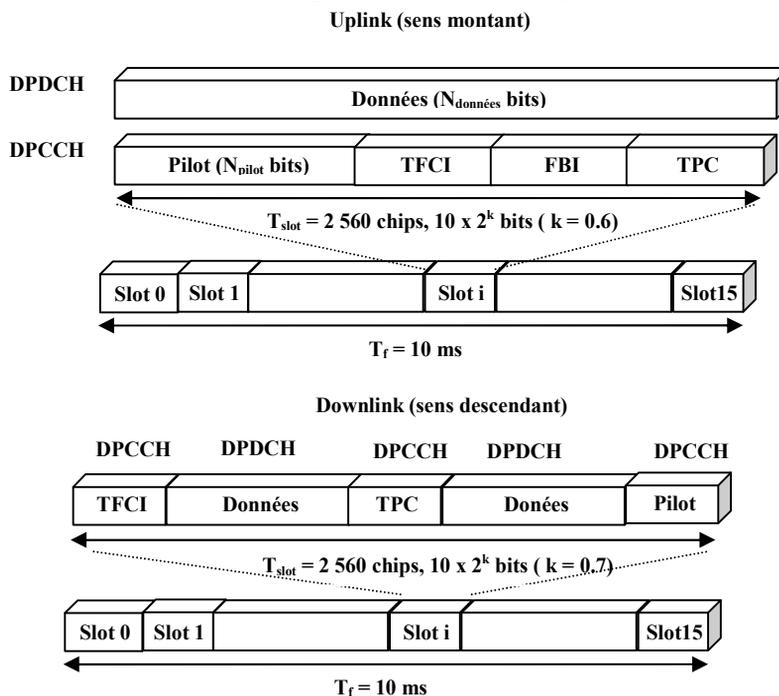


Figure 1.34 : Structure du canal physique dédié UL- DL

L'agencement des informations de contrôle et des données issues des couches supérieures dans le canal physique dédié, varie suivant les sens de la transmission: dans le sens montant, le canal physique de données (DPDCH) et le canal physique de contrôle (DPCCH) sont respectivement placés sur la voie en phase et la voie en quadrature du signal complexe, alors que, dans le sens descendant, la séparation se fait en temps, comme illustré sur la figure 1.34.

Chaque canal physique possède sa propre structure. Par exemple, le canal physique de synchronisation SCH consiste simplement à répéter une même séquence. La structure du canal d'accès aléatoire PRACH est adaptée à l'aloa discrétisé, qui est, comme en GSM et dans les autres systèmes radiomobiles, la technique d'accès utilisée. Néanmoins, si en GSM, deux stations lançant une requête en même temps entrent en collision, ici, le CDMA permet de jouer sur la dimension du code: le paquet d'accès initial est

étalé par un code parmi 16 choisis aléatoirement par le mobile. Ce code, aussi appelé signature, permet de distinguer plusieurs mobiles entrant dans le réseau quasi simultanément. Cela a un coût en terme de complexité de la SB qui doit surveiller en permanence les 16 codes possibles.

1.4.5.7 Canaux de transport

La couche physique, comme nous l'avons vu, offre des services de transport à la couche supérieure, par l'intermédiaire des canaux de transport. Un canal de transport est caractérisé par la façon dont est traitée l'information: taille des blocs transmis, type de codage, profondeur de l'entrelacement, facteur d'étalement, etc. Ces traitements sont décrits dans l'indicateur de format de transport, ou TFI (*Transport Format Indicator*), associé à chaque canal de transport. La couche physique peut combiner plusieurs canaux de transport, requérant une qualité de service similaire, dans un même canal physique. On parle alors de CCTrCH (Coded Composite Transport Channel). Les bits de contrôle TFCI servent au récepteur à démoduler et démultiplexer les différents canaux de transport.

Il n'y a qu'un seul canal de transport dédié, le canal DCH (Dedicated Channel), utilisé pour le transport des informations de l'utilisateur et des informations de contrôle des couches supérieures relatives à cet utilisateur (il n'existe pas, comme en GSM, de notion de canal de contrôle associé à un canal de trafic). Les canaux de transport communs sont:

- BCH (Broadcast Access Channel). Utilisé dans le sens downlink pour la diffusion des informations système propres à une cellule: type de code utilisé, présence des méthode de diversité d'antenne, etc.
- FACH (Forward Access CHannel). Après une demande d'accès initial par le canal RACH, le réseau répond au mobile dans ce canal.
- PCH (Paging CHannel). Ce canal descendant permet au réseau d'appeler un mobile dans la zone de localisation (ensemble de cellules dans lequel se trouve le mobile).
- DSCH (Downlink Shared CHannel). Ce canal descendant, partagé entre plusieurs utilisateurs, transporte des données dédiées à un utilisateur spécifique. Il est toujours utilisé associé avec un des DCH, pour indiquer aux utilisateurs que le DSCH leur est alloué dynamiquement.
- RACH (Random Access CHannel). Canal montant dans lequel un mobile effectue ses requêtes de demande de connexion.
- CPCH (Common Packet CHannel). Canal montant partagé qui étend les fonctionnalités du RACH. Les mobiles peuvent y envoyer des paquets de données sans nécessairement avoir une connexion ouverte.

Lorsque la taille des blocs d'information, définis au niveau des canaux de transport et après codage de canal, ne correspond pas exactement au nombre de bits que peuvent véhiculer les canaux physiques, il faut faire de l'adaptation de débit. Cette opération consiste à modifier le nombre de bit du bloc sur le canal de transport pour qu'il corresponde exactement au nombre de bit d'une trame physique. Le standard UMTS définit précisément ces règles d'adaptation.

Lorsque le bloc d'information du canal de transport est plus grand que la capacité d'une trame, il est réparti sur plusieurs trames. C'est la segmentation. En parallèle, plusieurs canaux de transport peuvent être multiplexés dans un même canal physique par l'utilisation des canaux composites (CCTrCH). Tous les bits sont mélangés par plusieurs entrelacements, ce qui modère sur tous les canaux de transport la probabilité de subir un évanouissement important sur le canal radio.

	DTCH (A)	DTCH (B)	DTCH (C)	3XDCCH
Débit	12.2kb/s	12.2kb/s	12.2kb/s	3.4kb/S
Transport block set	81	103	60	148
20	20	20	20	40
CRC	12	0	0	16
Codage canal	Code convolutif 1/3	Code convolutif 1/3	Code convolutif 1/3	Code convolutif 1/3

Exemple de configuration des canaux de transport.

Mode AMR	Débit source	Classe A	Classe B	Classe C
Amr 12-2k	12.2Kb/s	81	103	60
Amr 10-2k	10.2Kb/s	65	99	40
Amr 7-95k	7.95Kb/s	75	84	0
Amr 7-40k	7.40Kb/s	61	87	0
Amr 6-70k	6.70Kb/s	55	79	0
Amr 5-90k	5.90Kb/s	55	63	0
Amr 5-15k	5.15Kb/s	49	54	0
Amr 4-75k	4.75Kb/s	39	56	0

Ensemble des modes AMR définis pour l'UMTS.

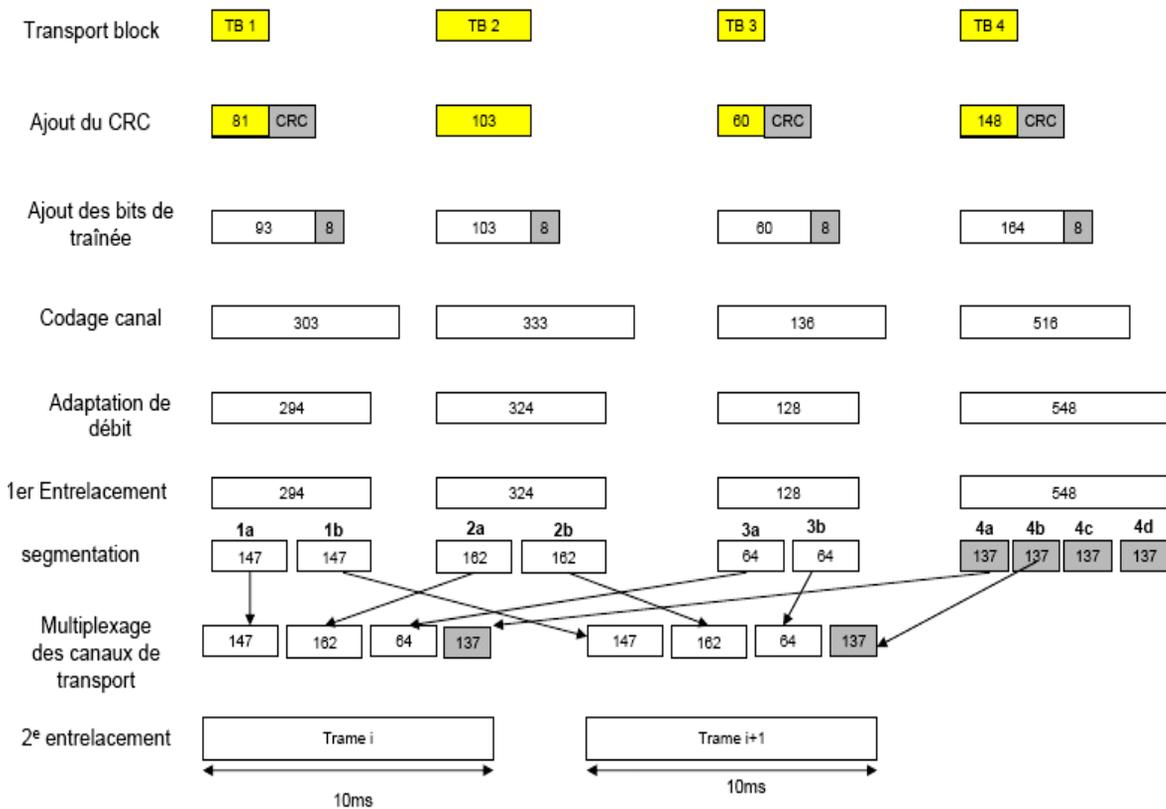


Figure 1.35 Exemple de canaux pour la voix.

Contrôle de puissance

A l'inverse du GSM, l'UMTS a un plus grand besoin de combattre l'effet 'Near-Far'. Un EU, proche d'un Node-B, transmettant avec la même puissance qu'un autre EU, à la frontière de la cellule, masquera ce dernier. Pour maintenir des liens fiables pour tous les EUs, la puissance reçue au niveau du Node-B devra être presque la même (figure 1.36). Cela signifie que ce qu'on appelle "propagation path loss" entre l'EU et le Node-B devra être pris en compte. Dans un environnement idéal, cela suffirait à lui seul, mais en réalité les conditions du canal de transmission varient, dans le court et le long terme. Ceci étant dit, nous pouvons relater les trois mécanismes de contrôle de puissance principaux dans l'UMTS:

- *Boucle ouverte de contrôle de puissance (Open loop power control)*: ce mécanisme se rapporte directement au "path loss". Il sert à établir la puissance initiale à laquelle l'EU transmettra. Cette opération s'effectue via une signalisation RRC. Ce contrôle est pris en charge dans l'EU et le RNC.
- *Boucle extérieure de contrôle de puissance (Outer loop power control)*: se rapporte aux variations à long terme du canal. Un SIR (Signal to Interference Ratio) cible est spécifié. Si le SIR reçu est inférieur à cette cible, la puissance de transmission aura besoin d'être diminuée. Sinon elle devra être augmentée. En pratique, la qualité cible de la voie downlink est évaluée en terme du BLER (Block error ratio). Le BLER peut être rapporté au SIR. Ce contrôle est situé dans l'EU et dans le RNC. Ce mécanisme est aussi connu par slow closed loop power control.
- *Boucle intérieure de contrôle de puissance (Inner loop power control)*: connu aussi par fast closed loop power contrôle, ce mécanisme combat l'évanouissement rapide du signal. Ce contrôle est situé dans l'EU et le Node-B. L'effet de ce contrôle est de maintenir la puissance reçue à un niveau constant afin de réaliser le BLER ciblé, même en cas d'évanouissement du canal.

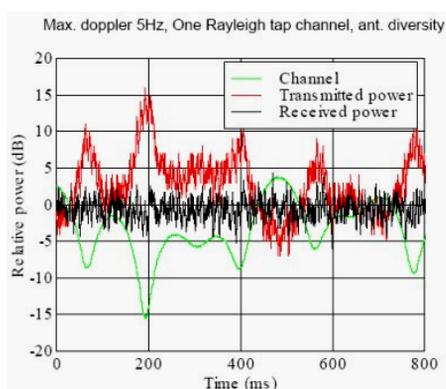


Figure 1.36 Contrôle de puissance rapide dans l'UMTS pour combattre l'évanouissement rapide.

Le contrôle de puissance rapide est important pour maintenir les interférences à un niveau minimum et pour améliorer la capacité. Sans ce moyen, la puissance de transmission sera trop élevée pour satisfaire la qualité recherchée. Il est à noter, que le BLER ciblé peut parfois ne pas être atteint. Par exemple, si le Node-B transmet déjà à sa puissance maximale possible, en aucun cas il ne peut répondre à une commande TPC "UP". Des décisions doivent être prises par la fonction de Control d'Admission dans le SRBC. Il est possible que certains appels doivent être coupés. Les débits doivent être, éventuellement, réduits pour réaliser le BLER ciblé. Il a été montré que la commutation dynamique des *bearers* dans les mauvaises conditions de canal peut améliorer les performances. Par exemple, dans de telles mauvaises conditions, lorsqu'il s'avère difficile de satisfaire le BLER ciblé, le service est basculé de 384 kpps (10 ms TTI, SF 8, 12 TBs) vers 128 kbps (20 ms, SF 16, 4 TBs). Ceci est beaucoup mieux que de couper carrément l'appel.

1.5 TAR mise en oeuvre dans le réseau WiMAX

WiMAX mobile est une technologie d'accès sans fil large bande basée sur les standards, de l'air-interface, IEEE 802.16-2004 et IEEE 802.16-2005e, dont elle implémente un sous ensemble de leurs fonctionnalités et recommandations. L'air-interface du WiMAX mobile, utilise l'OFDMA (Orthogonal Frequency Division Multiple Access) en tant que méthode d'accès radio dans des environnements de non-ligne-de-vue (non-line-of-sight), pour accroître les performances multi-trajets[25]. A son tour, la couche physique du WiMAX est basée sur la technique de modulation OFDM, dont elle hérite l'immunité contre l'interférence inter symboles, et l'évanouissement sélectif (selective fading).

L'utilisateur n'est pas lié aux puissances reçues par les autres utilisateurs. Contrairement à l'approche CDMA, la nécessité d'un contrôle de puissance adéquat est beaucoup moins présente dans les systèmes où

les transmissions des utilisateurs, dans la même cellule, sont maintenues orthogonales. Ceci est particulièrement important dans les systèmes conçus pour s'accommoder à plusieurs utilisateurs ayant chacun des débits moyens très bas: le surplus de traitements nécessaire pour effectuer un contrôle de puissance pour chaque utilisateur peut s'avérer très coûteux pour de tels systèmes. D'un autre côté, il y a une pénalité d'une médiocre réutilisation du spectre dans les systèmes à bande étroite, comparés au système CDMA. A la base, les systèmes à bandes étroites sont mal adaptés à une réutilisation totale des fréquences, parce qu'ils ne fonctionnent pas selon le principe du "average interference". Dans cette section, nous décrivons un système qui combine les points forts de deux systèmes précédents, à savoir: maintenir une orthogonalité entre les transmissions dans une même cellule (à l'image des systèmes à bandes étroites) et permettre une réutilisation totale des fréquences à travers les cellules (tel que le CDMA). Encore une fois, le deuxième point est rendu possible grâce au principe du "average interference".

1.5.1 OFDM (*Orthogonal Frequency Multiplexing*)

La technique OFDM est une variante du multiplexage par division de fréquence dans laquelle un seul canal utilise plusieurs sous porteuses sur des fréquences adjacentes. En plus, dans le système OFDM les sous-porteuses se recouvrent partiellement, pour maximiser l'efficacité spectrale (spectral efficiency). A l'origine, les canaux adjacents peuvent interférer entre eux. Cependant, les sous-porteuses dans le système OFDM sont précisément orthogonales entre elles, et ainsi, elles sont capables de se recouvrir (overlap) partiellement sans interférer[26]. De la sorte, les systèmes OFDM sont capables de maximiser l'efficacité spectrale sans causer d'interférences entre les canaux adjacents. Le domaine de fréquences d'un système OFDM est représenté dans le diagramme de la figure 1.37.

Notons, sur la figure, qu'il y a sept sous porteuses pour chaque canal individuel. Du fait que le débit symbole augmente quand la largeur de bande du canal augmente, cette implémentation permet des débits plus élevés que ceux des systèmes FDM (Frequency Division Multiplexing).

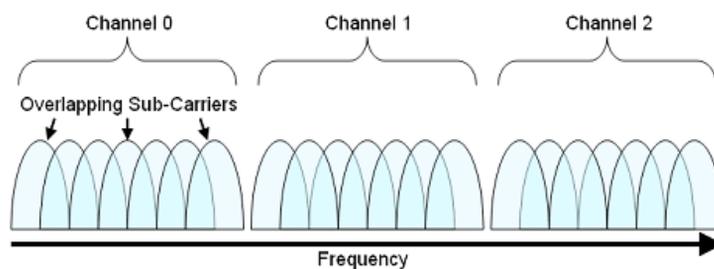


Figure 1.37 Canal à 7 sous-porteuses adjacentes se recouvrant partiellement.

source: . *OFDM and Multi-Channel Communication Systems*. 2007.

<http://zone.ni.com/devzone/cda/tut/p/id/3740>

1.5.1.1 Orthogonalité des porteuses des sous canaux

Les systèmes de communications OFDM sont, comme déjà signalé, capables d'utiliser plus efficacement le spectre de fréquences à travers le recouvrement des sous porteuses. Ces sous canaux peuvent se recouvrir partiellement sans interférer avec les sous porteuses adjacentes, car le maximum de puissance de chaque sous canal correspond directement au minimum de puissance de chacun des canaux adjacents. Comme on peut le voir sur la figure 1.38 chaque sous-porteuse est représentée par pic différent. En plus, le pic de chaque sous-porteuse correspond directement avec le passage à zéro de tous les autres canaux. Notons que les canaux OFDM sont différents des canaux FDM à bande limitée par la façon dont ils appliquent le filtre de mise en forme(d'une implusion). Avec les systèmes FDM, une mise en forme sinc

est appliquée dans le domaine du temps, pour mettre en forme chaque symbole et prévenir les ISI (Inter Symbol Interference). Avec les systèmes OFDM, une mise en forme sinc est appliquée dans le domaine des fréquences pour chaque canal. De ce fait, les sous-porteuses restent mutuellement orthogonales.

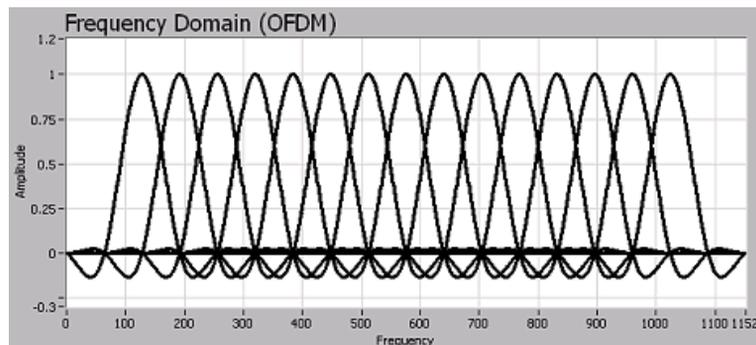


Figure 1.38 sous-porteuses orthogonales
 source: . *OFDM and Multi-Channel Communication Systems*. 2007.
<http://zone.ni.com/devzone/cda/tut/p/id/3740>

1.5.1.2 Implémentation de L'émetteur/récepteur (Génération du signal):

Pour transmettre un canal individuel en utilisant plusieurs sous porteuses, un système de communications OFDM doit effectuer plusieurs étapes, décrites dans la figure 1.39 ci-dessous.

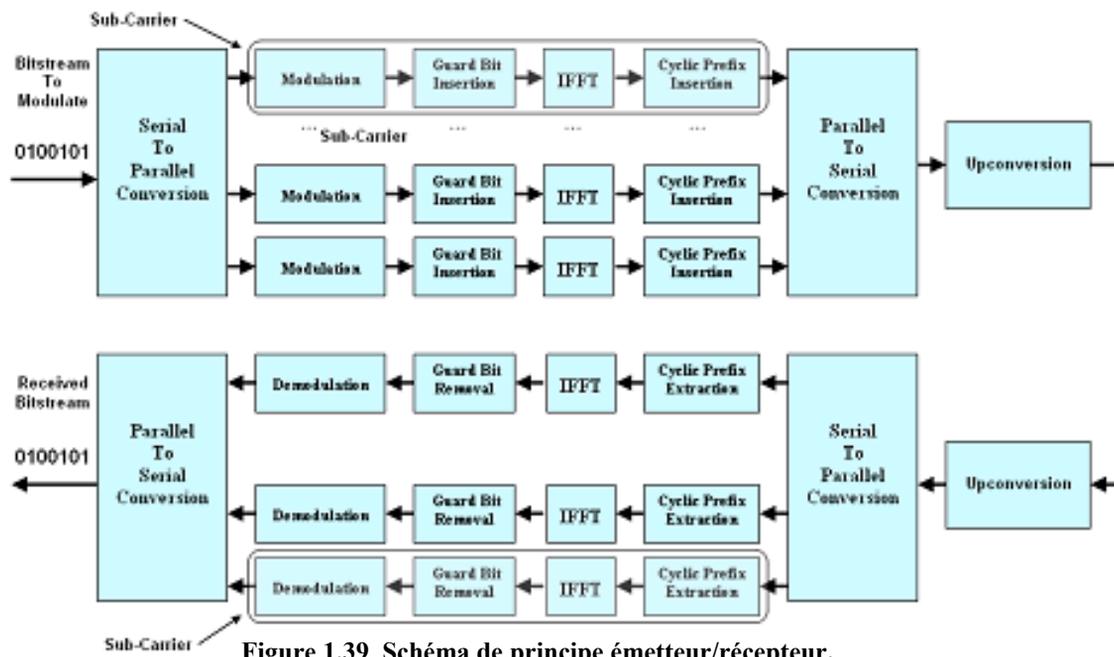


Figure 1.39 Schéma de principe émetteur/récepteur.
 source: . *OFDM and Multi-Channel Communication Systems*. 2007.
<http://zone.ni.com/devzone/cda/tut/p/id/3740>

Conversion Série Parallèle. Dans un système OFDM, l'efficacité dans l'utilisation du spectre de fréquences est accompagnée par des traitements additionnels au niveau de l'émetteur et du récepteur. Ces traitements additionnels sont nécessaires pour convertir un flot de bits en série en plusieurs flots de bits parallèles, à distribuer parmi les porteuses individuelles. Une fois cette opération réalisée, chaque sous porteuse est modulée comme si elles était un canal individuel, avant que tous les canaux ne soient recombines ensemble et transmis comme un tout. Le récepteur effectue le processus inverse pour diviser le signal arrivant en des sous-porteuses appropriées, et les démoduler, ensuite, individuellement avant de reconstruire le flot de bits original.

Modulation avec la FFT (Fast Fourier Transform). La modulation des données en une forme d'onde complexe se produit dans l'étape IFFT (Inverse Fast Fourier Transform) de l'émetteur. Ici, le schéma de modulation peut être choisi de manière tout à fait indépendante du canal spécifique qui est utilisé, et peut être choisi en fonction des exigences du canal. Le rôle de l'IFFT est de moduler chaque sous-canal sur la porteuse appropriée.

Insertion de Préfixe Cyclique. A cause de la sensibilité des systèmes de communications sans fil aux réflexions du canal multi trajets, un préfixe cyclique est ajouté pour réduire l'ISI (Inter Symbol Interference). Un préfixe cyclique est une répétition de la première section d'un symbole qui est ajoutée à la fin du symbole. Il est important parce qu'il permet aux représentations multi trajets du signal original de s'affaiblir pour qu'elles n'interfèrent pas avec le symbole suivant.

Conversion Série Parallèle. Une fois le préfixe cyclique ajouté aux canaux des sous-porteuses, ils doivent être transmis comme un signal. Ainsi donc, l'étape de conversion parallèle série est le processus de sommation de toutes les sous-porteuses, et leur combinaison en un seul signal. Par conséquent toutes ces sous-porteuses sont générées en parfaite simultanéité.

1.5.2 Avantages de l'OFDM

Dans cette section nous allons détailler les deux grands avantages de l'OFDM à savoir: une plus grande efficacité spectrale et une réduction des interférences inter symboles, **et la résistances aux distorsions multi trajets.**

Efficacité Spectrale. Dans les systèmes FDM traditionnels, chaque canal est espacé par environ 25% de la largeur du canal. Ceci est fait pour s'assurer que les canaux adjacents n'interfèrent pas. La figure 1.40 montre les bandes de garde entre les canaux individuels d'un système FDM.

A cause de l'exigence de bandes de garde, il est demandé au débit symbole de permettre l'existence de ces bandes de garde. En général, la largeur de bande permise est $2/R_s$, en conséquence de quoi les canaux

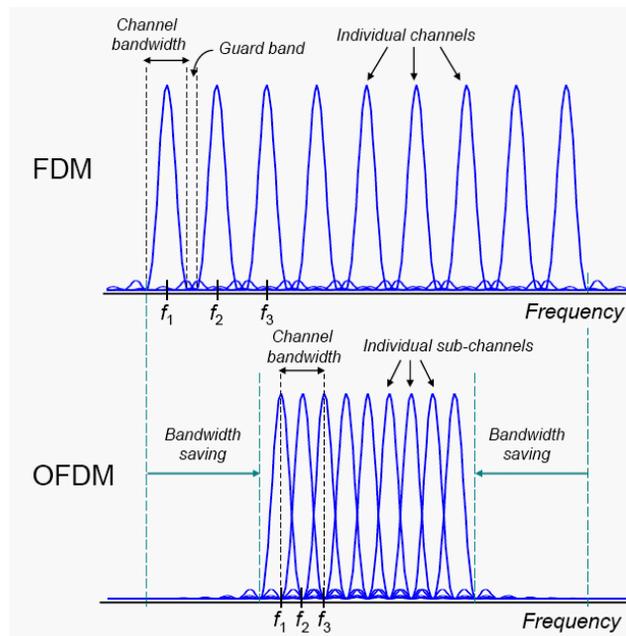


Figure 1.40 OFDM comparé à FDM.

source: Per Hjalmar Lehne, Frode Bøhagen *OFDM(A) for Wireless Communication* research report telenor 2008.

Dans les systèmes OFDM, d'un autre côté, les canaux se recouvrent en fait. Par conséquent, il est possible de maximiser le débit symbole, pour une largeur de bande donnée. Dans la figure 1.40, illustration est faite du recouvrement des sous-porteuses dans un système OFDM. Dans ce scénario, la largeur de bande

du canal avoisine $1/R_s$. Ainsi, lorsque le nombre de sous-porteuses s'approche de l'infini, les systèmes OFDM permettent presque le double de l'efficacité spectrale.

Notons que avec l'OFDM, il est toujours demandé d'avoir une bande de garde entre chaque canal individuel. Cependant, le débit symbole effectif pour les sous-porteuses combinées est plus grand que celui qu'aurait atteint l'utilisation d'une seule porteuse. Notons que le fait d'utiliser des sous-porteuses orthogonales se recouvrant exige aussi l'utilisation d'un préfixe cyclique pour prévenir les ISI. Donc, les avantages gagnés grâce au recouvrement des sous-porteuses sont quelque peu compromis. Cependant, l'avantage de l'efficacité spectrale est tel, qu'un plus grand débit est réalisable dans un système OFDM.

ISI(Inter Symbol Interference) Réduite. Dans les systèmes mono-porteuse, l'interférences inter symbole est souvent causée par les caractéristiques multi trajets (multi-path) du canal de communications sans fil. En effet, quand on transmet une onde électromagnétique sur une longue distance, le signal passe à travers une variété de médiums physiques. De ce fait, le signal réellement reçu contient le signal provenant du chemin direct recouvert par les réflexions du signal d'amplitudes plus petites. La figure 1.41 illustre comment, à de hauts débits symbole, les signaux réfléchis peuvent interférer avec les symboles subséquents.

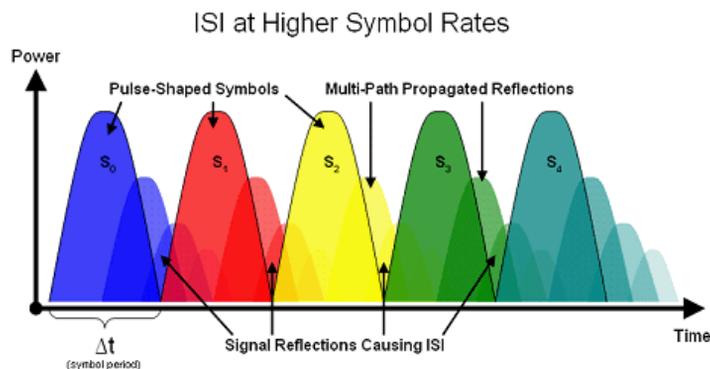


Figure 1.41 Interférences inter symboles à haut débit.

source: . *OFDM and Multi-Channel Communication Systems*. 2007.

<http://zone.ni.com/devzone/cda/tut/p/id/3740>

Dans les systèmes sans fil, ce genre d'interférence crée des difficultés parce que le signal reçu peut être légèrement distordu. Dans ce scénario, le signal du chemin direct arrive comme prévu, mais les réflexions légèrement atténuées arrivent plus tard dans le temps. Ces réflexions posent problème, car elles interfèrent avec les symboles subséquents transmis le long du chemin direct. Ces réflexions du signal sont typiquement adoucies par un filtre de mise en forme, qui atténue aussi bien la section de début que celle de la fin d'une période de symbole. Cependant, comme le montre la figure 1.41, ce problème devient beaucoup plus significatif aux hauts débits symbole. Puisque les réflexions constituent une bonne partie de la période de symbole, l'ISI sera aussi substantielle.

Les systèmes OFDM adoucissent ce problème en utilisant une période symbole comparativement plus

Reduced ISI with Low Symbol Rate (OFDM)

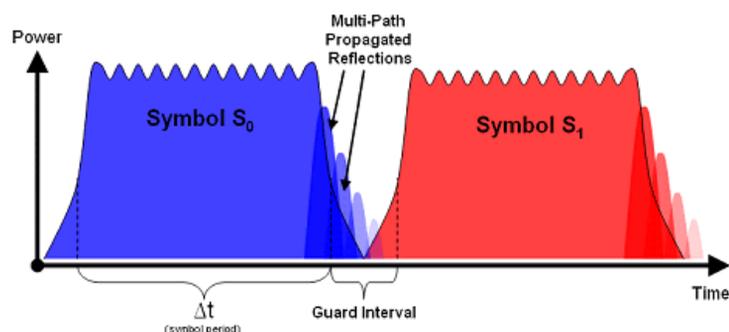


Figure 1.42 ISI réduite à bas débit.

source: . *OFDM and Multi-Channel Communication Systems*. 2007.

<http://zone.ni.com/devzone/cda/tut/p/id/3740>

longue. En plus, ils le font sans sacrifier de débit en utilisant plusieurs sous-porteuses par canal. La figure 1.42 ci-dessous illustre le domaine du temps des symboles OFDM. Notons que dans un système OFDM, le débit symbole peut être réduit tout en continuant à réaliser des débits similaires ou même supérieurs. Remarquons que, d'après l'illustration de la figure 1.42, le temps requis pour l'atténuation complète des réflexions est le même qu'avant. Cependant, en utilisant un débit symbole plus petit, les réflexions du signal constitueront seulement un petit pourcentage de la période symbole totale. Donc, il est possible de simplement ajouter un intervalle de garde pour enlever les interférences des réflexions sans diminuer significativement le débit système.

1.5.3 Technique d'accès OFDMA

L'OFDMA, comme l'OFDM, emploie plusieurs sous-porteuses faiblement espacées, mais avec cette différence que les sous-porteuses disponibles sont divisées en plusieurs groupes appelés sous-canaux. Dans le WiMAX mobile, basé sur la couche physique OFDMA, cette division en sous-canaux est permise aussi bien dans la voie uplink que dans la voie downlink, et là, les sous-canaux forment l'unité de ressource fréquentielle minimum, allouée par la station de base (SB). De la sorte, différents sous-canaux peuvent être alloués à différents utilisateurs en tant que mécanisme d'accès multiple.

Les sous-porteuses qui forment un sous-canal peuvent être soit adjacentes (contiguës), soit distribuées pseudo aléatoirement à travers le spectre de fréquences. Les sous-canaux formés en utilisant des sous-porteuses distribuées fournissent plus de diversité en fréquences, qui est particulièrement utile dans les applications mobiles. Le WiMAX définit plusieurs schémas de sous-canalisation (*subchannelization*), basés sur les porteuses distribuées pour, aussi bien le sens *uplink* que le sens *downlink*. Le schéma appelé PUSC (Partial usage of subcarriers) est obligatoire pour toutes les implémentations du WiMAX mobile. Le schéma basé sur les sous-porteuses contiguës dans le WiMAX est appelé AMC (Adaptive Modulation and Coding). Bien que la diversité en fréquence soit perdue avec ce schéma, l'AMC en bande permet aux concepteurs du système d'exploiter la diversité multi-utilisateurs, allouant les sous-canaux aux utilisateurs en se basant sur leurs réponses en fréquences. La diversité multi-utilisateurs peut fournir un gain significatif dans la capacité de tout le système, si celui-ci s'attelle à doter chaque utilisateur de sous-canaux qui maximisent le SINR reçu. En général, les sous-canaux contigus sont plus appropriés pour les applications fixes et de faible mobilité.

Structure de Slot et de Trame

La couche physique WiMAX est aussi responsable de l'allocation des slots et du tramage (framing) sur l'interface-air. Le minimum de ressource temps-fréquence qui peut être alloué par un système WiMAX à un lien donné est appelé *slot*. Chaque *slot* consiste en un sous-canal sur un, deux, ou trois symboles OFDM, suivant le schéma de "*subchannelization*" utilisé (figure 1.43). Des séries de slots contigus

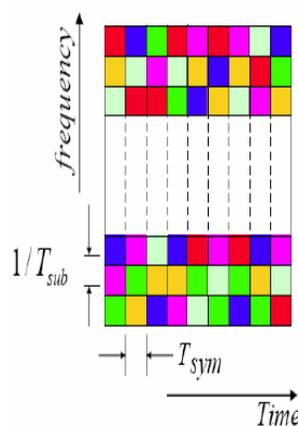


Figure 1.43 Sous-canalisation et notion de slot dans l'OFDMA.
 source: . *OFDM and Multi-Channel Communication Systems*. 2007.
<http://zone.ni.com/devzone/cda/tut/p/id/3740>

assignés à un utilisateur donné forment une région de données de l'utilisateur; les algorithmes de scheduling peuvent allouer les régions de données à différents utilisateurs, sur la base de la demande, des exigences de QoS, et des conditions du canal.

Les trames MAC s'étalent sur deux dimensions: le temps, en unités égales aux symboles OFDMA, et la fréquence, en unités égales aux sous-canaux logiques. Les paquets de données sont convoyés dans des régions de données bi-dimensionnelles(c.a.d. temps et fréquence), qui identifient des rectangles de slots dans la trame et sont annoncées par la SB via des messages de contrôle spécifiques (DL-MAP et UL-MAP). On peut considérer qu'une trame MAC est une matrice dont les lignes et les colonnes sont respectivement les sous-canaux et les symboles OFDMA. En général, la SB est libre de définir la forme et la position de n'importe quelle région de données, c.a.d sélectionner quels éléments de la matrice sont dédiés à cette région. La seule contrainte est que toutes les régions de données doivent tenir dans la sous-trame downlink sans recouvrement, c.a.d que chaque élément est assigné à une région au plus. D'un autre côté, la taille des régions de données (nombre d'éléments), dépend de la quantité de données dans les tampons de transmission de la SB et de la QoS négociés par les connexions admises. Le problème de la définition de la forme et de la position des régions de données de telle manière qu'elles tiennent sur une trame sans recouvrement ni dans le temps ni dans les fréquences, est appelé allocation de région de donnée. Une région de données est toujours transmise avec un même profile de burst.

La figure 1.44 montre une trame OFDMA dans le mode TDD. La trame est divisée en deux sous-trames: une trame *downlink* suivie d'une trame *uplink* après un court intervalle de garde. Le rapport sous-trame-downlink-à-uplink peut être varié de 3:1 à 1:1 pour supporter différents profils de trafic. Le mode TDD permet un partage de la largeur de bande plus flexible entre l'uplink et le downlink, ne requiert pas de spectre appairé, et à une conception plus simple de l'émetteur/récepteur. En contre partie, le mode TDD exige une synchronisation à travers plusieurs stations de base pour assurer une coexistence sans interférences. Les régulations de bandes appairées, cependant, pourraient forcer quelques opérateurs à déployer le WiMAX en mode FDD.

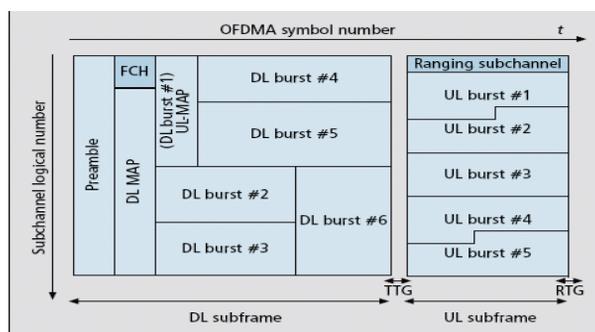


Figure 1.44 Exemple d'une structure de trame OFDMA en mode TDD.

source: ChingYao Huang, Hung-Hui Juan, Meng-Shiang Lin, Chung-Ju Chang. Radio Resource Management of Heterogeneous Services in Mobile WiMAX Systems. IEEE Wireless Communications • February 2007 .

Comme illustré sur la figure 1.44, la trame downlink débute par un préambule downlink qui est utilisé pour les procédures de la couche physique, telles que la synchronisation en temps et en fréquence et l'estimation initiale du canal. Ce préambule est suivi par un FCH (Frame control Header), qui fournit l'information de configuration de la trame, telle que la longueur du message MAP, la modulation et le schéma de codage, et les sous-porteuses utilisables. Les messages MAP dans l'uplink et le downlink (UL-MAP et DL-MAP) spécifient les régions de données allouées aux multiples utilisateurs. Ces messages sont diffusés sur FCH dans la sous-trame downlink. Les messages MAP incluent le "burst" de profile pour chaque utilisateur, qui définit la modulation et le schéma de codage utilisé dans ce lien là.

Le WiMAX est assez flexible quant à la façon dont sont multiplexés les multiples utilisateurs et paquets sur une seule trame. Une seule trame downlink contient de multiples bursts, de tailles et de types variables

transportant des données pour plusieurs utilisateurs. La taille d'une trame varie, sur une base de trame-par-trame, de 2ms à 20ms, et chaque burst peut contenir des paquets de taille fixe ou de tailles variables, ou bien des fragements de paquets reçus des couches supérieures. Initialement au moins, tous les équipements WiMAX supporteront seulement des trames 5ms.

La sous-trame uplink est faite de plusieurs bursts de différents utilisateurs. Une portion de cette sous-trame est laissée de côté pour l'accès par contention qui utilisé à des fins diverses. Cette sous-trame est principalement utilisée comme un canal de "ranging" pour effectuer une boucle fermée d'ajustement de fréquence, de temps et de puissance, durant l'entrée au réseau, et périodiquement par la suite. Ce canal pourrait être utilisé aussi par les stations pour émettre des requêtes de largeur de bande uplink. En plus du canal "ranging" et des bursts de trafic, la sous-trame uplink a un canal indicateur de la qualité du canal CQICH (channel Quality Indicator CHannel) et un canal pour accuser réception ACK. Le premier de ces deux derniers canaux est utilisé par les stations fixes WiMAX, pour retourner l'information sur la qualité du canal au scheduler de la station de base, le deuxième canal (ACK) est utilisé par les stations mobiles pour réagir aux accusés de réception downlink.

Pour s'occuper des variations de temps, WiMAX supporte optionnellement la répétition plus fréquente des préambules. Dans l'uplink, de courts préambules, appelés midambles, pourraient être utilisés après 8, 16, ou 32 symboles; dans le downlink, un court préambule est inséré au début de chaque burst. Il est estimé qu'avoir un miamble tout les 10 symboles autorise un mobilité jusqu'à 15 Km/h.

1.5.4 Allocation des sous-porteuses

Le WiMAX a trois classes de sous-porteuses.

- Sous-porteuses Data: utilisées pour le transport des symboles de données.
- Sous-porteuses Pilot: qui transportent des symboles (pilot symbols) connus à priori, et pouvant être utilisés pour l'estimation et la recherche du canal.
- Sous-porteuses Null.

Pour créer les symboles OFDMA dans le domaine fréquentiel, les symboles modulés sont affectés aux sous-canaux qui ont été alloués à la transmission du bloc de données. Un sous-canal, tel que définit par le standard IEEE 802.16e-2005, est une collection logique de sous-porteuses. Le nombre et la répartition exacte des sous-porteuses qui constituent un sous-canal dépendent du mode de permutation de sous-porteuses. Le nombre de sous-canaux alloués pour la transmission des blocs de données dépend de paramètres tels que la taille du bloc de données, le schéma de modulation et le taux de codage. La constitution de sous-canaux à partir de sous-porteuses s'effectue selon des schémas de permutation de sous-porteuses. Le standard IEEE 802.16e-2005 autorise les schémas suivants :

DL FUSC (Downlink Full Usage of Subcarriers)

Dans ce schéma, toutes les sous-porteuses de données sont utilisées pour créer les divers sous-canaux. Chaque sous-canal est fait de 48 sous-porteuses de données, qui sont distribuées de manière égale à travers la bande de fréquences(voire figure 1.45). Dans le FUSC, les porteuses pilot sont allouées en premier, et puis le reste des sous-porteuses sont affectées aux divers sous-canaux, en utilisant un schéma de permutation. Il est important de souligner que l'ensemble des porteuses pilot est divisé en deux

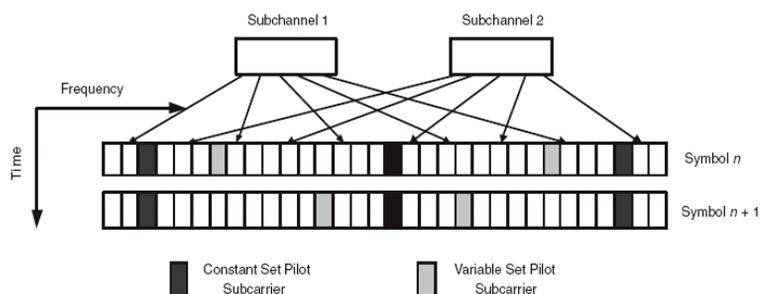


Figure 1.45 Schéma FUSC de permutation de sous-porteuses.

catégories les ensembles constants et les ensembles variables. L'indexe des sous-porteuses de l'ensemble variable change d'un symbole OFDMA à un autre, ce qui permet au récepteur d'estimer la réponse du canal de manière plus pointue à travers la bande de fréquences entière.

DL PUSC (Downlink Partial Usage of Subcarriers)

DL PUSC est similaire au FUSC sauf que toutes les sous-porteuses sont d'abord divisées en six groupes. La permutation des sous-porteuses pour créer les sous-canaux est effectuée indépendamment à l'intérieur de chaque groupe, donc, séparer logiquement chaque groupe des autres. Dans le cas du PUSC, toutes les sous-porteuses, excepté les sous-porteuses null, sont d'abord arrangées dans des clusters. Chaque cluster consiste en 14 sous-porteuses adjacentes sur deux symboles OFDM(figure 1.46). Dans chaque cluster, les sous-porteuses sont partagées en 24 sous-porteuses de données et 4 sous-porteuses pilot. Les clusters sont ensuite numérotés de nouveau en utilisant un schéma de numérotation pseudo aléatoire, qui redistribue l'identité logique des clusters. Après cette opération, les clusters sont divisés en six groupes, avec le premier un sixième des clusters appartenant au groupe 0, et ainsi de suite. Un sous-canal est créé en

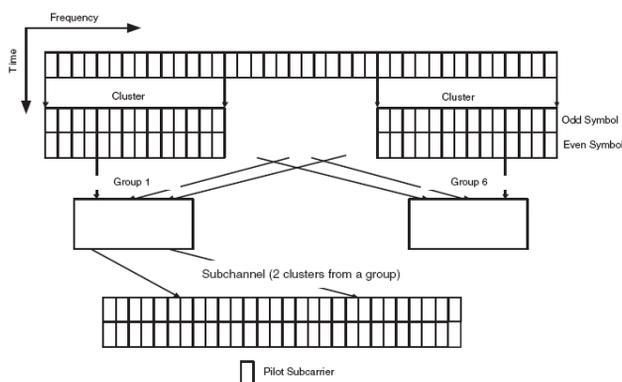


Figure 1.46 Schéma de permutation de sous-porteuses DL PUSC.

utilisant deux clusters du même groupe. En allouant des sous-ensembles disjoints des six groupes disponibles à des émetteurs voisins, il est possible de séparer leurs signaux dans l'espace sous-porteuse, permettant ainsi une réutilisation des fréquences plus serrée au prix du débit de données. Une telle utilisation des sous-porteuses est appelée segmentation. Par exemple, dans une SB, utilisant la segmentation, avec trois secteurs, il est possible d'allouer deux groupes distincts à chaque secteur, réutilisant ainsi la même bande de fréquences dans tous les secteurs. Cette possibilité donnée, avec le système WiMAX est très utile quand le spectre disponible n'est pas assez large pour permettre autre chose qu'une réutilisation de fréquence (1,1).

UL PUSC (Uplink Partial Usage of Subcarriers)

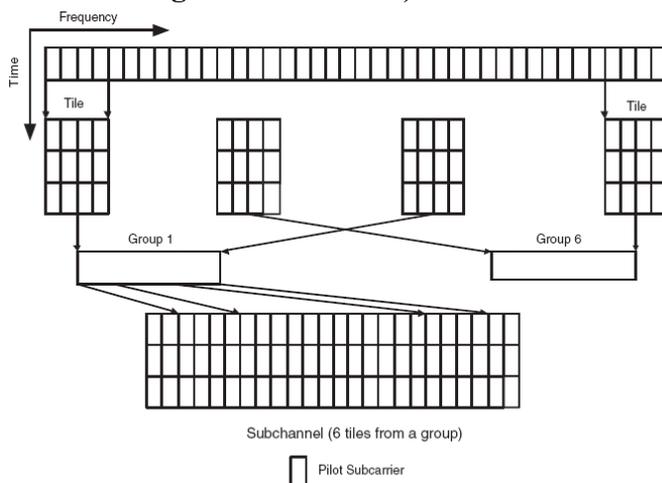


Figure 1.47 Schéma de permutation de sous-porteuses UL PUSC.

Dans ce cas, les sous-porteuses sont d'abord divisées en des carreaux (tiles) variés, comme indiqué sur la figure 1.47. Chaque carreau consiste en quatre sous-porteuses sur trois symboles OFDM. Les sous-porteuses d'un tile sont divisées en huit sous-porteuses de donnée et quatre sous-porteuses pilot. Les tiles sont ensuite numérotés de nouveau, en utilisant une séquence de numérotation pseudo aléatoire, et divisés en six groupes. Chaque sous-canal est créé par l'utilisation de six tiles d'un seul groupe. UL PUSC peut intégrer la segmentation pour opérer dans des cas de réutilisation de fréquence plus serrée.

Chapitre 2

Gestion des Ressources Radio (GRR)

2.1 Introduction

La Gestion des Ressources Radio (GRR) est l'une des fonctions les plus importantes des systèmes sans fil 3G. En effet, c'est sur cette fonction que les systèmes sans fil comptent le plus pour garantir une certaine Qualité de Service (QoS) cible, maintenir la région de couverture planifiée, et offrir une grande capacité. Ces objectifs tendent souvent à être contradictoires, puisque l'amélioration d'un aspect se paie par la détérioration d'un autre (par exemple, la capacité pourrait être augmentée mais au prix d'une réduction de la couverture ou au prix d'une diminution de la QoS). La planification radio du réseau et la GRR vont de pair pour répondre aux précédentes exigences, par une utilisation efficace de ces ressources radio, rares. La planification radio du réseau règle ces paramètres à un haut niveau (statiquement et à travers le composant OAM), alors que la fonction GRR fournit un mécanisme de réglage plus fin, pour équilibrer ces exigences de façon dynamique.

Les techniques de GRR visent à optimiser l'utilisation des ressources radio tout en satisfaisant les exigences de qualité, et du plus grand nombre d'utilisateurs possible, dans un environnement caractérisé par une disponibilité limitée du spectre, des ressources limitées aussi (énergie, puissance calcul, etc), des effets de distorsion du signal dus au canal multi trajets (multipath), les affaiblissements et interférences, la mobilité des utilisateurs, et l'hétérogénéité du trafic [2]. Du point de vue des utilisateurs, ils veulent un service avec les meilleures caractéristiques possibles, telles que le maximum de débit ou un taux de blocage et de coupure d'appel bas. Pour les fournisseurs de services ou les opérateurs, ils désirent servir le maximum d'utilisateurs possible en un temps donné, avec la qualité de service convenue, et au prix le plus rentable possible [3]. Ainsi donc, l'évolution dynamique des réseaux sans fil (notamment dans leur dimension interface air), appelle une gestion dynamique des ressources radio.

Unité de ressource radio. Une RRU (Radio Ressource Unit), pourrait être définie par l'ensemble des paramètres de transmission physique de base nécessaires pour supporter un signal transportant l'information de l'utilisateur final, correspondant à un service de référence [1]. Ces paramètres de transmission physique dépendent de la technique d'accès multiple utilisée. En particulier, dans l'FDMA (Frequency Division Multiple Access), une RRU est équivalente à une certaine largeur de bande autour d'une fréquence de porteuse donnée. Dans le TDMA (Time Division Multiple Access), une RRU est équivalente à un couple: fréquence de porteuse et slot de temps (time slot). Dans le CDMA (Code Division Multiple Access), une RRU est définie par une fréquence de porteuse, une séquence de codes et un niveau de puissance. Dans l'OFDMA (Orthogonal Frequency Division Multiple Access), une RRU pourrait être définie par une fréquence de porteuse, une durée de symbole, un débit, un mode de modulation, etc.

Comment les ressources radio sont-elles pourvues? L'objectif d'un opérateur de réseau sans fil est de déployer un tel réseau pour être en mesure de fournir à ces utilisateurs la QoS qu'il faut dans la région de couverture ciblée. A cette fin, l'ensemble du réseau implique plusieurs sous-problèmes couvrant les conceptions, du réseau sans fil, du réseau de transmission, et du réseau cœur. Pour la partie radio, le résultat de la planification du réseau sans fil s'exprimera en termes de fourniture des RRUs à travers la région de service, au moyen d'une certaine topologie de réseau et une configuration donnée, de sites de cellules. Il importe de noter qu'en raison de la variation dans le temps et dans l'espace de la pénétration du service et de son utilisation, la quantité de RRUs à fournir varie aussi, et que par conséquent, la planification radio du réseau est un processus évolutif. Néanmoins, l'inertie associée au déploiement du réseau sans fil (e.g. acquisition des sites, les travaux de génie civil de préparation des sites, etc) font que ce processus n'est capable de répondre qu'aux situations caractérisées par des paramètres d'entrée, de planification radio du réseau, assez stables ou variant à long terme. Dans ce contexte, la QoS est en général garantie par un surdimensionnement du réseau, et un excès de fourniture de RRUs. Il est clair, cependant, que le défi est d'être capable de fournir le niveau de QoS avec le minimum de ressources

possibles, minimisant, ainsi, l'investissement de l'opérateur tout en respectant les exigences de conception du réseau.

2.2 Les composants de la GRR

La GRR pour les communications sans fil consiste en un certain nombre de composants interdépendants et coopérants, dont les comportements conjoints devraient mener à une optimisation de tout le réseau d'accès radio. C'est composants sont, le contrôle d'admission, le contrôle de puissance, le contrôle du handover, le contrôle de charge, et les fonctionnalités d'ordonnancement des paquets (figure 2.1). Ces composants sont dédiés à la prise en charge les aspects suivants [3] :

- Les schémas d'allocation de la ressource temps/fréquence tels que l'allocation du canal, l'ordonnancement, les schémas de contrôle de la vitesse de transmission et la réservation de bande passante qui déterminent comment le trafic de données est contrôlé et géré,
- Les schémas de contrôle et d'allocation de puissance qui contrôlent la puissance de transmission des terminaux et des stations de base,
- Le contrôle d'admission, l'assignation de station de base, la gestion de la localisation et les schémas de handover/hand-off qui contrôlent l'accès et gèrent la mobilité.

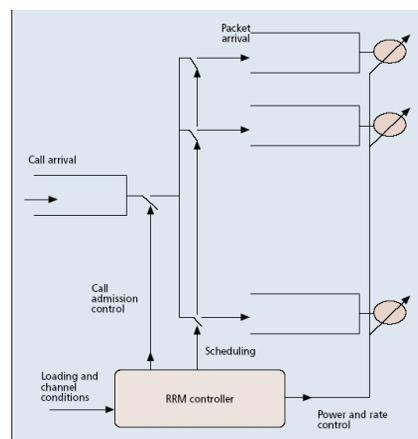


Figure 2.1 Modèle de Gestion des Ressources Radio

Les algorithmes de GRR peuvent être basés sur le degré d'importance du hardware présent dans le réseau (Hard blocking) ou bien sur les niveaux d'interférences dans l'interface air (Soft blocking). Le Hard blocking est défini comme le cas où c'est le hardware qui limite la capacité avant que l'interface air ne se surcharge. Le *soft blocking*, par contre, est défini comme le cas où la charge de l'interface air est estimée être au dessus de la limite planifiée. Il est montré [3] que la GRR basée sur le *soft blocking* apporte plus de capacité que la GRR basée sur le *hard blocking*, mais que la première nécessite un mesurage permanent de la charge de l'interface air.

Dans le reste de ce chapitre nous allons essayer de donner une vue générale des fonctionnalités liées à la GRR dans le contexte d'une seule technologie d'accès, avant d'insister sur les aspects spécifiques aux contextes hétérogènes où plusieurs technologies d'accès entre en jeu.

2.3 Le Contrôle d'Admission

Le contrôle d'admission régule le fonctionnement du réseau de telle manière qu'un service ininterrompu soit fourni, et que les nouvelles demandes de connexions soient accueillies de façon optimale. Le Contrôle d'Admission cherche à déterminer s'il serait possible d'admettre un nouvel utilisateur dans le réseau sans dégrader la qualité exigée par les utilisateurs déjà connectés. Si l'admission d'une nouvelle connexion est de nature à dégrader la qualité du service dans la cellule à un point tel, que la qualité des connexions en cours soit réduite et que la qualité de la connexion demandée elle-même ne soit pas garantie, l'utilisateur ne sera pas admis dans le système. Le Contrôle d'Admission pourrait être amené à bloquer les nouvelles demandes d'accès pour prévenir la congestion des connexions existantes. En effet, si la charge de l'interface air est autorisée à augmenter excessivement, la zone de couverture de la cellule

est réduite en déca des valeurs planifiées, et la QoS ne pourra plus être garantie. Par conséquent, le taux de blocage est augmenté, mais en contre partie le taux de coupure d'appels est diminué. Du point de vue des utilisateurs, le taux de coupure de connexion est plus critique, car il est plus ennuyant de perdre une connexion en cours, que d'échouer à en établir une (nouvelle).

Dans les réseaux sans fil de première et seconde génération, le CA a été développé pour des environnements avec un seul service. Dans les systèmes de troisième génération, des services multimédia tels que la voix, vidéo, données, et audio sont offerts avec des profils de QoS variés. Donc, des schémas de CA plus sophistiqués sont développés pour prendre en charge ces changements.

Les principes du contrôle d'admission adoptent des critères prédéfinis variés pour décider de l'admission ou non de nouvelles connexions au réseau (interférences ou charge que ces dernière ajouteraient aux connexions existantes, disponibilité des ressources, impact sur la QoS en cours, priorité des utilisateurs, etc). Dans ce qui suit nous allons passer en revue quelques approches représentatives des différentes implémentations de CA, proposées dans les travaux dédiés à cet aspect.

2.3.1 CA pour contrôler la Qualité du signal

Le CA est essentiel pour garantir la qualité du signal dans les réseaux sans fil à interférences limitées (interference-limited). Par exemple, les réseaux CDMA ont une capacité soft limite qui fait que plus le réseau est chargé, plus la qualité du signal est détériorée (pour les utilisateurs) en termes de niveau d'interférences ou du ratio signal à interférences (SIR). Par conséquent, Les schémas de CA de cette famille n'admettent de nouveaux utilisateurs dans le réseau que si une qualité minimum du signal peut être maintenue (aussi bien pour le nouvelle connexion que pour les connexions existantes). Dans cette approche les critères d'admission pourraient être:

2.3.1.1 Interférences et SIR (Signal to Inteference Ratio): Cette méthode vise à maintenir une valeur minimum du ratio SIR (SIR_{min}) pour contrôler la qualité du signal. A cette fin, on vérifie la valeur du SIR qu'une nouvelle connexion réaliserait en cas d'accès, et si cette valeur est supérieure à SIR_{min} , la demande d'accès est acceptée, et elle n'est acceptée qu'à cette seule condition. Dans [4] deux schémas basés sur le SIR ont été proposés, utilisant comme critère d'admission la capacité résiduelle (R_k) définie par:

$$R_k = \left[\left(\frac{1}{SIR_{th}} - \frac{1}{SIR_k} \right) \right] \quad (2.1)$$

Où SIR_k est le SIR du uplink dans la cellule k , SIR_{th} est le SIR seuil, qui est un paramètre de conception ($SIR_{th} > SIR_{min}$). La capacité résiduelle R_k est calculée périodiquement, et lorsqu'un nouvel utilisateur arrive dans la cellule k , la SB k contrôle si R_k est supérieure à zéro pour admettre le nouvel appel sinon cet appel est rejeté. Le second algorithme modifie le définition de la capacité résiduelle pour prendre en compte l'impact de l'admission d'un nouveau appel sur la cellule k elle-même ainsi que sur les cellules qui lui sont adjacente. Le critère d'admission choisi est la capacité résiduelle minimum des cellules cibles ainsi que des cellules adjacentes, où la capacité résiduelle dans les cellules adjacentes est modifiée comme suit :

$$R_{k,j} = \left[\frac{1}{\beta} \left(\frac{1}{SIR_{th}} - \frac{1}{SIR_k} \right) \right] j \in C(k) \quad (2.2)$$

Où β représente le couplage d'interférences entre des cellules adjacentes et $C(k)$ est l'ensemble des cellules adjacentes à la cellule k . Si la capacité résiduelle minimum est positive, la SB k accepte le nouveau utilisateur, sinon il est rejeté. Le premier schéma se comporte mieux lorsque le trafic est homogène alors que le second s'adapte plus à des zones "hot spot".

Au lieu d'utiliser le SIR en tant que critère d'admission, on peut utiliser la probabilité d'indisponibilité de service ($\Pr(SIR < SIR_{min})$) [5]. Cette probabilité est estimée pour chaque classe de service sur la base du nombre d'utilisateurs de la classe, et de la puissance allouée à chaque utilisateur actif. Un appel est admis

dans le système si la probabilité d'indisponibilité de service de chaque classe est inférieure au niveau requis correspondant.

Les schémas basés les interférence, utilisent le niveau de ces interférences en tant que critère d'admission pour garantir une qualité de signale en termes de SIR. Pour s'accommoder aux services à débits variables et au trafic non stationnaire et non uniforme, une bonne estimation du niveau d'interférence des appels arrivants doit être utilisée. Comme dans[6], deux parties sont distinguées: l'estimation et la décision. Dans la partie estimation, la moyenne et la variance du niveau d'interférence sont estimées en utilisant les valeurs correspondantes des connections actives. Les mêmes valeurs dues à un nouvel appel sont prédites. L'interférence totale est estimée grâce à la somme des interférences estimées de tout les appels actifs, de celles prédite due à une nouvelle demande de connexion, et un seuil de réservation (pour la variance estimée et les erreurs). Ensuite la partie décision vient trancher la question de savoir s'il faut admettre un nouvel appel en comparant le total des interférences au seuil d'interférence maximum I_{max} .

2.3.1.2 Charge de la cellule (ou le réseau): La qualité du signale en termes de SIR peut être réalisée en contrôlant la charge de la cellule/réseau, du moment qu'un nombre d'utilisateurs plus grand entraîne un SIR plus bas. Des schémas basés sur la charge ont été proposés [7,8]], pour contrôler le nombre d'utilisateurs actifs, en utilisant deux valeurs seuils M_v et M_d pour la voix et les données respectivement. M_v et M_d sont des paramètres de conception dont les valeurs relatives déterminent la priorité de l'un des deux services sur l'autre. La charge effective, définie comme la somme pondérée des nombres d'utilisateurs dans la cellule, avec le poids le plus élevé affecté aux utilisateurs qui sont loin de la SB (plus proches de la bordure de la cellule) a été employée en tant que critère d'admission.

Pour un trafic multi classes la charge de la cellule est exprimée en termes d'une nouvelle fonction appelée Bandwidth Utility (BU) [9] définie comme suit:

$$BU = \frac{1}{MN} \sum_k \frac{SIR_k}{1 + SIR_k} \quad (2.3)$$

Où M est le nombre d'éléments d'antenne, N est le facteur d'étalement, et SIR_k est la valeur SIR minimum des utilisateurs de la classe k . La dépendance du maximum de BU des interférences des autres cellules (I_{other}) est déterminée, où le maximum de BU est calculé en essayant toutes les combinaisons des nombres d'utilisateurs de chaque classe. Ensuite, la valeur maximale de BU pour laquelle le SIR cible est encore réalisable est choisie. Lorsqu'un nouveau appel arrive au système, I_{other} est mesuré et le maximum de BU est trouvé. Le nouvel appel est admis si la valeur courante de BU dans le réseau (due aux appels actifs) plus le BU du nouveau appel est inférieur au maximum de BU; sinon l'appel est rejeté.

2.3.1.3 Faisabilité de l'allocation de puissance (Power Allocation Feasibility): Comme son nom l'indique, cette méthode utilise la faisabilité d'allocation de puissance en tant que critère d'admission[10]. Dans le schéma SAS (Soft And Safe), quand un nouveau utilisateur arrive il est autorisé à transmettre en utilisant le DCPC (Distributed Constrained Power Contrôle) mais avec une puissance maximum limitée, qui augmente graduellement. Après conversion de la puissance, si le SIR cible est réalisé, l'appel est admis; sinon il est rejeté. Dans un autre schéma [11], la puissance est mise à jour en utilisant le DCPC alors que la puissance du nouvel utilisateur est surveillée pendant N itérations. Ensuite, deux procédures d'admission sont utilisées. Dans la première, appelée TPCAC (Transmitted Power CAC), l'utilisateur est bloqué si la puissance transmise est plus grande la valeur seuil; sinon, le nouveau utilisateur est admis. Dans la seconde procédure, appelée RPCAC (Received Power CAC), la puissance reçue est utilisée comme un critère d'admission où le nouveau utilisateur est rejeté si la puissance reçue est plus grande qu'une valeur seuil. Des schéma partant du même principe mais non interactifs sont aussi utilisables[12]. L'algorithme calcule les niveaux de puissances optimums qui satisfont les contraintes de probabilité d'indisponibilité de service pour tout les utilisateurs itérativement. Aussi bien les interférences inter cellules qu'intra cellule sont supposées suivre une certaine distribution (Gaussian distribution). Quant les appels arrivent au réseau, les niveaux de puissance optimums sont déterminés en utilisant le schéma d'allocation de puissance proposé. Ensuite la contrainte de la puissance totale(P_{tot}) est vérifiée en

supposant que P_{tot} aussi suit la même distribution. Si on peut assigner au nouvel utilisateur le niveau de puissance sans violer la contrainte, l'appel est admis dans le réseau; sinon, il est bloqué.

2.3.1.4 Contrôle d'Admission Optimum avec des contraintes de qualité du signal: La qualité du signal dans ces schémas de CA est contrôlée en résolvant un problème d'optimisation avec contraintes. De tels schémas optimisent une certaine fonction objective, étant donné des contraintes sur des mesurages de qualité de signal. Une telle approche a été utilisée[13], pour maximiser la capacité du système tout en maintenant une borne supérieure à la probabilité d'indisponibilité de service et à la probabilité de blocage. Des techniques de programmation linéaire modifiées ont été utilisées pour résoudre le problème d'optimisation et trouver la politique optimum. Une politique de CA est employée[14], pour minimiser le taux de blocage des appels vocaux tout en maintenant la qualité du signal en termes probabilité d'erreur paquet. La politique de CA est basée sur un processus de décision *semi-markovien*. Une fonction de coût, qui est égale au nombre d'appels bloqués, est minimisée en utilisant l'algorithme *value-iteration*.

Dans les réseaux TDMA traditionnels, une qualité de signal minimum est garantie par la présence d'un plan de réutilisation de fréquences épars (avec par exemple un facteur de réutilisation de 7 ou plus). Cette approche peut garantir un SIR_{min} d'une manière stochastique (lorsque les phénomènes de shadowing et fading sont considérées) au prix d'une limitation du nombre de canaux par cellule. Un plan de réutilisation de fréquences étroit (avec un facteur de réutilisation de 3 ou 1) peut être employé pour augmenter le nombre de canaux disponibles par cellule. Dans ce cas, la capacité du réseau TDMA est limitée par la qualité du signal (soft capacity) plutôt que par la disponibilité du canal (hard capacity) dans les réseaux TDMA traditionnels. Par conséquent, des schémas de CA, en même temps que d'autres techniques de gestion des ressources, doivent être employés pour garantir un SIR_{min} , puisque les co-canaux interférant ne sont plus assez distants l'un de l'autre.

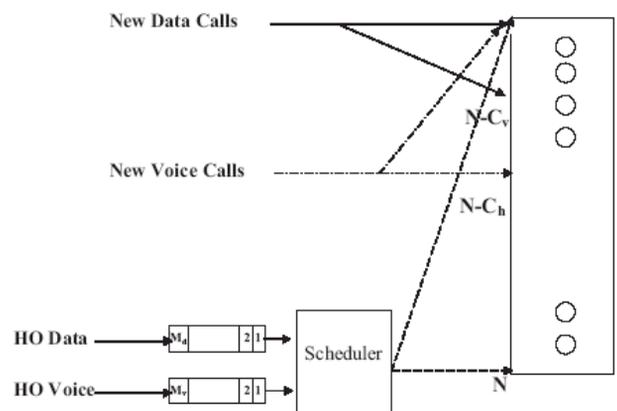
2.3.2 CA pour contrôler la probabilité d'échec du handover

Pour contrôler la probabilité d'échec du handover (P_{hf}), certaines mesures doivent être prises pour donner la priorité aux appels de handoff par rapport aux appels nouveaux, du moment qu'il est communément accepté que forcer la terminaison d'un appel en cours est plus ennuyant pour l'utilisateur que d'être bloqué lors d'un nouveau appel. Différentes approches sont utilisées pour contrôler la probabilité de handoff dans les réseaux sans fil. Dans toutes les approches, il y a un compromis entre la réduction de cette probabilité et la probabilité de blocage d'appel.

2.3.2.1 Politique des Bandes de Garde (Guard Band policy): Cette politique a été proposée par Hong et Rappaport dans [15]. Cette technique garde une certaine quantité de canaux pour les appels provenant de handovers seulement, alors que le reste des canaux peuvent être partagés par les deux types d'appels, c.a.d les nouveaux appels et ceux provenant de handovers. Ainsi, les appels de handoff sont rendus plus

- C_h est le nombre de canaux de garde utilisés exclusivement par les demandes de HO.
- C_v est le nombre de canaux qui ne devraient pas être utilisés par les nouveaux appels données tels que $C_v \geq C_h$.
- N est le nombre de canaux disponibles par cellule.
- M_v et M_d sont respectivement les capacités des files d'attente voix et données.

Figure 2.2 Politique des bandes de garde (Guard Band policy)



prioritaires par rapport aux appels nouveaux, et on peut dire que la réduction de la probabilité de handoff se fait au prix d'un plus grand taux de blocage. Par conséquent, la bande de garde (nombre de canaux) réservée pour les appels de handover doit être proprement choisie, comme un compromis entre la

probabilité de blocage d'un nouveau appel et la probabilité d'échec d'un handover. Cependant, même si la bande de garde est choisie proprement pour contrôler les deux probabilités P_b et P_{hf} se basant sur certaines suppositions sur les paramètres de charge de trafic (taux d'arrivés des nouveaux appels et des handoff), ceci ne garantit pas P_b et P_{hf} , du moment que les paramètres de charge de trafic sont variables. Des schémas à bande de garde dynamique sont aussi utilisés, qui adaptent le nombre de canaux réservés aux paramètres du trafic (dans la cellule seule ou avec les cellules voisines), de même qu'a été prise en charge l'adaptation au cas de plusieurs classes, de services, différentes (figure 2.2).

2.3.2.2 L'état de charge dans la cellule hôte et les cellules voisines en tant que critère de CA: Une manière de contrôler la probabilité de handoff est de prendre en compte dans le processus d'admission l'information de charge dans les cellules voisines. En effet, l'information sur l'état de charge dans les cellules voisines peut refléter l'admissibilité des appels de handoff potentiels de la cellule cible (cellule du nouveau appel arrivant) vers ces cellules voisines puisque, plus les cellules voisines sont chargées plus la probabilité d'échec d'un handoff est grande et vice versa. Cette information de charge donne aussi une indication sur le nombre de handoff potentiels en provenance des cellules voisines vers la cellule hôte. Le nombre d'utilisateurs dans la cellule hôte et ses cellules voisines a été utilisé[16,17] pour déterminer le nombre maximum de nouveaux appels qui peuvent être admis dans la cellule hôte. Ce nombre est déterminé en utilisant la probabilité, estimée, de surcharge ($P[\text{nombre d'utilisateurs} > N]$, où N est le nombre de canaux par cellule) comme une approximation de P_{hf} . Un appel de handoff est admis tant qu'il y a un canal libre.

2.3.2.3 Ressources disponibles dans la cellule hôte et les cellules voisines: Dans les systèmes sans fil à services multiples avec des exigences, en termes de ressources, différentes, P_{hf} peut être minimisé en assurant que les ressources nécessaires pour un nouveau appel seront disponibles pendant toute la durée de l'appel. Les ressources exigées pourraient (ou non) être réservées pour cet appel particulier dans la cellule hôte, ainsi que dans les cellules voisines. Un CA basé sur la réservation de ressources a été proposé dans[18]. Ce schéma vérifie si la bande passante requise dans la cellule hôte et dans les cellules voisines est disponible pour le trafic temps réel. Les nouveaux appels non temps réel sont admis dans la cellule hôte si celle-ci seulement elle possède la bande passante nécessaire. Un handoff est acceptés si le minimum de bande passante est disponible dans la cellule hôte, et si les autres cellules peuvent réserver la bande passante pour le trafic temps réel. Les appels de handoff non temps réel sont acceptés si n'importe quelle bande passante est disponible. Deux schémas de réservation ont été envisagés. Le premier schéma réserve dans chaque cellule la somme de la bande passante requise par tous les nouveaux appels temps réel, dans toutes les cellules voisines. Le second schéma utilise le multiplexage statistique et réserve seulement une portion de la somme de la bande passante requise. L'historique des handovers joue un grand rôle dans l'estimation de la probabilité de handoff à partir ou vers telle ou telle cellule. Dans certaines variantes, la réservation est effectuée dans les cellules voisines vers lesquelles des handoff de la cellule hôte vers celles-ci ont une grande probabilité de se produire. Dans d'autres propositions[19], lorsqu'un nouveau handoff arrive et qu'il ne trouve pas assez de ressources, les utilisateurs données sont instruits par la SB de réduire leur débits pour servir l'appel de handoff. Un tel appel n'est rejeté que si le débit réduit est inférieur à un débit minimum prédéfini r_{min1} . De manière similaire, les nouveaux appels sont admis même s'il n'y a pas assez de ressources en réduisant la vitesse de transmission des utilisateurs données, avec un débit minimum r_{min2} . En fixant les débits minimums tels que $r_{min2} > r_{min1}$ les handoff seront plus prioritaires que les nouveaux appels.

2.3.2.4 Estimation de la probabilité d'échec du handover / de surcharge: Bloquer des appels nouveaux qui auraient été, avec une grande probabilité, coupés par la suite est une manière efficace de contrôler P_{hf} , tout en maximisant l'utilisation des ressources. Ainsi, plusieurs schémas CA ont adoptés cette approche pour réduire P_{hf} . Le concept du shadow cluster[20] part du principe que l'arrivée d'un nouvel appel dans une cellule cible, crée une ombre (shadow) recouvrant la zone formée par les cellules que cet appel visitera vraisemblablement au cours de sa session de communication (shadow cluster).

L'idée, est de se baser sur les informations relatives à la mobilité pour estimer les besoins, futurs, en ressources dont aura besoin le nouvel appel dans chacune des cellules constituant le shadow cluster (figure 2.3). En se basant sur l'estimation des ressources, et l'information sur la mobilité, la survie de l'appel est déterminée en tant qu'indicateur du succès de la terminaison normale de l'appel, en prenant en compte les handovers potentiels du nouvel appel vers les cellules voisines. Des schémas similaires ont été proposés, avec cette différence que la transition estimée de la cellule cible vers les cellules voisines est supposée être prévisible d'après les statistiques, recueillies sur une longue période, de fonctionnement du réseau, d'un modèle de mobilité macroscopique, ou par un entraînement d'un modèle de Markov caché, plutôt que d'après des informations sur la mobilité comme c'est le cas avec le concept du shadow cluster.

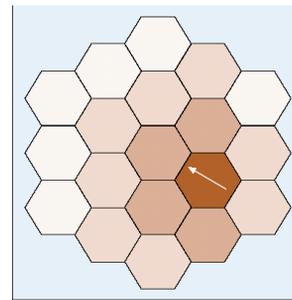


Figure 2.3 Shadow cluster formé par un terminal mobile

2.3.3 CA pour contrôler les paramètres de QoS du niveau paquet: Le CA peut aussi contrôler les paramètres de QoS du niveau paquet, tels que le délai paquet, le délai jitter, le taux de perte de paquets, etc. Le CA estime les valeurs de ces paramètres sur la base des ressources disponibles dans les réseaux sans fil ATM. Si ces paramètres peuvent être garantis, l'appel est admis; sinon il est rejeté.

2.3.3.1 CA pour le contrôle de la vitesse de transmission: Les schémas de CA peuvent assurer une vitesse de transmission minimum soit, en limitant la charge du réseau, soit en minimisant la dégradation de vitesse de transmission (définie comme la condition d'avoir la vitesse de transmission au dessous de la vitesse minimum), soit encore en estimant la vitesse de transmission allouée comme un critère d'admission. Le nombre d'utilisateurs par cluster est limité par une valeur maximale telle, qu'une vitesse de transmission minimum puisse être offerte à tout les utilisateurs(existants et nouveaux) même s'ils se déplacent vers une autre cellule, voisine[21,22]. Les handovers sont admis en premier (pour minimiser P_{hf} , puis les nouveaux appels sont admis sur la base de la disponibilité des ressources. S'il y a assez de ressources et que certains utilisateurs ont une priorité supérieure aux valeurs minimum déclarées, un processus de réduction des ressources est déclenché pour que les nouveaux appels puissent être accommodés.

2.3.4 Schémas de CA basés sur le revenu:

Lorsqu'une admission d'appel est demandée, ceci représente un gain potentiel pour les revenus du réseau à cause de la rémunération due à l'utilisation des ressources. Entre temps, les appels arrivant pourraient causer une congestion et une dégradation dans les ressources allouées aux appels déjà admis, ce qui cause à son tour une réduction dans les revenus du réseau. Par conséquent, la politique du CA peut jouer un rôle majeur dans l'optimisation des revenus des réseaux sans fil. Le nombre maximum d'utilisateurs de chaque classe à admettre dans un réseau sans fil multi classes pourrait être dérivé en offline, en résolvant le problème de maximisation du revenu à l'aide de la programmation linéaire. La dégradation dans l'utilisation des ressources, et la probabilité de chute (dropping probability) sont prises en charge en tant que contraintes du problème d'optimisation. Ensuite, un schéma de CA à base de seuil, utilise le nombre maximum d'utilisateurs par classe, qui est déterminé offline, pour admettre/rejeter les appels arrivant. Le revenu est modélisé comme une fonction linéaire de l'état du réseau (le nombre d'utilisateurs de chaque classe)[23,24,25,26].

2.4 Le Handover

La mobilité des utilisateurs finaux provoque des variations dynamiques aussi bien de la qualité du lien que du niveau des interférences, nécessitant parfois qu'un utilisateur particulier change sa station de base de service. Ce processus est appelé handover ou handoff (HO). Le Handover est un composant essentiel pour la prise en charge de la mobilité des utilisateurs finaux. Il garantit la continuité des services sans fil quand l'utilisateur mobile se déplace à travers les bordures des cellules.

Décision de déclenchement du handover

La décision de déclencher le processus de handover peut être centralisée ou décentralisée (c.a.d la décision peut être prise par la SM ou le réseau). De ce point de vue, il peut y avoir au moins trois types de prise de décision.

- **Network Controlled handoff** : Dans un protocole de handover contrôlé par le réseau, la décision de déclencher ou non ce processus est prise par le réseau, sur la base des mesurages des SMs dans un certain nombre de SBs. Ce type de handover est déclenché pour des raisons autres que le contrôle de liaison radio (pour contrôler la distribution du trafic entre les cellules, par exemple).
- **Mobile controlled handoff** : Avec ce protocole, chaque SM a le contrôle complet du processus de handover. Le SM effectue les mesurages des forces des signaux reçus des SBs avoisinantes et des niveaux des interférences sur tous les canaux. La SM choisit la meilleure SB, du moment, comme cible du handover, et exécute le processus si la liaison courante s'avère de qualité médiocre.
- **Mobile assisted handoff** : Dans ce mode, la SM effectue les mesurages et les communique au réseau, qui prend la décision de déclencher ou non le handover.

Types de handover

En général les handovers sont classés en deux catégories : Hard et Soft handovers. Souvent, le Hard handover est encore divisé en deux types différents, intra-fréquence et inter-fréquence handovers. Le handover intra-fréquence se produit entre cellules ayant la même fréquence porteuse, alors que l'inter-fréquence handover se produit entre cellules ayant des fréquences porteuses différentes. Le soft handover peut être divisé en deux types différents, multiway soft handover et softer handover.

2.4.1 Principes du Hard handover

Un hard handover est essentiellement une connexion "break before make"[1]. La SB relaie la connexion d'une SM à une autre cellule, puis se déconnecte de la SM. Dans un hard handover, le lien avec l'ancienne SB est terminé avant ou quand l'utilisateur est transféré vers la SB de la nouvelle cellule ; la SM n'est liée qu'à une seule SB, à un instant donné. Le hard handover est principalement utilisé dans les réseaux sans fil implémentant les techniques d'accès multiples FDMA(Frequency division Multiple Access) et TDMA(Time division Multiple Access), où des plages de fréquences différentes sont utilisées dans les canaux adjacents pour minimiser les interférences. Donc, quand une SM se déplace d'une SB à une autre, il devient impossible pour elle de communiquer avec les deux SBs(puisque des fréquences différentes sont utilisées). La figure 2.3 montre une SM se déplaçant de la cellule couverte par la station de base BS_1 vers celle couverte par la station de base BS_2 . La force moyenne du signal reçu de BS_1

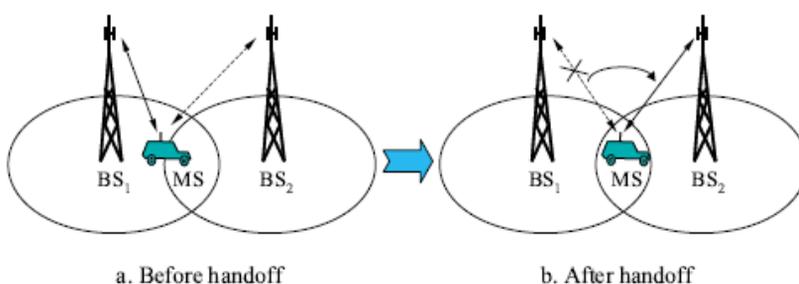


Figure 2.3 Hard Handover entre la SM et les SBs

diminue quand la SM s'éloigne d'elle. De manière similaire, la force moyenne du signal reçu de BS_2 augmente quand la SM s'en approche.

2.4.1.1 Critères de déclenchement du hard handover

L'évaluation des performances du hard handover est basée sur des critères de déclenchement variés [2]. On suppose que le signal est considéré en moyenne, pour que les fluctuations rapides dues à la propagation selon des trajets multiples puissent être éliminées. Puisque un mesurage ancien peut se révéler non fiable, on utilise en général des fenêtres de longueurs et formes déterminées, les valeurs moyenne permettant de décider de l'instant de déclenchement du hard handover. La figure 2.4 est utilisée dans les sous-sections suivantes pour décrire les approches adoptées à cette fin.

La Relative force du signal : Cette méthode choisit toujours la SB dont le signal reçu par la SM est le plus fort. Cette décision est basée sur un mesurage moyen du signal reçu. Sur la figure 2.4 le handover se déclenchera à la position A. Cette méthode est néanmoins connue par le trop grand nombre de handovers inutiles, qu'elle génère même dans les cas où le signal de la SB est encore à un niveau acceptable.

La Relative force du signal avec Seuil : Cette méthode n'autorise une SM à déclencher un handover que si, le signal dans la cellule courante est suffisamment faible (inférieur au seuil), et que celui de l'autre SB est le plus fort des deux. L'effet du seuil dépend de sa valeur relative comparée aux forces des signaux des deux SBs, au point où ses forces sont égales. Si le seuil est supérieur à cette valeur, disons T_1 dans la figure 2.4, ce schéma se comporte exactement comme les schéma précédent, donc le handover se produira au point A. Si le seuil est inférieur à cette valeur, disons T_2 dans la figure 2.4, la SM reportera le

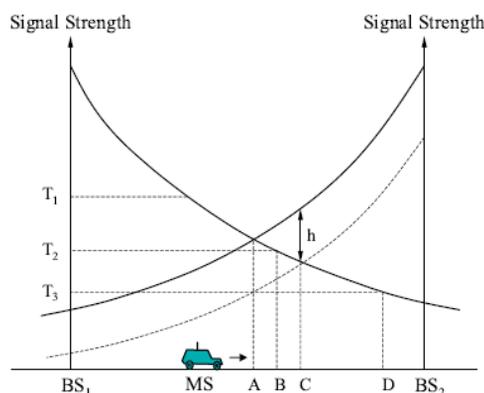


Figure 2.4 Force du signal et Hysteresis entre deux SBs adjacentes pour un handover potentiel

handover jusqu'à ce que le niveau du signal dans la cellule courante traverse le seuil à la position B. Dans le cas de T_3 , le délai pourrait être tellement long que la SM se soit engagé très loin dans la nouvelle cellule. Ceci réduit la qualité du lien de communication avec BS_1 et pourrait résulter en une coupure de la communication. De plus, cela produira des interférences additionnelles pour les utilisateurs du même canal dans les deux cellules (cochannel interference). Ainsi, ce schéma pourrait résulter en recouvrement des zones de couverture des cellules. En pratique un seuil n'est pas utilisé seul, parce que son efficacité dépend de la connaissance à priori du croisement des forces des signaux entre la SB courante et les SBs candidates.

La Relative force du signal avec Hysteresis : Avec ce schéma un utilisateur n'est autorisé à déclencher un handover que si le signal reçu de la nouvelle SB est suffisamment plus fort (avec une marge, *hysteresis margin*, h dans la figure 2.4) que le signal de la SB courante. Dans ce cas, le handover se produira au point C. Cette technique prévient l'effet ping-pong, qui résulte du handoff répété entre deux SBs suite à des fluctuations rapides des forces des signaux reçus des deux SBs. Mais le premier handover, cependant, pourrait ne pas être nécessaire si le signal de SB courante est encore suffisamment fort.

La Relative force du signal avec Hysteresis et Seuil: Ce schéma remet une SM à une nouvelle SB dans le cas où le signal dans la cellule courante chute au dessous d'un seuil donné, et que le signal de la SB

cible est plus fort que le signal courant avec une marge *hysteresis* donnée. Sur la figure 2.4, le handover se produira au point D si le seuil est T_3 .

Techniques de prédiction : Les techniques de prédiction basent la décision de déclencher un handover sur la valeur future attendue de la force du signal reçu. Il semble que ces techniques donnent de meilleurs résultats en termes de réduction du nombre de handovers non nécessaires.

2.4.2 Principes du soft handover

Le soft handover est différent du traditionnel hard handover. Avec la hard handover, la décision de l'exécuter ou non est claire, et le mobile ne communique qu'avec une seule SB à la fois. Avec le soft handover, cette décision est conditionnelle. Selon les changements dans la force du signal pilote reçu des deux (ou plus de deux) SBs impliquées, une décision hard sera prise de ne communiquer qu'avec une seule des ces SBs. Ceci se produit normalement après qu'il devient clair que le signal provenant d'une SB est considérablement plus fort que ceux reçus des autres. Dans la période intermédiaire, la SM communique simultanément avec toutes les SBs de l'active set (ensemble des cellules avec lesquelles le mobile communique présentement). Le hard handover se produit en un certain point du temps, alors que le soft handover dure un certain temps. La figure 2.5 montre le processus de base du hard et soft handover (cas 2-way). Supposons qu'il y a un terminal mobile à l'intérieur de la voiture qui se déplace de la cellule cell1 vers la cellule cell2, BS_1 est la SB servante originale du mobile. En se déplaçant, le mobile mesure continuellement la force du signal pilote reçu des SBs situées aux alentours. Avec le hard handover comme indiqué sur la figure 2.5 (a), le déclenchement du handover peut être simplement décrit comme :

```

If  $(pilot\_E_c/I_0)_2 - (pilot\_E_c/I_0)_1 > D$  and  $BS_1$  is the serving BS
  Handover to  $BS_2$ ;
else
  do not handover ;
end
  
```

Où $pilot_E_c/I_0)_1$ et $pilot_E_c/I_0)_2$ sont respectivement le pilote E_c/I_0 reçu de BS_1 et BS_2 ; D est la marge hysteresis (dans le ratio E_c/I_0 E_c est l'énergie par bit, et I_0 est densité spectrale de la puissance du bruit). La raison de l'introduction de la hysteresis margin dans le hard handover est d'éviter l'effet ping-pong, qui renvoie aux fréquents hard handovers qui se produisent lorsque la SM se déplace à l'intérieur et à l'extérieur de la limite de la cellule. La mobilité mise à part, l'effet ping-pong est aussi rendu plus sérieux par l'affaiblissement du canal radio. En introduisant hysteresis margin, cet effet est adouci car le mobile n'exécute pas immédiatement la handover vers la meilleure SB. Plus grande est la marge, moins important est l'effet ping-pong. Cependant, une grande marge, signifie plus de délai. Plus encore, le

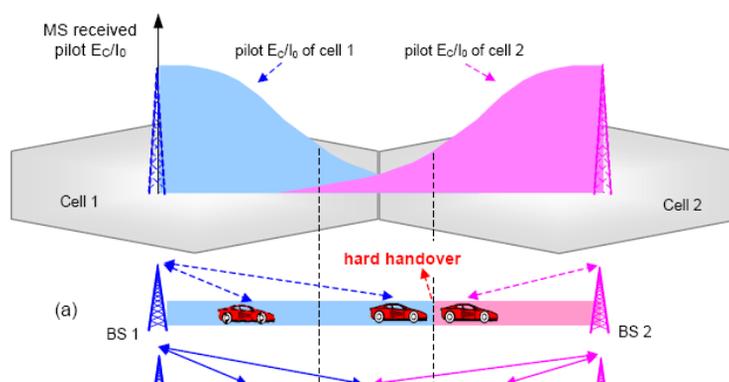


Figure 2.5 Comparaison entre le hard et le soft handover

mobile cause des interférences additionnelles aux cellules avoisinantes à cause de la médiocrité de la liaison pendant ce délai. Ainsi, pour déclencher un hard handover, la valeur de la hysteresis margin est vraiment importante.

Dans le cas du soft handover, montré en (b) de la figure 2.5, avant que $(pilot_E_c/I_0)_2$ n'aille au-delà de $(pilot_E_c/I_0)_1$, et tant que la condition de déclenchement du soft handover est satisfaite, le mobile se met dans l'état de soft handover, et une nouvelle liaison est établie. Avant que BS_1 ne soit coupée (condition de coupure de handover satisfaite), le mobile communique avec BS_1 et BS_2 simultanément. Ainsi, et contrairement au hard handover, le soft handover est un processus « make before break ». Jusqu'à présent, plusieurs algorithmes ont été proposés pour le support du soft handover, utilisant différents critères. Le soft handover n'est pas le même dans les différentes directions de transmission (comme l'illustre la figure 2.6). Dans l'uplink, le mobile transmet les signaux dans l'air à travers ses antennes omnidirectionnelles. Les deux SB dans l'active set reçoivent les signaux simultanément. Donc, dans l'uplink il n'est pas besoin d'un extra canal pour le support du soft handover.

Dans le downlink, les mêmes signaux sont transmis à travers les deux SB et le mobile peut combiner de manière cohérente les signaux reçus de différentes SB puisqu'il les voit juste comme composants multi trajets additionnels. Cependant, pour supporter le soft handover dans le downlink au moins un canal additionnel (2-way SHO) est requis dans le downlink. Ce canal agit pour les autres utilisateurs comme une interférence additionnelle dans l'interface air. Donc, pour supporter le SHO dans le downlink, plus de

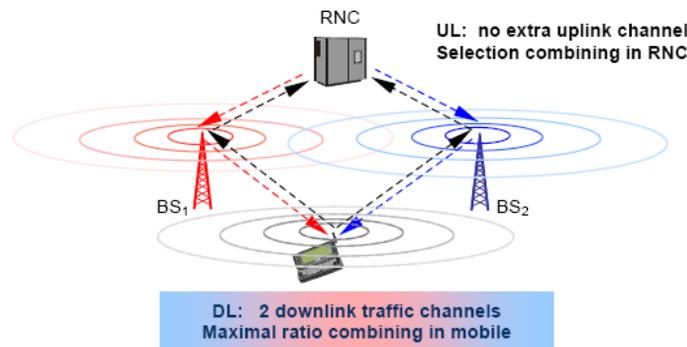


Figure 2.6 Principes du soft handover

);

ressources sont nécessaires. Par conséquent, dans la direction downlink, les performances du SHO dépendant du compromis entre le gain de macro diversité, et la consommation de ressources supplémentaires.

2.4.2.1 Algorithme du soft handover

La performance du soft handover est étroitement liée à l'algorithme utilisé. La figure 2.7 montre l'algorithme du soft handover pour IS-95A (appelé aussi cdmaOne) [2]. L' "active set" est la liste des cellules qui sont actuellement connectées au mobile; le "candidate set" est la liste des cellules qui ne sont pas présentement utilisées dans la connexion soft handover, mais dont les valeurs du pilot E_c/I_0 sont assez fortes pour être ajoutées dans l' "active set"; le "neighbouring set" (monitored set) est la liste des cellules que le mobile mesure continuellement, mais dont les valeurs du pilot E_c/I_0 ne sont pas assez fortes pour être ajoutées à l' "active set". Dans IS-95A, le seuil du handover est une valeur fixée du pilot E_c/I_0 . Il est facile à implémenter, mais a des difficultés pour traiter les changements de charge dynamiques. Basés sur

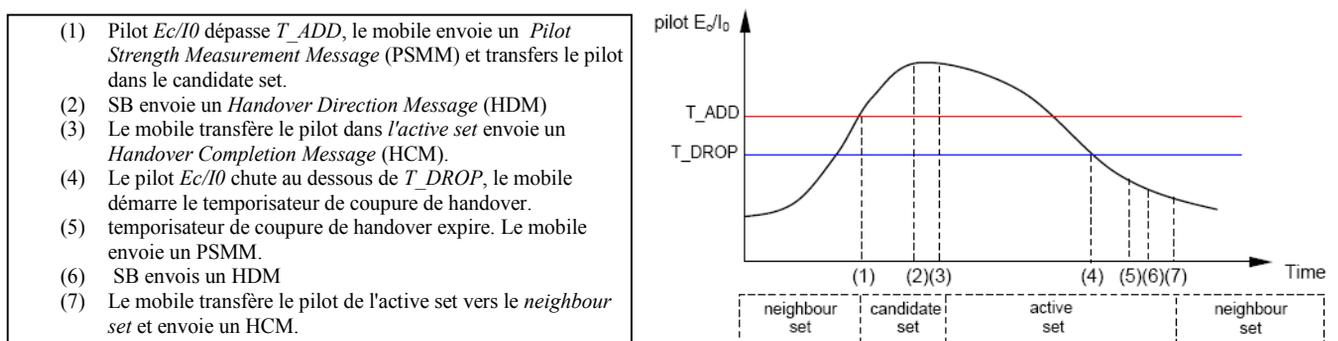


Figure 2.7 Algorithme du soft handover

l'algorithme du IS-95A, plusieurs algorithmes cdmaOne modifiés furent proposés pour les systèmes IS-95B et cdma2000 avec des seuils dynamiques.

Dans le WCDMA, des algorithmes plus compliqués sont utilisés, comme illustré dans la figure 2.8 [2]. L'algorithme de soft handover WCDMA peut être décrit comme suit:

- Si $pilot_E_c/I_0 > Best_pilot_E_c/I_0 - (AS_Th - AS_Th_Hyst)$ pour une période ΔT et l'active set n'est pas plein, la cellule est ajoutée à l'active set. Ceci est appelé Event 1A ou Radio Link Addition.
- Si $pilot_E_c/I_0 < Best_pilot_E_c/I_0 - (AS_Th + AS_Th_Hyst)$ pour une période ΔT , alors la cellule est retirée de l'active set. Ceci est appelé Event 1B ou Radio Link Removal.
- Si l'active set est plein et $Best_candidate_pilot_E_c/I_0 > Worst_Old_pilot_E_c/I_0 + AS_Rep_Hyst$ pour une période ΔT , alors la cellule la plus faible dans l'active set est remplacée par la cellule candidate la plus forte. Ceci est appelé Event 1C ou Combined Radio Link Addition and Removal. La taille maximale de l'active set est supposée être 2, dans la figure 2.8.

Où $pilot_E_c/I_0$ est la quantité mesurée et filtrée de E_c/I_0 du CPICH; $Best_pilot_E_c/I_0$ est la plus forte cellule mesurée dans l'active set; $Best_candidate_pilot_E_c/I_0$ est la plus forte cellule mesurée dans le *monitored set*; $Worst_Old_pilot_E_c/I_0$ est la plus faible cellule mesurée dans l'active set.

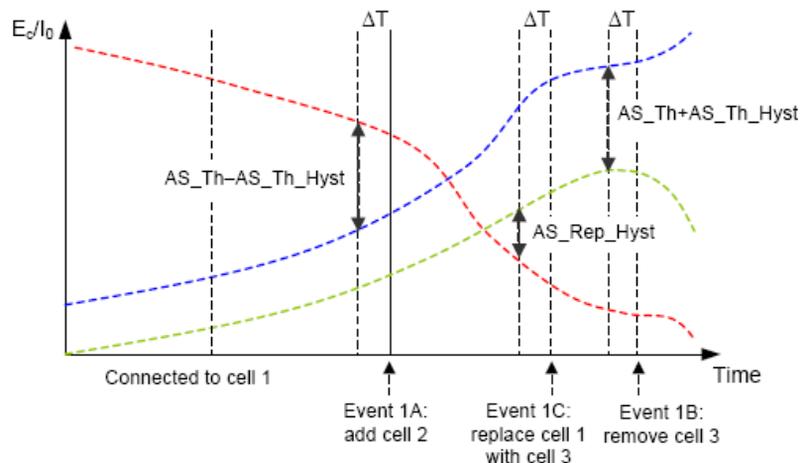


Figure 2.8 Algorithme de soft handover pour le WCDMA

Dans l'algorithme du WCDMA, des seuils relatifs plutôt que des seuils absolus sont utilisés. Comparé à IS-95A, le plus grand bénéfice de cet algorithme est son paramétrage facile, sans aucun réglage requis pour les zones de faibles interférences, à cause des seuils relatifs.

2.5 Contrôle de puissance

Le contrôle de puissance est un élément nécessaire dans les systèmes mobiles à cause du problème de longévité des batteries et pour des raisons de sécurité, mais dans les systèmes CMDA, par exemple, le contrôle de puissance est essentiel parce que ces systèmes sont limités par les interférences (interference limited). Dans le GSM, un contrôle de puissance lent (fréquence approximative 2 Hz) est employé. Dans IS-95 un contrôle de puissance rapide avec 800 Hz est supporté dans le lien uplink, mais dans le lien downlink, une boucle de contrôle de puissance relativement lente (à peu près 50 Hz) contrôle la puissance de transmission. Dans le WCDMA un contrôle de puissance rapide avec une fréquence de 1.5 kHz est supporté dans les deux sens uplink et downlink[OP98]. Le contrôle de puissance rapide et étroit est le plus important aspect des systèmes WCDMA.

Les raisons de l'utilisation du contrôle de puissance sont différentes dans l'uplink et le downlink. L'ensemble des objectifs du contrôle de puissance peut être résumé comme suit :

- Venir à bout de l'effet near-far dans le sens uplink,
- Optimiser la capacité du système en contrôlant les interférences,

- Maximiser la durée de vie des batteries des terminaux mobiles.

La figure 2.9 illustre le problème near-far [SL-SWJ99] dans le sens uplink. Les signaux des différentes SMs, dans les systèmes WCDMA, sont transmis simultanément dans la même bande de fréquences. Sans un contrôle de puissance, le signal provenant de la SM qui est plus proche de la SB pourrait bloquer les signaux provenant des autres SMs qui sont beaucoup plus éloignées de la SB. Dans la pire situation, une SM émettant une grande puissance peut bloquer toute la cellule. La solution est d'appliquer un contrôle de puissance pour garantir que les signaux provenant des différents terminaux ont la même puissance ou le même *SIR* (Signal-to-Interference Ratio) lorsqu'ils arrivent à la SB.

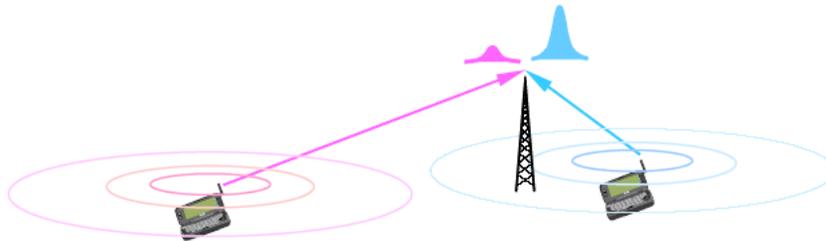


Figure 2.9 Effets Near-Far (control de puissance dans le UL)

Dans la direction downlink, il n'y a pas de problème near-far à cause du scénario one-to-many. Le contrôle de puissance est responsable de la compensation des interférences inter cellules, subies par les mobiles, spécialement ceux qui sont proches des limites de la cellule, comme montré sur la figure 2.10. Plus encore, le contrôle de puissance dans le downlink est responsable de la minimisation des interférences totales en maintenant la QoS à sa valeur cible.

Dans la figure 2.10, le mobile 2 subit plus d'interférence inter cellules que le mobile 1. Par conséquent, pour atteindre la même qualité cible, plus de puissance doit être allouée au canal downlink entre la SB et le mobile 2.

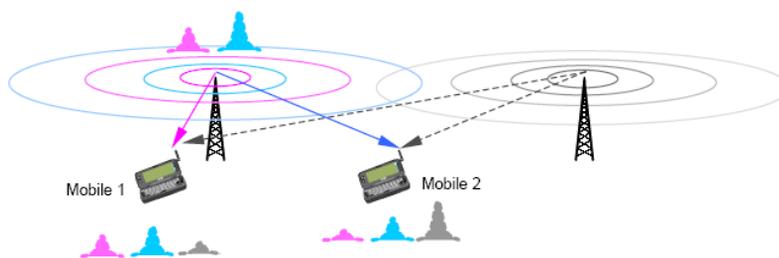


Figure 2.10 Compensation des interférences inter-cellules (control de puissance dans le DL)

Il y a trois types de contrôle de puissance: contrôle de puissance en boucle ouverte, contrôle de puissance en boucle fermée, et contrôle de puissance en boucle extérieure.

2.5.1 Contrôle de puissance en boucle ouverte

Le contrôle de puissance en boucle ouverte est utilisé dans mode FDD de l'UMTS pour l'initialisation de la puissance du mobile. Le mobile estime la perte de chemin (path loss) entre la SB et le mobile, en mesurant la force du signal reçu en utilisant un circuit spécial appelé AGC (Automatic Gain Control). Selon cette estimation de la perte de chemin, la SM peut décider de la puissance de transmission dans la direction uplink. Le contrôle de puissance en boucle ouverte est efficace dans les systèmes TDD parce que les directions uplink et downlink sont réciproques, mais il n'est pas très efficace avec les systèmes FDD car les canaux uplink et downlink opèrent sur des bandes de fréquences différentes, et car les affaiblissement de Rayleigh dans l'uplink et le downlink sont indépendants. Donc, ce type de contrôle de

puissance peut approximativement, seulement, compenser l'atténuation due à la distance. C'est pourquoi il est utilisé seulement comme un réglage initial de la puissance dans les systèmes FDD.

2.5.2 Contrôle de puissance en boucle fermée

Appelé aussi contrôle de puissance rapide (fast power control) dans les systèmes WCDMA, ce type de contrôle de la puissance est responsable du contrôle de la puissance transmise de la SM (uplink) ou de la SB (downlink) pour contre carrer l'affaiblissement du canal radio et atteindre le SIR cible, fixé par la boucle extérieure. Par exemple, dans l'uplink, la SB compare le SIR reçu de la SM avec le SIR cible une fois tout les time slots (0.666ms). Si le SIR reçu est supérieur au SIR cible, la SB envoie une commande TPC "0" à la SM via le canal downlink DCCH (Dedicated Control Channel). Si le SIR reçu est inférieur au SIR cible, la SB émet une commande TPC "1" à la SM. Parce que la fréquence du contrôle de puissance en boucle fermée est très rapide il peut compenser aussi bien l'affaiblissement rapide que l'affaiblissement lent.

2.5.3 Contrôle de puissance en boucle extérieure

On a besoin de ce type de contrôle de puissance pour maintenir la qualité de la communication au niveau requis, en fixant l'objectif pour le contrôle en boucle fermée. Il vise à fournir la qualité requise: ni mieux, ni pire. La fréquence de ce contrôle est typiquement de 10-100 Hz. La figure 2.11 montre l'algorithme général de ce type de contrôle. Le contrôle de puissance en boucle extérieure compare la qualité reçue à la

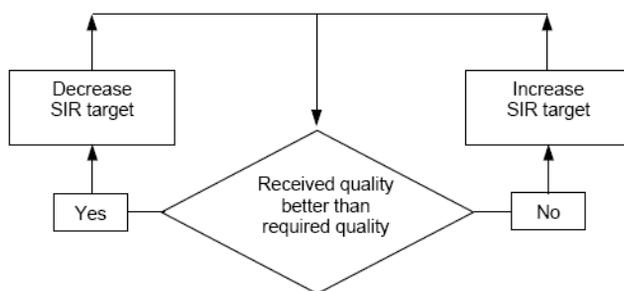


Figure 2.11 Algorithme générale de control de puissance de la boucle extérieure.

qualité requise. Souvent, la qualité est définie comme un certain BER(Bit Error Rate) cible ou une certain FER (Frame Error Rate). La relation entre le SIR ciblé et la qualité cible dépend de la vitesse du mobile et du profile multi trajets. Si la qualité reçue est meilleure, cela signifie que le SIR courant est assez élevé pour garantir la QoS requise. Dans le but de minimiser ce surplus de qualité, le SIR ciblé va être réduit. Cependant, si la qualité reçue est pire que la qualité requise, le SIR ciblé aura besoin d'être augmenté pour garantir la QoS exigée.

2.6 Ordonnancement (Scheduling)

Les algorithmes d'ordonnancement fournissent des mécanismes d'allocation de bande passante et de multiplexage au niveau paquet. Ils permettent donc, le support du mode de commutation de paquet sur l'interface radio. Ceci implique le partage de l'interface air entre les utilisateurs de paquets de données, la prise de décision concernant le canal de transport de chaque utilisateur et la surveillance de la charge du système, et l'allocation des paquets. Les paquets peuvent être ordonnancés selon le délai, jitter, le débit ou encore l'équité pour la QoS de bout en bout. L'opération d'ordonnancement des paquets est étroitement liée aux autres fonctions de la GRR telles que le contrôle de charge et l'admission d'accès.

Les caractéristiques des réseaux sans fil posent des problèmes spéciaux qui n'existent pas dans les réseaux filaires. Celles-ci incluent :

- 1) des erreurs sous forme de rafales et à taux élevé;
- 2) une capacité de la liaison sans file variable dans le temps est dépendante de la position;
- 3) bande passante rare;
- 4) mobilité de utilisateurs;
- 5) contrainte de puissance des SMs.

Toutes ces caractéristiques constituent autant de défis à relever pour le développement d'algorithmes d'ordonnancement (pour des réseaux sans fil) efficaces.

2.6.1 Aspects majeurs de l'ordonnancement sans fil

L'ordonnancement dans les réseaux sans fil présente quelques aspects propres, qui le distinguent des réseaux filaires. Ces aspects ont trait à ce qui suit :

Variabilité de la liaison sans fil

La plus grande différence entre un réseau sans fil et un réseau filaire est la nature variable de la liaison de transmission du premier réseau. Les canaux sans fil sont, en effet, plus sujets aux erreurs et souffrent des interférences, affaiblissements, et du shadowing. En conséquence, la capacité d'une liaison sans fil est, par nature très variable, et peut même atteindre des états dans lesquels la liaison sans fil est tellement mauvaise qu'aucune transmission de paquet ne peut être réussie. L'occurrence de ces états est intempestive et se produit par rafales d'où leur appellation de "*bursty error state*". En plus des ces fluctuations dans le temps, la capacité d'une liaison sans fil est aussi dépendante de la localisation des stations utilisateurs(SUs). En raison des localisations physiques différentes, certaines SUs pourraient jouir de communications sans erreurs, alors que d'autres ne peuvent pas communiquer du tout. Ce cas est appelé "*location-dependant error*". Plus encore, la mobilité des SUs augmente la variabilité des liaisons de transmission. De telles variations exigent que les algorithmes d'ordonnancement soient dotés de mécanismes dynamiques qui prennent en charge ses changements (aussi bien ceux liés au temps que ceux liés à la localisation).

L'équité

L'équité dans les réseaux sans fil est un problème plus compliqué. Il peut arriver qu'un paquet soit ordonnancé pour une transmission sans fil selon une certaine discipline de service ou directive d'équité, qui est indépendante de l'état de la liaison, et que celle-ci soit en fait dans un état d'erreur. Si le paquet est transmis, il sera altéré et la transmission gaspillera des ressources de transmission. Dans ce cas, différer la transmission de ce paquet jusqu'à ce que la liaison se rétablisse de cet état d'erreur est un choix raisonnable. Le flot affecté perdra, donc, momentanément son partage de la bande passante. Pour assurer une équité, le flot devrait être dédommagé de cette perte plus tard lorsque l'état de la liaison le permettra. Mais déterminer comment le dédommagement se fera n'est pas une tâche aisée. La définition et les objectifs de l'équité deviennent plus ambigus dans un environnement sans fil. La granularité de l'équité, c.a.d, équité à court terme versus équité à long terme, est un autre facteur qui affecte la politique d'ordonnancement. Une interprétation appropriée de l'équité pour un ordonnancement sans fil devrait dépendre du modèle de service, du type de trafic, et des caractéristiques du canal

QoS

Les réseaux sans fil large bande vont offrir des services pour des classes de trafic hétérogènes avec des exigences de QoS différentes. Par conséquent, la différenciation en termes de QoS, et la garantie de cette dernière doivent être supportées. Le support de la QoS dans l'ordonnancement sans fil est dicté par le modèle de service. Pour les services de type DiffServ, au moins un service d'ordonnancement à priorités pour un trafic agrégé et une différenciation par la QoS devrait être implémenté dans l'algorithme d'ordonnancement. Pour les services de type IntServ, l'algorithme d'ordonnancement devrait avoir des performances de QoS (délai, limite de jitter) garanties au niveau flot. Il reste bien entendu que si une liaison sans fil souffre de dégradations de canal fréquentes, il est très difficile de garantir la QoS pour les flots utilisant cette liaison. Néanmoins, la QoS devrait être garantie pour les flots, soit de manière déterministe ou statistiquement, sur les liaisons où la dégradation du canal physique n'excède pas un certains seuils.

Débit de données et utilisation du canal

La ressource la plus précieuse dans un réseau sans fil est la bande passante. Un algorithme d'ordonnancement sans fil devrait donc viser la minimisation des transmissions improductives sur les liaisons à erreurs, et au même temps, maximiser le service effectif délivré et l'utilisation des canaux sans fil.

La contrainte de puissance et la simplicité

Dans les réseaux cellulaires l'algorithme d'ordonnancement est souvent exécuté dans la SB. La puissance électrique exigée pour le calcul de l'ordre du service paquet ne devrait pas constituer une grande préoccupation, en raison de la disponibilité de puissance adéquate dans la SB. Cependant, les SUs ont cette contrainte de la puissance. Un bon algorithme d'ordonnancement devrait être conçu de manière à ce qu'un nombre minimal de messages de contrôle liés à l'ordonnancement (qui pourraient contenir des informations sur la file d'attente des SUs, les temps d'arrivés de paquets, et les états des canaux) sont exigés de la SM. De même, l'algorithme d'ordonnancement ne devrait pas être trop complexe, pour qu'il puisse être exécuté à une grande vitesse, pour ordonnancer le trafic multimédia temps réel à forte contraintes temporelles.

2.6.2 Quelques approches pour l'ordonnancement sans fil

CSDPS (Channel State Dependent Packet Scheduling)

L'un des premiers papiers à avoir traité le problème des erreurs liées à la localisation et des erreurs de nature intempesive et courte, de l'ordonnancement dans les réseaux sans fil est [4]. Il propose un algorithme d'ordonnancement appelé CSDPS (Channel State Dependent Packet Scheduling). La figure 2.12 présente les composants majeurs du système d'ordonnancement CSDPS de la SB. Une file d'attente séparée est maintenue pour les paquets d'un mobile. Dans une file d'attente, les paquets sont servis dans l'ordre FIFO. A travers les files d'attentes la politique de service pourrait être décidée suivant les

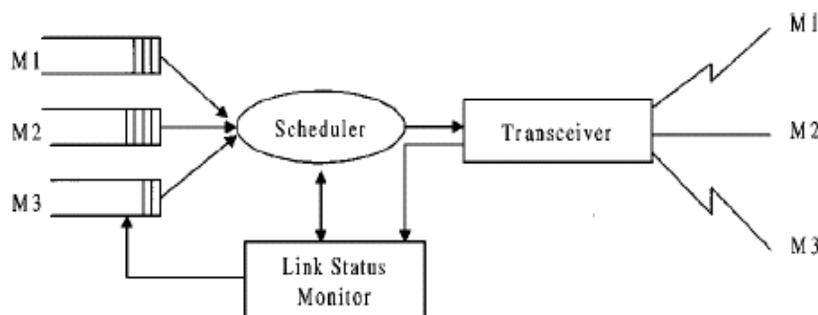


Figure 2.12 Schéma d'un ordonnanceur CSDPS

exigences du service (round robin RR, earliest timestamp first ETF, longest queue first LQF). L'idée de base du CSDPS est très simple. Lorsqu'une liaison sans fil souffre des erreurs "*bursty errors*", l'algorithme d'ordonnancement diffère la transmission des paquets sur cette liaison. Les LSM (Link Status monitor) surveillent les états des liaisons pour toutes les SMs. Lorsque le LSM détecte une liaison dans l'état "*bad*", il marque la file d'attente affectée. L'ordonnanceur ne servira pas les files d'attente marquées. La file est démarquée après une période de temps (time-out period), qui peut être, par exemple, la durée d'erreur moyenne de la liaison. Une liaison est déclarée dans l'état "*bad*" lorsque l'accusé de réception du paquet de données n'est pas reçu du mobile. Le CSDPS améliore les performances de l'ordonnancement en prenant en compte les états du canal qui sont dépendants de la localisation et du temps (location-dependant and time-dependant). Cet algorithme a aussi des insuffisances. Il n'a aucun mécanisme pour garantir de la bande passante pour un utilisateur mobile. Un utilisateur mobile pourrait recevoir beaucoup moins d'opportunités de service que ne le permet un partage équitable lorsqu'une liaison est "*supposée*" dans un état "*bad*", les opportunités de service de l'utilisateur mobile affecté seront dans temps données à d'autres utilisateurs, bien que ces autres utilisateurs pourraient avoir déjà dépassé ce que leurs permet un partage de service équitable. Aucune limite n'est imposée sur la quantité de service reçu par utilisateur mobile. Plus encore, l'algorithme ne donne aucune garantie sur le délai du paquet.

CSDPS + CBQ (Class-Based Queuing)

Pour résoudre le problème de la non équité du partage de la bande passante dans CSDPS, un schéma d'ordonnancement combinant CBQ et CSDPS a été proposé dans [5]. Dans ce schéma, les utilisateurs ou les flots de trafic sont groupés dans des classes, et chaque classe est assurée d'une certaine quantité de bande passante. Le composant CSDPS est utilisé pour traiter les variations de la liaison sans fil, et le composant CBQ est utilisé pour fournir un mécanisme pour l'équité dans le partage de tout la canal sans fil.

CBQ est un ensemble de directives hiérarchiques de partage de canal pour assurer que les classes reçoivent leurs parts du partage de bande passante sur une période prédéfinie de temps. Il vise aussi la distribution, de manière équitable, de l'excès de bande passante. CBQ garde trace de la quantité de service reçu par chaque classe dans une certaine fenêtre d'intervalle de temps. Une classe est appelée *unsatisfied* si elle a des arriérées persistantes, et que le service qu'elle a reçue récemment est inférieur à la fraction allouée. Les directives de partage de canal du CBQ sont qu'une classe devrait être empêchée de recevoir un servir lorsqu'elle excède son partage de bande passante allouée et qu'elle contribue à l'état *unsatisfied* de n'importe quelle autre classe. Une telle file d'attente est appelée *restricted queue*.

IWFQ (Idealized Wireless Fair Queueing)

Un modèle IWFQ est défini par référence à un système de service WFQ (weighted fair queueing) sans erreurs. Etant donné la séquence d'arrivée des paquets de chaque flot de données, un service sans erreurs est défini comme le service WFQ pour les flots avec les séquences d'arrivées identiques et les liaisons complètement sans erreurs. Le service reçu par un flot dans un système IWFQ (avec des erreurs de liaison) est comparé avec un système sans erreurs. Un flot est dit être *leading*, *lagging*, ou *in sync* à un instant donné si la taille de sa file d'attente est inférieure, supérieure, ou égale à la taille de la file d'attente du système sans erreurs. En supposant une parfaite connaissance des liaisons sans fil et un contrôle d'accès multiple(MAC) parfait, IWFQ fonctionne comme suit. Chaque flot a sa propre file d'attente. Lorsqu' aucune liaison ne souffre des erreurs, le fonctionnement est comme celui d'un WFQ ordinaire dans un réseau filaire. Lorsqu'un paquet de numéro de séquence n d'un flot i arrive, il est étiqueté avec le temps de départ (*start time*) $S_{i,n}$ et le temps de terminaison (*finish time*) $f_{i,n}$ du service virtuel. Plus spécifiquement

$$S_{i,n} = \max\{v(A(t)), f_{i,n-1}\};$$
$$f_{i,n} = S_{i,n} + L_{i,n} / r_i$$

où

- $L_{i,n}$ taille de paquet du paquet arrivé;
- $v(A(t))$ temps virtuel du système;
- r_i la vitesse de service allouée au flot i .

Les paquets sont rangés dans l'ordre non décroissant des *finish times* dans chaque file d'attente. L'ordonnanceur choisit toujours de servir le paquet avec le plus petit *finish time*. Toutes ces opérations sont exactement ce qu'un ordonnanceur WFQ filaire exécute. La différence se manifeste lorsqu'une erreur de liaison se produit. Si le paquet choisi ne peut pas être transmis à cause d'un état de liaison *bad*, le paquet, dans les autres files d'attente, avec le plus petit *finish time* suivant est pris. Le processus continuera et recommencera à partir du début, si nécessaire, jusqu'à ce que l'ordonnanceur trouve un paquet avec un état de liaison *good*. Puisque les étiquettes de temps d'un paquet ne sont normalement pas changés après l'arrivée du paquet, un flot qui perd son opportunité de service à cause d'une erreur de liaison aura des paquets avec des *finish times* plus petits, comparés aux paquets des autres flots. Par conséquent, il aura la priorité dans l'accès à la bande passante du service lorsque sa liaison sort de l'état d'erreur. Donc la compensation est garantie.

CIF-Q (Channel-condition Independent packet fair Queueing)

CIF est très similaire à IWFQ en ce sens qu'il utilise aussi un système de files d'attente équitable est sans erreurs pour approximer le service réel au système idéal sans erreur. L'une des contributions importantes du CIF-Q est peut être qu'il a clairement spécifier les propriétés qu'un ordonnanceur sans fil équitable devrait avoir. L'ensemble des propriétés inclus quatre objectifs.

- Garanties de délai et de débit: une limite de délai et un débit pour les flots sans erreurs devraient être garantis, et non affectés par les autres flots dans l'état *error*.
- Equité de long terme : après qu'un flot sorte d'une erreur de liaison, et tant qu'il a assez de demande de service, il devra être compensé, sur une période suffisamment longue, pour tout son service perdu alors qu'il été dans l'état *error*.
- Equité à court terme : la différence entre les services normalisés de n'importe quelle pair de flots sans erreurs qui sont continuellement surchargés et qui sont dans le même état (*leading*, *lagging*, ou *satisfied*) durant un intervalle de temps devrait être bornée.
- Dégradation gracieuse des flots dans l'état *leading* : pendant n'importe quel intervalle de temps, lorsqu'il est sans erreur, un flot *leading* chargé devrait être assuré de recevoir au moins une fraction du service qu'il aurait reçu dans un système sans erreurs.

Chapitre 3

Etat de l'Art de la Prise de Décision De Gestion des Ressources Radio Dans les Réseaux Sans Fil Hétérogènes

3.1 Introduction

Le concept de réseau hétérogène est destiné à proposer une architecture flexible et ouverte pour une large variété de technologies d'accès radio (*TAR*), applications, et services, avec des demandes de QoS différentes, ainsi que des piles de protocoles différentes [1]. Les technologies d'accès sans fil (ou radio) incluent les réseaux cellulaires, et aussi d'autres réseaux publics non cellulaires (*WLAN*). Dans ces scénarios, plusieurs *TARs* vont coexister et fonctionner de manière coordonnée. Les réseaux sans fil diffèrent les uns des autres par la technologie de l'interface air, la taille de la cellule, les services, les prix, l'accès, la couverture, et le propriétaire. Les caractéristiques, complémentaires, des différentes *TARs* rendent possible l'exploitation du "*trunking gain*" menant à une plus grande performance globale, que les performances agrégées des réseaux individuels. Il est clair que le gain potentiel de ces systèmes hétérogènes ne deviendra réalité qu'à travers une gestion convenable des ressources radio disponibles dans tout le système.

La gestion des ressources radio (*GRR*) dans un réseau sans fil et de mobiles hétérogène (*RSMH*) soulève des défis spécifiques par rapport à la *GRR* dans les à technologie d'accès unique. Il s'agit, en effet, d'exploiter l'hétérogénéité à l'avantage des chacun des réseaux sans fil constituant, tout en réduisant les coûts inhérents à leur composition. Plus spécifiquement, et en termes de ressources radio, on devrait être en mesure de gérer comme un tout l'ensembles de pools de ressources radio propres à chacune des *TAR* constituant le *RSMH*. Cela n'implique aucunement que des changements profonds des algorithmes de *GRR* spécifiques à chaque *TAR* doivent être opérés, puisque de telles *TARs* ont été conçues avec beaucoup de soins en tenant compte des spécificités liées au spectre alloué, à la méthodes d'accès, ... etc; de même qu'elles ont aussi fait leurs preuves sur le terrain. Ainsi, une station utilisateur (*SE*), une fois admise dans le réseau et sa *TAR* d'accès choisie, sera prise en charge par les algorithmes *GRR* spécifiques à cette même *TAR*, tant que cela n'empiète pas sur les attributions de la gestion commune des ressources radio (Common Radio Resources Managment) du *RSMH*. Il faut cependant préciser que le fonctionnement global est régi par trois principes majeurs, indépendamment du caractère centralisé, décentralisé, ou hybride, de l'exercice de cette fonction, et du couplage entre les différentes *RATs*. Ces principes sont:

- L'entité (ou entités) responsable la *GRR* au niveau global devrait être tenue informée, aussi souvent que nécessaire et avec la granularité appropriée, de l'état (en termes de ressources radio notamment) de chaque *TAR* et du contexte en général, pour une prise décision judicieuse.
- C'est à cette même entité (ou entités) que revient la prise de décision à l'occurrence de certains évènements particuliers, dépendant de la stratégie globale de gestion suivie et comprenant le Contrôle d'Admission (*CA*), le HandOver Vertical (*HOV*), et le choix de la *TAR* initiale. Les deux dernières fonctionnalités caractérisent les *RSMHs*.
- Les décisions prises par l'entité (ou entités) de *GRR* au niveau global devraient être respectées au niveau de chacune des *RATs* constituant le *RSMH*.

La revue des publications traitant de la *GRR* dans les réseaux sans fil hétérogènes fait ressortir, d'une part, la diversité des approches proposées (théorie des jeux, logique floue, chaînes de Markov, algorithmes génétiques, méthodes d'optimisation, modèle économique,...etc) et, d'autre part, les aspects importants considérés (contrôle d'admission, choix de technologie d'accès initiale, handover inter-technologies, partage de charge, pour ne citer que les plus importants). En général, la gestion des ressources radio dans les réseaux mobiles hétérogènes, implique trois activités principales[2]: la collecte des informations

(information harvesting), la prise de décision (decision making), et l'exécution des décisions (decision enforcement). Le rôle de chacune de ces activités est sommairement décrit dans ce qui suit:

- **La collecte d'information:** cette activité est relative à la collecte des informations concernant aussi bien les utilisateurs que le réseau. En effet, ces informations sont importantes pour la prise de décision, et servent, selon certaines approches, à définir le contexte des différentes opérations du réseau. D'abord l'information concernant chaque utilisateur est collectée au niveau des terminaux de ces derniers, ensuite elle est propagée le long des points d'accès et des routeurs d'accès pour avoir plus d'informations sur les conditions régissant au niveau cellule et réseau respectivement.
- **La prise de décision:** c'est au cours de cette activité qu'est mis en œuvre le mécanisme de prise de décision particulier à une approche donnée. Chaque mécanisme de prise de décision, dans la gestion des ressources, a besoin d'informations sur des facteurs d'entrée spécifiques. Ces entrées sont essentielles pour une bonne prise de décision; selon leurs natures, ils peuvent être classés en deux catégories: prédéterminés (probabilité de coupure de connexion, minimum de bande passante, de perte, de latence, ...etc) et variants dans le temps (charge du réseau, puissance de transmission, SNR, BER, ...etc). Les facteurs prédéterminés restent pour une longue période de temps et sont pris en considération comme valeurs pour une politique initiale ou comme préférences; ils incluent aussi les contraintes d'applications et les aptitudes de la technologie et des équipements. D'un autre côté, les facteurs variants dans le temps sont continuellement surveillés et reportés.
- **L'exécution des décisions:** cette activité veille au respect de l'application des décisions prises. Plusieurs mécanismes peuvent être utilisés pour assurer l'exécution des décisions prises. Le contrôle d'admission est l'un de ces mécanismes.

Une fois en possession des informations sur les facteurs de prise de décision, des schémas de gestion déploient des techniques de prise de décision variées, qui donnent les meilleures solutions possibles pour les fournisseurs de services. En général, elles permettent de contrôler l'accès au réseau et déterminer la meilleure technologie d'accès et le point d'attachement aussi bien pour les nouvelles connexions que pour les connexions en cours. Le partage de charge et la gestion de la mobilité sont aussi parmi les fonctions importantes prises en charges par les techniques de GRR.

3.2 La Prise de Décision

La GRR dans les réseaux sans fil hétérogènes a fait l'objet d'une pléthore de travaux de recherche, dans lesquels divers mécanismes de prise de décision, en la matière, sont explorés. Dans le reste de ce chapitre nous présentons un certain nombre de ces mécanismes selon les groupes suivants:

3.2.1 Mécanismes basés sur une fonction

Dans cette catégorie, les mécanismes de décisions sont simplement basés sur le résultat de fonctions. Nous avons choisi de présenter quatre fonctions représentatives de cette catégorie de mécanismes.

Fonction Objectif – Dans les fonctions objectives [3], les entrées sont dérivées de trois sources différentes: les données utilisateurs, les données réseau, et les informations de politique. D'abord, les utilisateurs (ou stations utilisateurs) sont interrogés sur la liste de points d'accès (Access Points) visibles, avec la qualité de signal correspondante, la liste des services demandés avec le débit nominal, et le délai toléré correspondant. Ensuite, les données réseau telles que, la bande passante du AP au niveau interface air et le délai dans la file d'attente entre le routeur d'accès et le backbone, sont collectées. Puis, la politique telle que le coût, compatibilité, confiance, préférences et aptitudes ainsi que leurs poids sont définis. Les poids peuvent être dynamiquement changés selon les conditions du réseau. Enfin, avec tout les facteurs et leurs poids respectifs, l'algorithme calcule l'allocation qui maximise la fonction objective pour tout le réseau. Les auteurs définissent pour l'algorithme d'accès et de sélection de l'interface une fonction OF définie comme suit: pour un service s appartenant au total des services S , et un point d'accès ap appartenant au total des points d'accès AP , la fonction objective OF est donnée par (3.1).

$$OF(\forall s \in S, \forall ap \in AP) = F(s, ap) + OF(\forall s' \in S, s' \neq s, \forall ap \in AP) \quad (3.1)$$

La valeur de OF pour s' représente l'allocation pour le reste des services. La séquence par laquelle OF est calculée affecte tout le réseau, parce que l'allocation d'une application à un AP diminue sa bande passante disponible. Donc, toutes les permutations possibles doivent être considérées. La fonction F consiste en la partie qualité Q et la partie politiques PT , avec leurs poids correspondants ($w_q + w_{pt} = 1$).

$$F = w_q Q + w_{pt} PT \quad (3.2)$$

Plus encore, les fonctions Q et PT sont analysées comme:

$$Q = w_{bi} BI + w_{di} DI + w_{sqi} SQI \quad (3.3) \quad \text{et} \quad PT = w_{cci} CCI + w_{npi} NPI + w_{tti} TTI \quad (3.4)$$

Chaque terme dans les sommes (3.3) et (3.4) représente un facteur spécifique, qui est calculé comme le produit d'un indicateur avec son poids correspondant. Noter que $w_{bi} + w_{di} + w_{sqi} = 100$ de même que $w_{cci} + w_{npi} + w_{tti} = 100$. BI = Bandwidth Indicator, DI = Delay Indicator, SQI = Signal Quality Indicator, CCI = Cost and Compatibility Indicator, NPI = Network Provider Indicator, and TTI = Terminal Type Indicator.

Surplus consommateur (Consumer surplus) – Avec le surplus client [4], les auteurs proposent une solution centrée utilisateur, qui signifie que les décisions sont prises du côté de l'utilisateur. Le schéma a été conçu pour le trafic temps réel avec la stratégie suivante. Premièrement, les utilisateurs surveillent l'interface radio et déterminent une liste des réseaux disponibles. Ensuite, ils prédisent les vitesses de transfert des réseaux de la liste, en prenant la moyenne des cinq derniers transferts de données et dérivent les temps de terminaison. Après cela, les utilisateurs calculent l'utilité, qui est une relation entre le budget et la flexibilité de l'utilisateur dans le temps de terminaison du transfert. Enfin, les utilisateurs calculent le "consumer surplus", qui est la différence entre l'utilité et le coût imposé par le réseau, pour chaque réseau candidat; et ils choisissent le meilleur, pour une demande de connexion. Les auteurs proposent de calculer d'abord le temps de terminaison prévu T_c (3.5) et donc l'utilité prévue (U_i), ainsi que le surplus du client CS (3.6) pour chaque réseau candidat. $T_c - ideal$ représente le temps de terminaison de transfert idéal de l'utilisateur et $T_c - max$ le temps de terminaison de transfert maximum qu'un utilisateur est prêt à attendre. Les auteurs définissent le temps de terminaison de transfert (en secondes) comme T_c . Il est lié à la taille du fichier et dépend de la vitesse suivant l'équation:

$$T_c = F_i / r \quad (3.5)$$

où F_i est la taille du fichier I en bits et r la vitesse moyenne pour le transfert total en bps. L'utilisateur vise à maximiser CS , sous les contraintes de dernier délai pour la terminaison du transfert du fichier, elle est calculée comme:

$$CS = U_i(T_c) - C_i \quad \text{sous} \quad T_c \leq T_c - max \quad (3.6)$$

Où CS est le surplus du consommateur, $U_i(T_c)$ est la valeur monétaire que l'utilisateur place sur le transfert du fichier i dans le temps de terminaison de transfert donnée (T_c) et C_i le coût chargé par le réseau pour la terminaison du transfert de fichier.

Fonction de profit – Dans cette fonction [5], les auteurs associent chaque handoff avec un profit qui est décidé par une fonction cible avec deux paramètres: le bande passante et le coût du handover. Pour le fonctionnement de la fonction de profit, les auteurs distinguent les handoffs réactifs des handoffs proactifs. Un handover réactif est initié lorsqu'un mobile est sur le point de traverser la frontière de la cellule courante, alors que le handover proactif ne peut être initié qu'à des périodes de temps discrètes quand une amélioration de la qualité d'une connexion est possible. Les paramètres utilisés dans le calcul du gain incluent les réseaux d'accès avec leurs bandes passantes maximales pouvant être, respectivement, délivrées à un seul utilisateur et le pourcentage de capacité utilisée, le maximum de bande passante requise par un application, et les bandes passantes des réseaux d'accès utilisées par un mobile pour le handover. Les auteurs définissent le coût d'un handover comme la perte de volume de données due au délai de handoff; il correspond au volume de données qui aurait pu être transmis durant le délai de handoff. Ainsi, le profit est la différence entre le gain et le coût. A chaque période de handover, la station mobile va comparer le profit à partir de différents réseaux et choisit celui qui procure de profit maximum.

Plus concrètement les auteurs associent chaque handover avec un profit (P), qui est décidé par une fonction cible f avec deux paramètres: le Gain (G) de bande passante et le Coût (C) du handoff.

$$P = f(G, C) \quad (3.7)$$

1. Gain de Bande passante.

Les paramètres utilisés dans le calcul du gain incluent:

N_i : le $i^{\text{ème}}$ réseau d'accès;

n_i : le maximum de bande passante pouvant être procurée à un seul utilisateur par N_i ;

η_i : le pourcentage de capacité de transmission déjà utilisée dans N_i ;

r_M : le maximum de bande passante requis par une application; des niveaux de QoS inférieurs avec $r_{M-1}, r_{M-2}, \dots, r_1$ peuvent être tolérés;

$m(i, tk)$: la bande passante de N_i utilisée par une station mobile durant deux "handoff decision epochs" $[t_k, t_{k+1})$

Le gain de bande passante G d'une décision de handoff à t_k est défini comme

$$G_i(t_k) = \begin{cases} m(i, t_k) - m(j, t_{k-1}) & k \geq 1, \\ m(i, t_k) & k = 0 \end{cases} \quad (3.8)$$

Dans (3.8), $k = 0$ signifie que la station mobile initie sa connexion pour la première fois.

$G_i(t_k)$ donne la différence en bande passante entre la prochaine période la présente période.

2. Coût du handover

Dans le modèle, les auteurs définissent le coût du handoff comme la perte de volume de données due au délai du handoff.

$$C(t_k) = m(i, t_{k-1})d(x, y) \quad (3.9)$$

Où $d(x, y)$ est le délai du handover lorsqu'une SM exécute un handover de la station de base x vers la station de base y

3. La fonction de profit

$$P_i = (t_{k+1} - t_k)G_i(t_k) - m(i, t_{k-1})d(x, y) \quad (3.10)$$

Dégradation de service – Les auteurs proposent l'idée de dégradation de service [6] pour traiter des utilisateurs de différentes priorités. En dégradant le trafic de priorité inférieure, plus de bande passante peut être libérée pour les utilisateurs de plus grande priorité. D'abord, les fournisseurs de services ont besoin de spécifier les niveaux de services en termes de bande passante offerte, classée comme, excellente, bonne, basic, et rejetée, pour chaque type d'application tel que la voix, la vidéo, et les données. Cette spécification de bande passante va être utilisée pour calculer la bande passante libérée (différence des bandes passantes avant et après la dégradation). Ensuite, une table de récompenses pour chaque classe de priorité d'utilisateur est définie: il y a trois QoS (excellent, bonne, et basic) et trois sortes de déconnexions (déconnexion forcée, coupure de handover, et rejetée), associées chacune à une récompense pour chaque type d'application. Cette table servira à calculer les points de récompense perdus (différences des points de récompense avant et après la dégradation). Enfin, la dégradation de service est le rapport de la bande passante libérée sur les points de récompense perdus. Lorsqu'une nouvelle connexion est demandée, le fournisseur de service trouve toutes les connexions dégradables potentielles, calcule leurs dégradations de service et commence la dégradation de la connexion qui a la plus grande utilité de dégradation.

En se référant aux tables 3.1 et 3.2, prenant un exemple. Considérons une connexion: priorité d'utilisateur 1; type d'application: vidéo; niveau de qualité: excellent. Lorsque la connexion est dégradée vers le niveau Bon:

La Bande passante libérée = $2000 - 384 \text{ Kbit/s} = 1616 \text{ Kbit/s}$;

Les points de récompense perdus = $700 - 600 = 100$;

L'utilité de dégradation = $1616/100 = 16.16$.

Les quatre types de fonctions précédemment décrits sont récapitulés dans la table 3.3 avec leurs entrées et leurs buts.

Table 3.2 Récompenses pour la classe de priorité d'utilisateur 1

Niveau de Qualité	Voix	Vidéo	Données
Excellent	300	700	1000
Bon	300	600	800
Basic	300	500	400
Déconnexion forcée	-5000	-5000	-5000
Coupure de Handover	-5000	-5000	-5000
Rejet	-2500	-2500	-2500
Niveau de Qualité	Voix	Vidéo	Données
Excellent	300	700	1000
Bon	300	600	800
Basic	300	500	400
Déconnexion forcée	-5000	-5000	-5000
Coupure de Handover	-5000	-5000	-5000
Rejet	-2500	-2500	-2500

Table 3.1 Bande passante pour différentes QoS

Application	Excellent (kbit/s)	Good (kbit/s)	Basic (kbit/s)	Rejected
Voix	30	30	30	0
Vidéo	2000	384	256	0
Données	100	50	10	0

Fonction	Entrées	But
Fonction Objectif	Indicateur de Qualité (bande passante, délai, et qualité du signal) et Indicateur de Politique (coût et compatibilité, fournisseur de réseau, type de terminal)	Maximise la somme des entrées pondérées
Surplus Consommateur	Utilité (réseau disponible et temps de terminaison prévu) et coût imposé par le réseau	Maximise la différence entre l'utilité et le coût
Fonction Profit	Gain de Bande Passante (bande passante disponible, pourcentage de capacité utilisée, exigences de l'application) et Coût de Handover (perte de données due au délai de handover)	Maximise la différence entre gain et coût
Utilité de Dégradation	Bande passante libérée et points de récompense perdus (perte de points durant la dégradation, selon la class de trafic/ qualité, priorité de l'utilisateur)	Maximiser le ratio de la bande passante libérée sur les points de récompense

Table 3.3 Fonctions: entrées et but

3.2.2 Mécanismes basés sur un modèle mathématique

En plus des fonctions de gestion qui viennent d'être décrites, des techniques mathématiques variées ont aussi été proposées pour la gestion des ressources radio dans les réseaux sans fil. Quatre techniques, basées sur des modèles mathématiques sont décrites présentées ci-dessous, où il est question de donner une idée des domaines appliqués plutôt que rechercher l'exhaustivité de la présentation.

Programmation Stochastique (stochastic programming SP)

Les auteurs de [7], discutent l'utilisation de la programmation stochastique ou "stochastic programming" (SP), dans la conception d'un module d'allocation proactive qui interagit avec les modules de contrôle d'accès dans chaque réseau d'accès sans fil. SP, est une technique de programmation mathématique qui est utilisée dans la prise de décision sous la contrainte de l'incertitude. En abrégé, il est expliqué comment l'incertitude de la demande pourrait être prise en charge pour décider de politiques d'allocation robustes et optimales. Le schéma proposé utilise, en fait, un sous ensemble de SP appelé SLP (Stochastic Linear Programming) ou programmation stochastique linéaire, et considère un scénario dans lequel un hotspot WLAN se recouvre avec une cellule d'un réseau cellulaire, et où un service de données unique (Single Common Service) d'une bande passante fixée (notée Q) est fourni dans les deux réseaux sans fil (N_1 pour le réseau le plus large, et N_2 pour le hotspot). Les capacités maximales de N_1 et N_2 sont notées B_1 et B_2 , respectivement. La capacité maximale d'un réseau renvoie à la capacité qui peut être fournie aux utilisateurs correspondants, de telle manière qu'au moins Q puisse être fournie à chaque utilisateur. Ces capacités représentent celles disponibles au moment de l'arrivée de la demande. Donc, le nombre maximum d'utilisateurs qui peuvent être supportés par N_1 et N_2 est B_1 / Q et B_2 / Q , respectivement.

Notons, D_{ij} , R_{ij} , et A_{ij} , la demande, le rejet (demande insatisfaite) et l'allocation pour l'utilisateur ij , tels que:

$$R_{ij} = D_{ij} - A_{ij}. \quad (3.11)$$

Lorsque l'allocation totale dans un certain réseau est inférieure aux ressources disponibles, c.a.d:

$$\sum_i A_{ij} = A_j < \frac{B_j}{Q} \quad \sum_i A_{ij} = A_j < \frac{B_j}{Q} \quad (3.12)$$

Le réseau j est sous utilisé de U_j , défini comme : $U_j = \frac{B_j}{Q} - A_j$ (3.13)

Une interconnexion entre les deux réseaux est dédiée à l'échange de messages de contrôle et au relais dans la transmission des paquets en vol, des utilisateurs subissant un handover vertical. Cette interconnexion est supposée suivre un certain profil délai-débit. Soit B_v la borne supérieure du nombre d'utilisateurs au dessus de laquelle l'échange ne peut plus être effectué avec un délai acceptable. Aussi, soit U_v la sous utilisation associée avec cette interconnexion. L'égalité suivante peut être établie:

$$\sum_{i \neq j} A_{ij} + U_v = B_v \quad (3.14)$$

Les constituants de l'opération de gestion comprenant la fonction de profit peuvent maintenant être détaillés. Soit x_{ij} le profit par utilisateur alloué ij . Aussi, soit y_j et y_v les coûts unitaires de sous utilisation pour le réseau j et l'interconnexion, respectivement. Enfin, notons z_{ij} le coût par rejet d'utilisateur. La fonction de profit total peut être écrite comme:

$$\prod_{SCS-PD} = \sum_{\forall i,j} x_{ij} \cdot A_{ij} - \sum_{c=\{j,v\}} y_c \cdot U_c - \sum_{\forall i,j} \sum_{s \in S} p_{ij}(s) \cdot z_{ij}(s) \cdot R_{ij}(s) \quad \prod_{SCS} = \sum_{\forall i,j} x_{ij} \cdot A_{ij} - \sum_{s=\{j,v\}} y_s \cdot U_s - \sum_{\forall i,j} z_{ij} \cdot R_{ij} \quad (3.15)$$

Vu ce qui précède, le programme linéaire peut être écrits comme suit:

Programme SCS-DD (Single Common Service-Demande Déterministe)

Max \prod_{SCS}

Sujet à : $A_j + U_j = \frac{B_j}{Q} \quad \forall j$ (3.16)

$$\sum_{i \neq j} A_{ij} + U_v = B_v$$

$$A_{ij} + R_{ij} = D_{ij} \quad \forall i, j$$

Toutes les variables integer positif

Le but étant de s'adapter au nombre maximum d'utilisateurs, dans les deux réseaux, résidants dans la zone de hotspot WLAN, tout en minimisant les coûts de la sous utilisation des ressources et le rejet des demandes.

Procéder à des allocations proactives avec une demande incertaine est un problème que la programmation linéaire ne peut maîtriser. La programmation linéaire stochastique (SLP), un sous ensemble des méthodes de programmation stochastique (SP), représente des techniques qui rendent viable la considération de la nature probabiliste des demandes dans les réseaux sans fil hétérogènes. Un problème formulé en programmation linéaire peut prendre la forme: $\max \{c^T x \mid Ax = b, x \in X\} \quad X \subset R^n$. Une formulation de cette forme implique souvent que le problème comprend seulement des paramètres déterministes, spécifiés dans les coefficients de la fonction objectif et dans les conditions, respectivement c^T et A , ainsi que dans les conditions contraignant les valeurs, b . Il y a, cependant, des problèmes où les solutions sont requises avec certains paramètres possédant un degré considérable d'incertitude. Limitant la discussion au cas où l'incertitude touche la partie droite, c.a.d :

$$\max \{c^T x \mid Ax = b, Tx = \xi, x \in X\} \quad \text{où } \xi \text{ est aléatoire,} \quad (3.17)$$

la résolution d'un tel problème se réduit au changement de la formulation en une formulation de programmation linéaire déterministe équivalente, en énumérant les scénarios de résultats possibles tout en associant chaque scénario avec sa probabilité. Une manière de réaliser ceci est par l'instauration d'une pénalité pour la non satisfaction de la contrainte posée par chaque scénario. Supposons, par exemple, qu'il y a K scénarios possibles et que la probabilité associée avec le $k^{ième}$ scénario est noté p_k . Soit q_k la pénalité par unité de la non satisfaction de la $k^{ième}$ contrainte, c.a.d, par unité de différence entre Tx et ξ_k . Ainsi, la formulation du problème déterministe équivalent devient:

$$\max \left\{ \begin{array}{l} c^T x + \sum_{k=1}^K p_k (q_k)^T y_k | Ax=b, \\ T_k x + y_k = \xi_k, k=1:K, x \in X \end{array} \right\} \quad (3.18)$$

Ici, y_k représente le relâchement du à la non satisfaction de la condition à cause de la nature probabiliste de ξ_k . Pour plus de tolérance, SLP laisse de la place pour des types de problèmes où les contraintes stochastiques n'ont pas besoin d'être absolument respectées, et il est acceptable que ces contraintes soient respectées avec des probabilités prescrites. De telles contraintes, appelées chances ou contraintes probabilistes, sont ajoutées à la formulation comme suit:

$$P(T_k x + y_k \leq \xi_k) \geq p, p \in (0,1) \quad (3.19)$$

Il est maintenant possible de présenter la formulation pour le problème SCS avec des demandes probabilistes. Soit S , l'ensemble de tous les scénarios possibles. Dans chaque scénario $s \in S$ la demande $D_{ij}(s)$ prend des valeurs spécifiques avec des probabilités prédéterminées. La probabilité que la demande $D_{ij}(s)$ courante ait une valeur spécifique, c.a.d $P(D_{ij}(s) = D)$ est $p_{ij}(s)$. L'incertitude de la demande peut être imposée au Programme SCS-DD à travers les contraintes allocation-rejet-demande, où la pénalité peut être appliquée au rejet. De cette manière, la pénalité (coût) d'une unité de rejet est $z_{ij}(s)$. Comme telle, la fonction de profit à maximiser devient:

$$\prod_{SCS-PD} = \sum_{\forall i,j} x_{ij} \cdot A_{ij} - \sum_{c=\{j,v\}} y_c \cdot U_c - \sum_{\forall i,j} \sum_{s \in S} p_{ij}(s) \cdot z_{ij}(s) \cdot R_{ij}(s) \quad (3.20)$$

Le SLP peut être écrits comme suit:

Programme SCS-PD

Max \prod_{SCS}

$$\text{Sujet à : } A_j + U_j = \frac{B_j}{Q} \quad \forall j \quad (3.21)$$

$$\sum_{i \neq j} A_{ij} + U_v = B_v$$

$$A_{ij} + R_{ij} = D_{ij} \quad \forall i, j$$

Toutes les variables integer positif

Logique Floue et Réseaux de neurones

Le travail présenté dans [8] vise à proposer un schéma de *JRRM* (Joint Radio Resource Management) basé sur la logique floue et les réseaux de neurones dans un réseau hétérogène comprenant UMTS, GERAN, et WLAN 802.11b, comme *TARs*. En particulier, un algorithme de *JRRM* à deux étapes, basé sur les deux techniques citées, y est présenté, pour la gestion du pool commun de ressources radio disponibles. La première étape sélectionne une combinaison adéquate de cellules construite autour des trois *TARs* disponibles, alors que la seconde étape choisit la *TAR* la plus appropriée à laquelle l'utilisateur devrait être rattaché. Un débit correct est aussi sélectionné pour chaque utilisateur dans cette seconde étape. Le schéma de *JRRM* proposé comprend trois fonctions principales de GRR: la sélection de cellule et du *TAR* (c.a.d., décider de du *TAR* et de la cellule à laquelle le mobile doit être rattaché au démarrage de la session), le débit alloué (c.a.d., décider du débit ou de la bande passante le plus approprié pour chaque *TAR* et chaque utilisateur admis), et le contrôle d'admission(c.a.d., décider si une demande de connexion peut ou non être acceptée).

Les données d'entrées disponibles pour les décisions de *JRRM* sont principalement les suivantes:

1. Les *TARs* déployées, la bande passante disponible pour chaque *TAR*, et la configuration envisagée (par exemple, niveau de puissance de transmission maximum des stations de bases, les séquences de codes disponibles);
2. Les mesures provenant des différentes *TARs* (niveaux de charge, par exemple) ainsi que les mesures provenant des équipements utilisateurs (*EUs*), telles que les niveaux de puissance reçues, la perte de chemin (path lost), ou la ratio E_c/I_o (Energie chip sur densité spectrale du bruit et des interférences) dans le cas des *TARs* basées sur le CDMA;
3. Aspects technico- économiques et subjectifs, incluant les politiques de l'opérateur, qui pourrait préférer l'utilisation de certaines *TARs* plutôt que d'autres pour différentes raisons (stratégies commerciales, propriété du réseau sans fil, etc.) ainsi que les profils des abonnés et les préférences des utilisateurs (QoS contre coût).

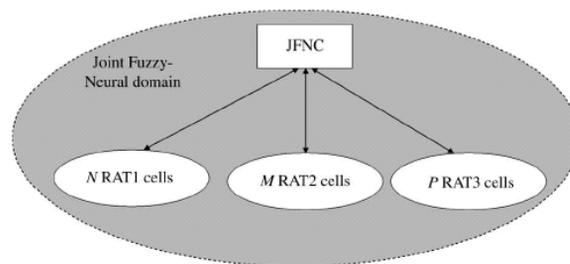


Figure3.1 Architecture proposée pour l'exécution de l'algorithme neuronal flou de GCRR

Comme le souligne les auteurs, les informations à la base de la prise de décision sont caractérisées par leur imprécision et leurs dissimilarités. La logique floue et les réseaux de neurones font partie des techniques intelligentes capables de traiter de pareilles informations; et c'est ce qui explique leur adoption dans le cas qui nous concerne. Alors que les réseaux de neurones sont bons pour reconnaître l'occurrence d'instances d'un certain modèle, ils ne sont pas bons dans l'explication du cheminement qui les guide vers leurs décisions. En contre partie, les systèmes de logique floue sont bons dans l'explication de leurs décisions à partir d'informations imprécises, mais ne peuvent pas acquérir automatiquement les règles qu'ils utilisent pour prendre les décisions ou pour régler les fonctions qui convertissent les valeurs numériques en des quantités floues. A cause des limitations de ces deux techniques, des systèmes hybrides intelligents les combinant, et donc outrepassant ces limitations, ont été créés. En termes d'architecture réseau, une allocation de haut niveau des fonctions de *JRRM* dans un scénario de *TARs* hétérogènes est montrée sur la figure 3.1. La région de service est subdivisée en des domaines flous et neuronaux, incluant chacun un sous ensemble des cellules appartenant à des *TARs* différentes. Chaque domaine est géré par un JFNC (Join Fuzzy Neural Controller), qui est en charge de l'exécution de l'algorithme de *JRRM* pour l'ensemble des cellules sous son domaine.

L'algorithme de *JRRM* floue et neuronal

Le schéma proposé pour l'implémentation de l'algorithme de *JRRM* basé sur mécanisme floue et neuronal, consiste en trois blocs principaux montrés sur la figure 3.2, et identifiés comme "fuzzy-based decision", "reinforcement learning", et "multiple objective decision making", respectivement. Ces trois blocs sont exécutés à chaque fois qu'un nouvel utilisateur demande à être admis au système, et durant la session utilisateur. Ils assurent l'allocation et la récupération des ressources radio dans le scénario et la sélection de la *TAR* la plus appropriée, tout en gardant les exigences de QoS désirées de tous les utilisateurs admis. Les entrées du bloc "fuzzy-based decision" sont un ensemble de variables linguistiques LV_i qui correspondent à différentes mesures. La sélection de ces variables linguistiques a été faite en tenant en compte des paramètres les plus influant sur la sélection de la *TAR* et l'allocation de bande passante (force du signal en rapport avec les *TARs* considérées, la charge des cellules, et la vitesse des

mobiles). En plus, des critères subjectifs et technico-économiques, sous la forme de UP (User Preferences) et OP (Operator Preferences), sont aussi des entrées pour le bloc " multiple objective decision making ", puisqu'ils sont aussi des éléments clé qui devraient jouer un rôle dans les décisions de JRRM. D'un autre côté, les sorties de l'algorithme neuronal et floue sont subdivisées en deux groupes, qui sont les indicateurs moteurs de la sélection cellule/TAR et de l'allocation de bande passante.

1. Pour effectuer la sélection cellule/TAR, chaque TAR est caractérisée par un indicateur, qui prend des valeurs dans l'intervalle [0, 1], référencé par valeur FSD (Fuzzy Selected Decision), qui évalue

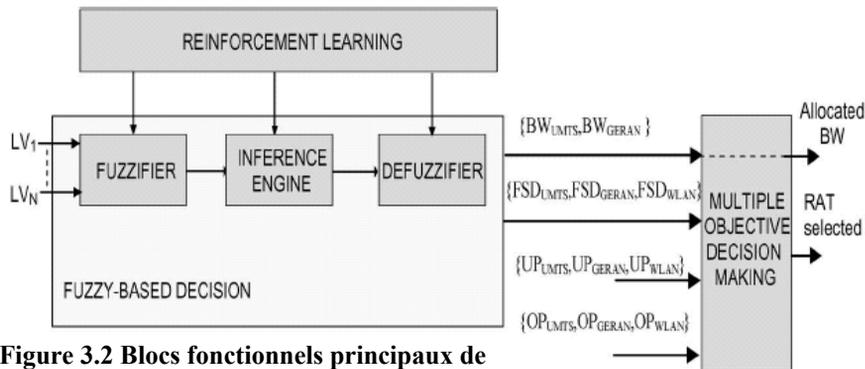


Figure 3.2 Blocs fonctionnels principaux de l'approche JRRM proposée

le bien fondé de la sélection d'une TAR avant les autres.

2. Pour effectuer l'allocation du débit, une valeur de sortie (B) est associée à chaque TAR, qui donne une indication de la bande passante qui devrait être assignée à l'utilisateur.

Dans la suite nous décrivons chaque bloc représenté dans la figure 3.2.

A-Fuzzy-Based Decision

Ce processus est exécuté en trois étapes par les blocs fuzzifier, moteur d'inférence (inference engine), et defuzzifier.

Etape 1. Fuzzifier: Ce bloc prend en charge la fuzzification dont l'objectif est d'assigner pour chaque variable linguistique LV_i , une valeur (entre 0 et 1) donnée par la fonction $\mu_{x_j}(LV_i)$, qui correspond au degré d'appartenance de cette entrée à un ensemble flou donné $X_j(LV_i)$. Un total de sept variables linguistiques sont considérées dans l'approche de JRRM proposée, et incluent les variables suivantes:

1. SS_{UMTS} , SS_{UMTS} , SS_{WLAN} (force du signal reçu de chacune des TARs considérées);
2. RA_{UMTS} , RA_{GERAN} , RA_{WLAN} (disponibilité des ressources dans chacune des TARs considérées);
3. MS (vitesse du mobile)

Une variable linguistique est caractérisée par un ensemble de termes (set term) qui prennent leurs valeurs dans l'ensemble $\{H, M, L\}$ (c.a.d., High, Medium, Low). Les ensembles de termes considérés sont:

$$\begin{aligned} X(SS_{UMTS}) &= X(SS_{UMTS}) = X(SS_{WLAN}) = X\{L, H\} \\ X(RA_{UMTS}) &= X(RA_{UMTS}) = X(RA_{WLAN}) = X\{L, M, H\} \\ X(MS) &= X\{L, H\}. \end{aligned} \quad (3.22)$$

Le nombre de valeurs dans chaque "term set" est choisit de manière à avoir un nombre limité de combinaisons entre eux. La forme de la fonction d'appartenance peut être un triangle, un trapèze ou une fonction en forme de cloche.

Etape 2. Moteur d'Inférence (Inference Engine): Dans un contrôleur flou, le comportement dynamique du système est caractérisé par un ensemble de règles linguistiques exprimant les politiques de décisions et sont définies comme suit:

IF (un ensemble de conditions satisfaites) THEN (un ensemble de conséquences peuvent être inférées).

La sortie du moteur d'inférence est encore une valeur floue définie par rapport à deux groupes de variables linguistiques de sortie. D'un côté, les variables linguistiques de sortie de décision D (D_{UMTS} , D_{GERAN} , D_{WLAN}) ont été définies avec les "term sets" suivants:

$$X(D_{UMTS}) = X(D_{GERAN}) = X(D_{WLAN}) = X\{Y, PY, PN, N\} \quad (3.23)$$

Où Y est pour dire "Yes", PY pour "probablement Yes", PN pour "probablement No", et N pour "No". De manière similaire, il y aura deux variables linguistiques de sortie correspondantes avec le débit alloué B (B_{UMTS} , B_{GERAN}), chacune avec les "term sets" suivants:

$X(B_{UMTS}) = X(B_{GERAN}) = X\{H, M, L\}$ où H correspond à "High", M à "Medium", et L à "Low". Un exemple de trois règles d'inférence est montré dans la table 3.4. En supposons que la $j^{\text{ème}}$ règle correspond à la première ligne de la table 3.4, la valeur d'appartenance (membership) de la sortie de la $j^{\text{ème}}$ règle est définie comme:

$$\begin{aligned} \mu_Y(D_{UMTS})_j &= \mu_N(D_{GERAN})_j = \mu_N(D_{WLAN})_j \\ &= \mu_H(B_{UMTS})_j = \mu_L(B_{GERAN})_j \\ &= \min [\mu_H(SS_{UMTS}), \mu_L(SS_{GERAN}), \mu_L(SS_{WLAN}), \mu_H(RA_{UMTS}), \mu_H(RA_{GERAN}), \mu_M(RA_{WLAN}), \mu_L(MS)]. \end{aligned} \quad (3.24)$$

Donc, finalement, les conséquences des règles dans la base des règles floues conduisant au même ensemble flou d'une variable linguistique de sortie donnée, doivent être combinées dans le but d'obtenir la

TABLE I
EXAMPLE OF INFERENCE RULES

IF							THEN				
SSUMTS	SSGERAN	SSWLAN	RAUMTS	RAGERAN	RAWLAN	MS	DUMTS	DGERAN	DWLAN	BUMTS	BGERAN
H	L	L	H	H	M	L	Y	N	N	H	L
H	L	L	M	H	M	L	PY	N	N	M	L
H	L	L	L	H	M	L	PN	PN	N	L	L

Tables 3.4 exemple de règles d'inférence

valeur d'appartenance de l'ensemble flou. En particulier, elle est calculée comme la somme des valeurs d'appartenance de toutes les règles ayant cet ensemble flou comme conséquence.

Etape 3. Defuzzifier: Enfin, le defuzzifier exécute l'opération de "defuzzification", qui consiste à convertir les sorties du moteur d'inférence en valeurs numériques, dénotées par FSD (Fuzzy Selected Decision), indiquant l'utilité de la sélection de chaque TAR . Les trois sorties, FSD_{UMTS} , FSD_{GERAN} , et FSD_{WLAN} , sont alors obtenues. De manière similaire, ce processus fournit aussi les débits alloués au moyen des sorties BW_{UMTS} BW_{GERAN} .

Reinforcement Learning

Les étapes de la procédure de décision ci-dessus peuvent être représentées au moyen d'une structure à cinq couches utilisée dans les réseaux de neurones, qui permet l'utilisation d'une procédure d'apprentissage par renforcement pour l'ajustement des différentes fonctions d'appartenance. Cette structure neuronale floue est montrée sur la figure 3.4 et se trouve composée d'un ensemble de nœuds appartenant aux différentes couches. Chaque nœud de la $k^{\text{ème}}$ couche ($k = 1, \dots, 5$) est numéroté par un indice qui varie de 1 jusqu'au nombre de nœuds dans la $k^{\text{ème}}$ couche. La structure basique de chaque nœud est représentée sur la figure 3.3, où u_{ki} dénote le $i^{\text{ème}}$ signal d'entrée pour la $k^{\text{ème}}$ couche, et p le nombre

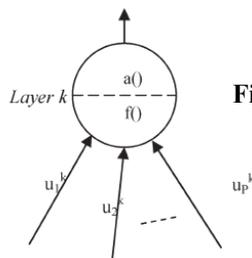


Figure 3.3 Structure basique d'un nœud

d'entrées connectées au nœud. Chaque nœud est caractérisé par une fonction d'intégration $f(u_{k1}, u_{k2}, \dots, u_{kp})$, qui combine les différentes entrées, et par une fonction d'activation $a(f)$, qui fournit la sortie. Une fois le réseau neuronal défini au moyen des cinq couches, des variables linguistiques d'entrée/sortie, et des "term sets" correspondants, il est nécessaire d'initialiser la structure du réseau neuronal flou en

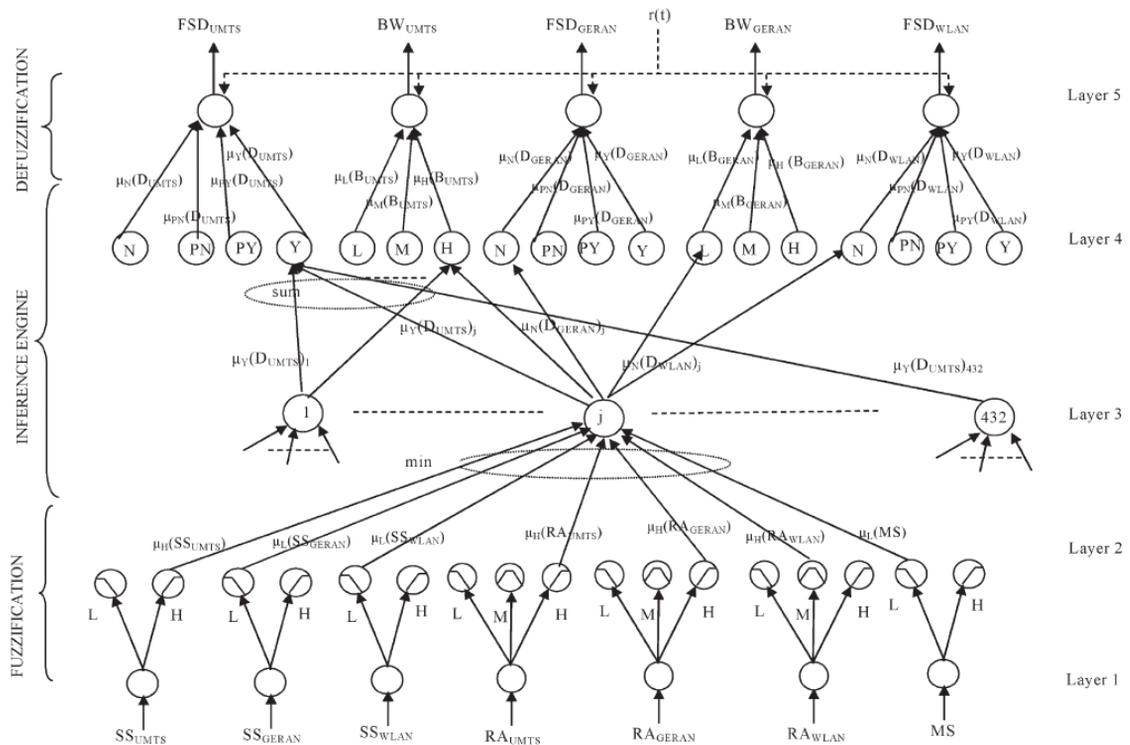


Figure 3.4 Structure neuronale floue

définissant les règles d'inférence floues contenues dans la base des règles floues et la forme initiale et la position des fonctions d'appartenance. Il importe de noter que cette phase d'initialisation est effectuée off-line et qu'après cela, l'apprentissage par renforcement est en charge de l'ajustement, online, des paramètres définissant la structure du réseau neuronal flou.

Multiple Objective Decision Making

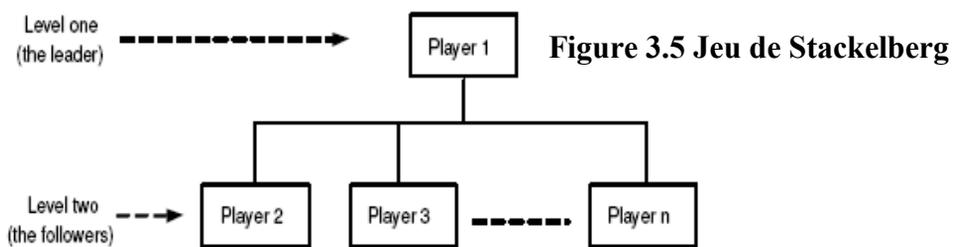
Dans un architecture hétérogène, la sélection de la *TAR* pourrait ne pas dépendre uniquement des aspects relatifs à l'interface radio (force du signal, ressources disponibles), mais dépendrait aussi d'entrées qualitatives ou technico-économiques. En particulier et comme montré sur la figure 3.2, le bloc de prise décision multi objectifs proposé prend en entrée la sortie du bloc FSD, ainsi que le UP (User Preferences) et OP (Operator Preferences). Les critères considérés pour chaque TAR sont $C1 = FSD$, $C2 = UP$, $C3 = OP$. Du au fait que les différents critères pourraient avoir différentes importances subjectives, un nombre $\alpha \geq 0$ indicateur de l'importance du critère est introduit. La décision est prise alors selon :

$$D = C_{i_1}^{\alpha_1} \cap C_{i_2}^{\alpha_2} \cap C_{i_3}^{\alpha_3} \quad \text{où} \quad \frac{1}{N} \sum_{n=1}^N \alpha_n = 1 \quad (3.25)$$

N étant le nombre de critères ($N=3$ dans ce cas). De cette manière, le critère de faible importance a moins d'influence sur la décision de sélection toute entière.

3.2.3 Mécanismes basés sur la Théorie des jeux

La théorie des jeux appliquée à la GRR dans les réseaux sans fil hétérogènes a fait l'objet de plusieurs travaux. Celui de [9], par exemple, examine le problème de l'allocation des ressources dans un réseau hétérogène en utilisant les jeux coopératifs de Stackelberg. En formulant le problème comme un jeu coopératif de Stackelberg il est permis aux réseaux individuels de coopérer entre eux pour former des coalitions. Par conséquent, l'objectif de chaque réseau est de maximiser l'objectif global du réseau hétérogène et de satisfaire les demandes d'allocation de ressources provenant des utilisateurs. Chaque membre de la coalition fournit une partie des ressources demandées, suivant ses contraintes opérationnelles. Un jeu coopératif de Stackelberg [10], est joué dans lequel le joueur dominant (Stackelberg leader) choisit une stratégie dans la première étape, qui prend en compte la réaction la plus vraisemblable des partisans. Dans la seconde étape, les partisans (Stackelberg followers) choisissent leurs propres stratégies après avoir observé la décision du leader, c.a.d., ils réagissent à la stratégie du leader. La figure 3.5 représente un jeu de Stackelberg sur laquelle le réseau/joueur 1 assume le rôle de leader, et le reste des réseaux celui des partisans.



Un jeu coopératif est celui dans lequel les joueurs ont cette option de planifier en tant que groupe, avant de choisir leurs actions. Un jeu coopératif consiste en deux éléments: un ensemble de joueurs et une fonction caractéristique spécifiant la valeur créée par différents sous ensembles (coalitions) de joueurs dans le jeu. Il est fait usage, dans le travail proposé, des définitions suivantes:

Coalition : Soit $N = \{1, 2, \dots, n\}$ un ensemble de joueurs. Des sous ensembles non vides de N , $S, T \subseteq N$; sont appelés coalition. Une coalition de n joueurs est donnée par la paire (N, v) , où v est la fonction caractéristique.

Grand Coalition : Une "Grand coalition" est une coalition qui inclut tous les joueurs.

Fonction Caractéristique : C'est une fonction notée v , qui affecte à chaque coalition S son gain maximum, le revenu total attendu de la coalition, noté $v(S)$.

Le noyau: Le noyau est l'ensemble de tous les résultats faisables qu'aucun joueur ou coalition ne peut améliorer en agissant pour lui-même. Ce concept a été développé pour les solutions des jeux coopératifs. Il consiste en toutes les allocations non dominées dans le jeu (undominated). Une allocation dans le noyau d'un jeu sera toujours une allocation efficace.

Pareto-optimale : L'optimalité de Pareto est une mesure de l'efficacité. Un résultat d'un jeu est Pareto-optimale s'il n'y a pas d'autres résultats qui rendent plus nanti au moins un joueur, sans rendre les autres joueurs moins nantis. Cela revient à dire qu'un résultat Pareto-optimale ne peut pas être amélioré sans léser au moins un joueur.

Imputation : Le partage du résultat qui peut être réalisé entre tous les joueurs coopérants est appelé une imputation du jeu. Les éléments $x = (x_1, x_2, \dots, x_n)$ représentent le résultat de chaque joueur (imputation) i sous la coalition P . Le couple (x, P) est appelé configuration de solution.

Pour ce qui est de l'approche de l'allocation, les auteurs considèrent un problème d'allocation pour un utilisateur mobile ayant accès à un réseau sans fil hétérogène (utilisateur équipé d'interfaces réseaux multiples). Ils supposent que chaque joueur N_i ($i = 1, \dots, n$ où n est le nombre de réseaux participants), a une capacité c_i de ressources. La configuration des capacités des réseaux sans fil est donnée par $c = (c_1, c_2, \dots, c_n)$, et le total des ressources des réseaux est donné par:

$$cf_i = e^{\frac{l_i}{c_i}} \quad (3.26)$$

Chaque réseau utilise ses ressources disponibles c_i pour la contribution $r_i \geq 0$ à la demande d'allocation de ressources totale R , telle que $r_i \leq c_i$. Dans les jeux de Stackelberg originaux, chaque joueur tente d'optimiser son propre objectif par rapport aux décisions prises par les autres joueurs. L'objectif de chaque joueur est donné par le problème d'optimisation suivant:

$$\begin{aligned} \max_{r_i} \quad & u_i(r_i) \\ \text{tel que} \quad & r_i \leq c_i, \\ & r_i \geq 0. \end{aligned} \quad (3.27)$$

Où u_i et r_i sont respectivement la fonction utilité et les ressources allouées se rapportant au joueur i . Allouer les ressources selon les capacités des réseaux individuels, sans considérer les conditions globales du réseau hétérogène ne résultera pas en une solution Pareto-optimale. Pour se faire, la coopération entre les réseaux individuels par la formation d'une coalition entre eux, est requise. Une structure de coalition est notée par $P = P_1, P_2, \dots, P_k$, où $P_i \cap P_j = \emptyset$ pour tout $i \neq j$. Après la formation d'une coalition, les joueurs individuels agiront, chacun pour le bénéfice de la coalition à laquelle il appartient. Par conséquent, l'objectif de chaque membre dans la coalition devient, l'optimisation de l'objectif de la coalition soumise aux contraintes opérationnelles. La fonction objectif de la coalition peut être formulée comme suit:

$$u'_{P_j}(R) = \sum_{i \in P_j} u_i(r_i) \quad (3.28)$$

Où u'_i est la fonction utilité de la coalition qui maximise son résultat total pour l'allocation de R ressources. Ainsi, le total des ressources allouées à l'utilisateur mobile dans le réseau hétérogène est donné par : $R = \sum_{i=1}^m r_i$.

Les auteurs signalent aussi qu'il est dans l'intérêt d'une coalition de prendre en compte, avec les buts de l'allocation des ressources et la maximisation du résultat, l'équilibrage de charge dans les réseaux, et de minimiser les délais subis par l'utilisateur de données sur les multiples réseaux. Ils emploient, à ce titre, une fonction de coût exponentielle pour affecter un facteur de congestion (qui est une mesure de l'utilisation d'un réseau) aux réseaux individuels. Le facteur de congestion est calculé pour chaque réseau avec:

$$cf_i = e^{\frac{l_i}{c_i}} \quad \text{où } l_i \text{ est la charge du réseau et } c_i \text{ sa capacité.} \quad (3.29)$$

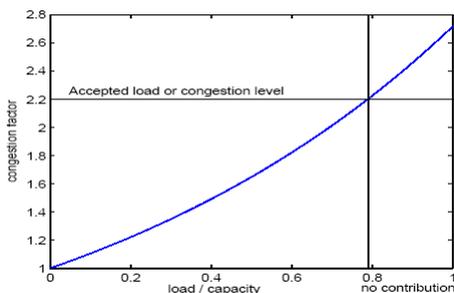


Figure 3.6 Courbe du facteur de congestion

Sur la base du facteur de congestion, les réseaux décident d'allouer ou non des ressources. Ce facteur (représenté par sa courbe sur la figure 3.6) est utilisé pour optimiser les performances du réseau, comme par exemple, minimiser la charge ou les délais du réseau. Le revenu des allocations de ressources sera, avec le facteur de congestion, différent pour chaque réseau est dépendra de la quantité de ressources déjà en utilisation. Si une requête ne peut être satisfaite sur aucun réseau, elle est rejetée. Sinon, elle est allouée sur les réseaux où la somme des revenus est maximisée. Soit, alors le revenu de l'allocation de

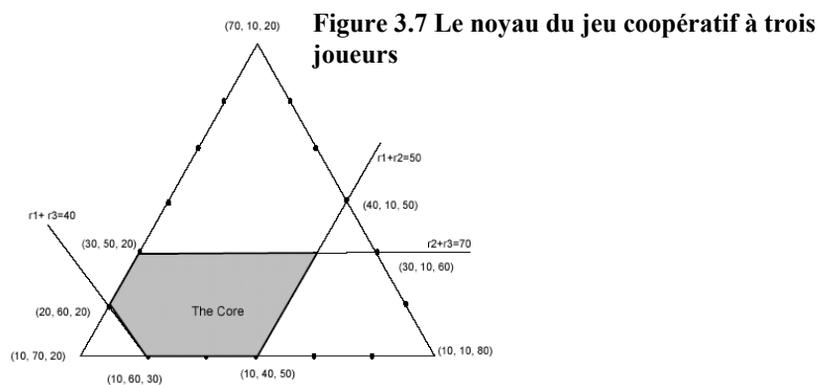
ressources sur le réseau i est donné par $x_i = \lambda_i * r_i$. La coalition est caractérisée par le revenu total maximum dénoté par $\mathbf{x} = (x_1, x_2, \dots, x_n)$, où x_i est le revenu alloué à joueur i .

Exemple illustratif

Considérons un jeu avec trois joueurs/réseaux. Soit $\mathbf{r} = (r_1; r_2; r_3)$ de dimension 3 représentant le nombre de ressources allouées à l'utilisateur par les trois réseaux. L'objectif est d'allouer les ressources de telle sorte que le revenu total de la coalition soit maximisé, sous les contraintes des capacités en ressources de chaque réseau. Il est supposé que le demande de débit de l'utilisateur est égale à 100 Mbits/s. par conséquent, $r_1 + r_2 + r_3 = 100$. Soit $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ un vecteur de dimension 3 représentant $1/cf$ pour les multiples réseaux. Le vecteur $\mathbf{x} = (x_1, x_2, x_3)$ représentant le revenu total est donnée par $\mathbf{x} = \lambda r^T$. La fonction caractéristique du jeu est donnée par la table 3.5. Les imputations sont les points $(r_1; r_2; r_3)$ tels que $r_1 + r_2 + r_3 = 100$ et $r_1 \geq 10, r_2 \geq 10, r_3 \geq 20$. L'ensemble des imputations pour le jeu est représenté graphiquement sur la figure 3.7. Cette figure montre le noyau (core) et d'autres points solution variés pour le jeu. Le noyau consiste en toutes les imputations de région à l'intérieur du trapèze.

Table 3.5 Structure de coalition

Coalition	Coalition value
$\{\emptyset\}$	0
$\{1\}$	10
$\{2\}$	10
$\{3\}$	20
$\{1,2\}$	50
$\{1,3\}$	40
$\{2,3\}$	70
$\{1,2,3\}$	100



3.2.4 Mécanismes basés sur AHP et GRA (Analytic Hierarchy Process et Grey Relational Analysis)

Dans [11], l'objectif de conception est de garder les utilisateurs mobiles toujours connectés, de la meilleure manière possible (Always Best Connected). Deux techniques mathématiques sont combinées, dans un mécanisme, pour décider du réseau optimum pour les utilisateurs, à travers la recherche d'un équilibre entre trois facteurs de décisions. Les facteurs de décision adoptés, sont les caractéristiques de QoS en matière de préférences de l'utilisateur, d'application de service et de conditions du réseau. Dans le but de supporter la meilleure QoS possible, les auteurs tiennent compte de plusieurs de facteurs de QoS. Les facteurs sont construits dans une hiérarchie AHP basée sur leurs interrelations. AHP est un moyen de décomposition d'un problème compliqué en une hiérarchie de sous problèmes plus abordables. Ces sous problèmes sont souvent appelés facteurs de décision et pondérés selon leurs relatives dominance dans le problème. Les facteurs de décision en bas de la hiérarchie sont en fait les alternatives de la solution. AHP sélectionne la solution alternative avec le poids synthétisé le plus grand. AHP effectue les quatre étapes principales suivantes : décomposition, comparaison par paires, calcul des poids locaux, synthèse des poids. La figure.3.8 représente une hiérarchie pour le "choix d'une maison".

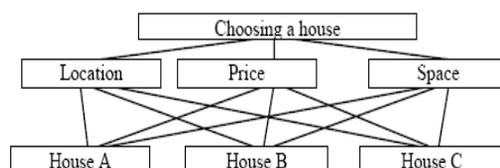


Figure 3.8. Un exemple d'une hiérarchie AHP

Pour le cas d'un réseau hétérogène, la QoS est placée au sommet de la hiérarchie comme objectif. Six facteurs principaux de QoS sont définis dans [] pour décrire la condition du réseau (*availability*, *throughput* (α), *timeliness* (β), *reliability* (γ), *security* (δ), *cost* (ϵ)), qui sont placés dans le second niveau

de la hiérarchie. Pour une prise de décision correcte et efficace, le facteur β (*timeliness*) est décomposé en trois éléments ζ (*delay*), η (*response time*), et θ (*jitter*), tandis que le facteur γ (*reliability*) est divisé en les éléments λ (*BER*), μ (*burst error*), ν (*average number of retransmissions per packet*), et σ (*packet loss ratio*). Ces sous facteurs sont arrangés dans le troisième niveau. Pour éviter aux utilisateurs se déplaçant à grandes vitesses, des handovers fréquents, les auteurs adoptent *RSS* (*received signal strength*) et *CA* (*coverage area*) en tant que sous facteur de "availability". Enfin, à la base de la hiérarchie se trouvent les deux réseaux sans fil choisis en tant qu'alternatives dans le travail proposé.

Le mécanisme de sélection de réseau est divisé en trois blocs de fonctions logiques principaux, (représentés sur la figure 3.9) : "collecting data", "processing data" et "making decision".

Deux types de données (préférences de l'utilisateur et conditions du réseau) en matière de paramètres de QoS sont collectés. Les données basées utilisateur sont traitées par AHP pour dériver les poids globaux GW(Global Weights) des facteurs de second niveau et les poids locaux LW1,LW2(Local Weights) des facteurs de troisième niveau. Si $GW = \{w_\alpha, w_\beta, w_\gamma, w_\delta, w_\epsilon, w_\kappa\}$, $LW1 = \{w_\zeta, w_\eta, w_\theta\}$ et $LW2 = \{w_\lambda,$

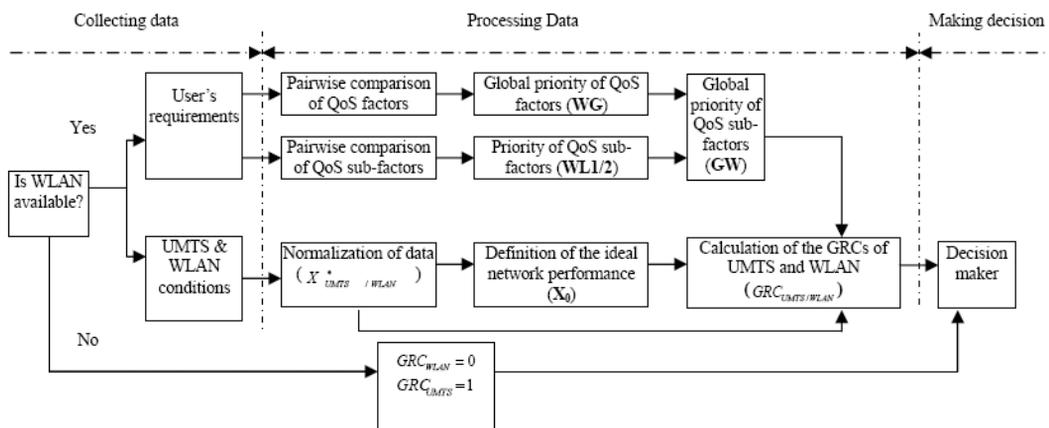


Figure 3.9 Modèle de sélection de réseau basé sur AHP GRA

$w_\mu, w_\nu, w_\sigma\}$, les poids finaux de tous les facteurs sont la synthèse des poids locaux et des poids globaux correspondants,

$$w = \{w_1, w_2, \dots, w_{10}\} = \left\{ \begin{array}{l} w_\alpha, w_\beta, w_\xi, w_\beta, w_\eta, w_\beta, w_\theta, \\ w_\gamma, w_\mu, w_\gamma, w_\nu, w_\gamma, w_\epsilon, w_\delta, w_\epsilon \end{array} \right\}. \quad (3.30)$$

Les données basées réseau sont normalisées par GRA selon deux situations: *larger-the-better* et *smaller-the-better*. GRA est une méthode d'analyse du degré d'interrelation pour les séquences discrètes. Une des séquences est définie comme séquence de référence représentant la situation idéale. Les interrelations *grey* entre la séquence de référence et les autres séquences peuvent être réalisées en calculant le coefficient GRC (Grey Relational Coefficient) selon le niveau de similitude et de variabilité. La séquence la plus désirable est celle qui a le plus grand coefficient GRC. GRA est souvent implémenté suivant trois étapes: normalisation des données, définition de la séquences idéale et le calcul des GRCs.

Chapitre 4

SCHEMA POUR LA GESTION DES RESSOURCES RADIO DANS UN RESEAU SANS FIL ET MOBILE HETEROGENE

4.1 Introduction

Des travaux de recherche sur les réseaux mobiles 3G et 4G futures ont mis en évidence, en plus du fait que ces derniers seront hétérogènes en termes de technologies d'accès, leur complexité en termes de solutions de gestion et, en particulier, de gestion des ressources radio (**GRR**).

Notre approche pour la GRR dans un réseau sans fil et mobile hétérogène (RSMH) est une exploration de la possibilité d'utiliser le mécanisme du handover vertical (**HOV**) en tant qu'outil de base pour cette gestion. Cette exploration s'étendra aux aspects connexes de la question, pour une prise en charge qui soit cohérente avec le principe retenu. En effet, traiter de la GRR dans les RSMH oblige à la considérer dans ses rapports avec les autres fonctions du réseau où elle intervient à des degrés divers. Les exigences propres à ces dernières sont autant de contraintes dont la GRR doit tenir compte (ou de critères qu'elle doit utiliser) pour jouer pleinement son rôle.

Cette approche s'articulera autour d'un schéma de redistribution périodique des ressources radio, dont le but est de mettre en œuvre une politique agressive de réutilisation des ressources radio "oisives", dans tout le réseau hétérogène. Cela se traduira par une quête permanente des ressources radio "oisives" en vue de leur utilisation là où le besoin se fait sentir. Le schéma devra répondre à des critères d'efficacité, de flexibilité, d'équité, de simplicité, d'extensibilité ... etc.

Ce schéma est complété (pour une approche globale) par des propositions traitant du volet GRR en rapport avec les autres fonctions importantes d'un RSMH, à savoir: la gestion de la mobilité, les handovers inter technologies, le contrôle d'appels d'admission, la prise en compte de la QoS et, de manière indirecte, le contrôle de puissance. Des questions critiques envers notre approche seront posées et leurs réponses présentées dans le but de montrer la faisabilité, sur le plan théorique et pratique, de cette dernière, sa globalité ainsi que sa cohérence.

La suite de ce chapitre sera organisée comme suit : dans la première section nous présenterons le principe sous tendant notre approche ainsi que ses motivations. Dans la deuxième section nous détaillerons notre modélisation d'un RSMH où nous introduirons la notion et le rôle des "capteurs de contextes". La troisième section traitera dans le détail du schéma proposé et de ses caractéristiques et sera suivie par une section qui discutera de certaines questions critiques envers l'approche elle-même. Les aspects GRR ayant trait à la gestion de la mobilité, à la prise en charge de la QoS ainsi qu'au contrôle d'appels d'admission, se verront consacrer trois autres sections différentes. Enfin, le chapitre sera terminé par une conclusion.

4.2 Modélisation d'un réseau sans fil hétérogène

Un réseau sans fil hétérogène (**RSH**) est un composite formé de deux ou plusieurs technologies d'accès radio (**TAR**), ayant chacune ses propres caractéristiques en termes de couverture, de QoS assurée, d'implémentation, de coût, ...etc. Lorsqu'un RSH supporte aussi la mobilité des stations clientes on parle alors d'un réseau sans fil et de mobiles hétérogène (**RSMH**). Pour profiter pleinement de ces réseaux les utilisateurs devront être équipés de **terminaux mobiles multimodaux**, c.a.d de terminaux capables de communiquer en utilisant plusieurs technologies d'accès radio (simultanément ou non). Un exemple de réseau hétérogène sans fil est schématisé sur la figure 4.1.

La présence sur un même espace géographique de plusieurs réseaux sans fil différents (et donc des TAR différentes aussi) crée entre leurs zones de couverture de base sans fil (**ZBS**) des **zones de chevauchement de couverture sans fil** (wireless overlays), caractéristiques des RSMH. L'existence de ces zones de chevauchement de couverture sans fil (**ZCS**) dans un RSMH est génératrice d'opportunités très intéressantes de coopération, entre les différents réseaux constituants. Ce fait, a constitué tout au long de notre travail, une sorte de fil conducteur de notre approche pour la GRR dans les RSMH. Cela tient à

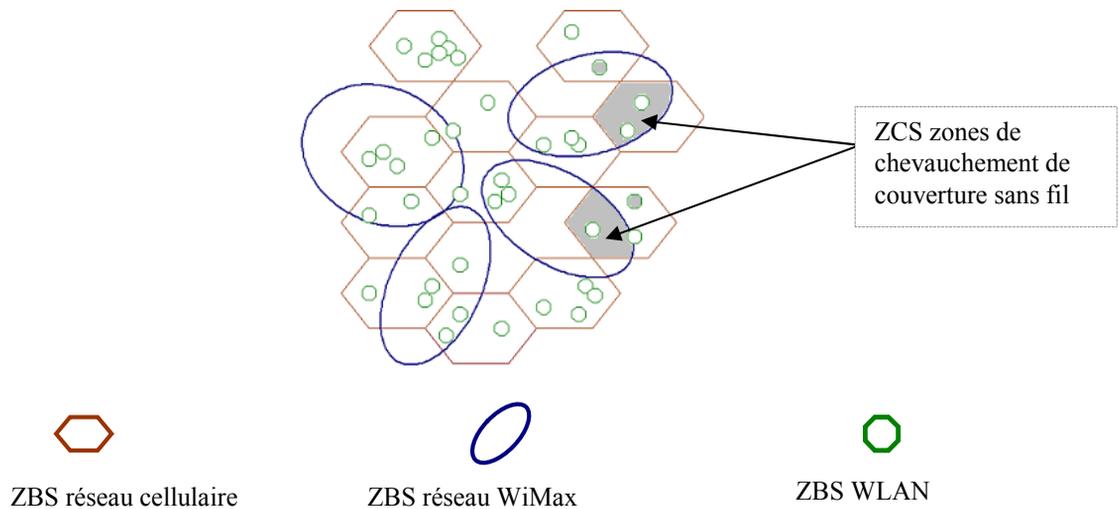


Figure 4.1. Un réseau hétérogène formé par trois technologies d'accès différentes

plusieurs raisons (explicitées dans nos développements ultérieurs), mais déjà à ce stade nous pouvons, commencer par évoquer la raison suivante :

La GRR dans les RSMH telle que nous l'avons conçue, s'appuie sur une connaissance assez précise du contexte des différentes ZBS de tout le RSMH. Selon cette conception, le système devrait répondre aux impératifs suivants :

- nécessité de connaître le nombre de stations (fixes et mobiles) dans chaque ZBS, et particulièrement leur répartition dans les diverses zones (ZCS,ZUS), ainsi que le volume de ressources alloué, les interfaces, la classe de QoS de chacune de ces stations,
- pouvoir estimer l'évolution de ces paramètres à court et moyen terme,
- décompter le nombre de handover verticaux (**HOV**) dans les divers sens de chaque ZBS,
- suivre les itinéraires des diverses stations (ou certaines d'entre elles) à travers tout le RSMH jusqu'à terminaison de leurs sessions de communication,
- estimer l'état de charge de chaque ZBS ainsi que la qualité de l'environnement radio qui y règne (interférences, erreurs, puissances d'émission, ... etc).

Pour permettre au système d'être au fait du contexte à tout moment, et à un coût raisonnable, nous avons adopté une démarche qui consiste à le doter d'entités logiciel appelées **capteurs de contextes**, et à dresser une sorte de carte du RSMH faisant ressortir ses différentes **ZCS** et **ZUC** (zone d'unicité de couverture sans fil). Chaque capteur de contexte (**CP**) est attaché à une ZBS particulière et peut communiquer avec les autres CPs du réseau pour les besoins de sa tâche. La figure 4.2 illustre un exemple de situation où les CP sont interpellés, et qui montre l'utilité de découper les zones ZBS en zones ZCS et zones ZUS. Sur la figure 4.2 sont placés, à des endroits différents, cinq équipements d'utilisateurs (A,B,C,D,E) qu'il est question, pour les CP, de localiser. Nous signalons tout de suite qu'il ne s'agit pas de donner les positions exactes de ces équipements (à la manière des systèmes GPS), mais plutôt de les situer par rapport aux différentes zones (ZBS, ZCS, ZUS) du RSMH. A cet effet les CP interrogeront les équipements des

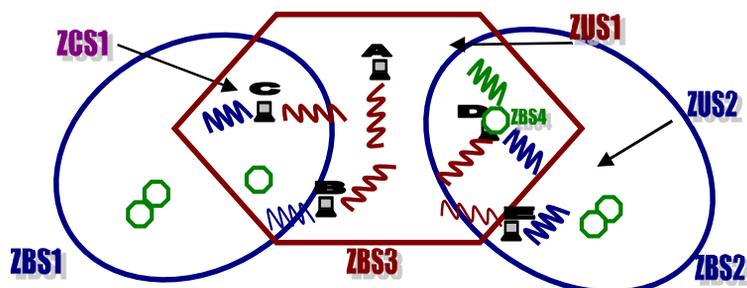


Figure 4.2 Rôle des zones de recouvrement dans la localisation des EU

utilisateurs (EUs) pour leur fournir des données ayant trait aux niveaux des différents signaux que ces derniers équipements reçoivent, à travers les diverses interfaces radio qu'ils sont capables de scruter. Ces données seront comparées aux valeurs des seuils définis pour chaque ZBS, afin d'en déduire les localisations recherchées, des EUs.

Ainsi, et comme indiqué sur la figure 4.2, on peut déduire que la station A est située dans la zone ZBS3, et plus précisément dans ZUS1, du fait que cette station ne reçoit qu'un signal seulement (fort) provenant de la station de base de ZBS3. La station B quant à elle reçoit deux signaux, l'un (fort) détecté sur l'interface avec la zone ZBS3, l'autre (faible) détecté sur l'interface avec la zone ZBS1. Ces informations seront interprétées comme une indication que la station B est située dans la zone ZBS3, près de la limite avec la zone ZBS1, c'est à dire aux frontières de la zone ZCS1. L'équipement utilisateur C est en pleine zone ZCS1 car pouvant communiquer, aussi bien sur l'interface de la zone ZBS1 que sur celle de la zone ZBS3 (c'est ce qui caractérise les zones de chevauchement de plusieurs TAR, comme déjà signalé). La station D se trouve dans une zone (ZBS4) de couverture où se chevauchent trois TAR, lui permettant de communiquer à travers trois interfaces possibles (trois signaux forts). Il faut remarquer que nous avons là un cas où la zone de chevauchement coïncide complètement avec la zone de base. Enfin, la station E est aussi située près d'une limite, car recevant un signal faible de la zone ZBS3, mais se trouve, en contre partie, en plein zone ZBS2, car y recevant un signal fort.

A ce stade de notre développement, nous avons montré qu'en utilisant une technologie qui existe dans les stations mobiles actuelles, il nous est possible de localiser assez précisément les stations dans un RSMH. La mobilité des stations ne change en rien le principe retenu, bien au contraire, elle confirme sa puissance et son efficacité. En effet, en combinant leur connaissance de la "topologie" du RSMH, avec les données qu'ils reçoivent des stations mobiles, les capteurs de contextes sont capables de suivre les itinéraires des mobiles à travers le réseau (figure 4.3), et tenir à jour les contextes de leurs zones de rattachement respectives.

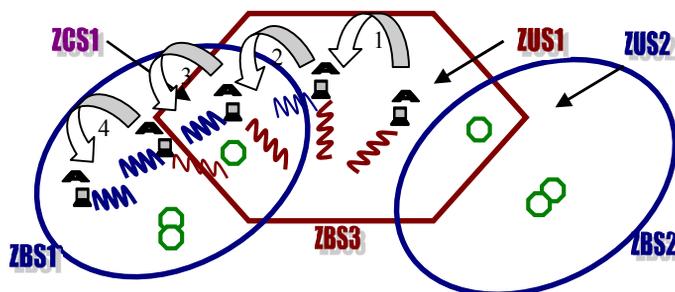


Figure 4.3: suivi du mouvement d'une station mobile en fonction des niveaux des signaux reçus à travers les différentes interfaces.

A la question de savoir à quel rythme les CP devront rafraîchir leurs connaissances, nous considérons, qu'il y a au moins deux approches possibles: le rafraîchissement synchrone et le rafraîchissement asynchrone. Dans l'approche synchrone c'est au réseau de prendre l'initiative, de scruter périodiquement les mobiles, à la recherche de stations ayant changé de localisation (sur la base des niveaux des signaux que ces dernières transmettraient). Dans une approche asynchrone, c'est les stations qui prennent l'initiative de signaler au réseau les changements des niveaux des signaux qu'elles reçoivent des différentes TAR (indiquant par là, un changement de localisation).

La figure 4.4 illustre le résultat escompté de l'activité des capteurs de contextes, et de leurs interactions avec les divers équipements des utilisateurs. Les différents traits et formes représentent les zones remarquables de répartition des stations dans le RSMH. Les CPs sont aussi chargés de fournir des données chiffrées en rapport avec ces zones, constituant ainsi une sorte de base de données, à laquelle nous recourons à diverses occasions tout au long de nos développements ultérieurs.

Prévoir l'importance des mouvements des stations (et leurs sens) vers les diverses zones, connaître le peuplement de celles-ci, particulièrement les ZCS d'entre elles, estimer le volume et le délais des besoins en ressources radio (ou en ressources libérables) dans chaque ZBS, déterminer les stations sur le point

d'effectuer des HO, compter ces HO, et effectuer une compensation entre eux, enregistrer leurs zones de premier accès et garder un œil sur certaines stations durant leur traversée de plusieurs frontières pour signaler la fin de leurs sessions de communication, déterminer l'état de charge de chaque ZBS et celui des interférences ou des puissances d'émissions pratiquées, sont autant de préoccupations auxquelles répond la disponibilité d'une BD telle que celle construite par les CPs.

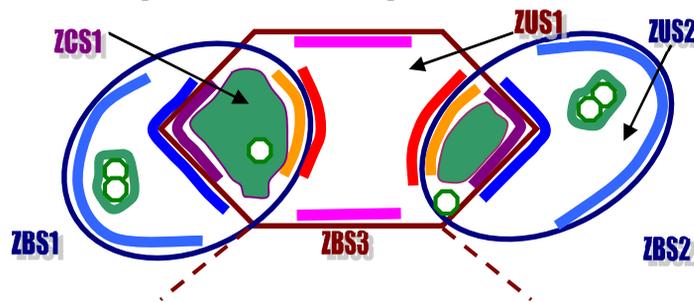


Figure 4.4 : Zones remarquables de répartition des stations des utilisateurs dans le RSMH.

4.3 Schéma proposé pour la redistribution des RR

4.3.1 Principe

Dans ce qui suit nous allons décrire le scénario selon lequel les diverses zones, mises en évidence sur la figure 4.4, sont exploitées pour la GRR dans le RSMH. Tout d'abord, il faut attirer l'attention sur le statut particulier des zones de chevauchement de couverture sans fil (ZCS), dans le réseau hétérogène. En effet, on admet que la présence de plusieurs technologies d'accès radio (TAR) différentes dans une même région dénote une importance fondée et partagée, accordée par des opérateurs différents (ou le même opérateur) à la couverture de celle-ci. Il est donc justifié de s'attendre à ce que l'activité dans ces zones soit importante, sinon intense, et de les ériger en zones sources (et puits), de ressources radio. Notre conception de la GRR dans les RSMH adopte ces principes et associe à chaque ZBS une entité (objet, agent ...etc) ayant autant de ports de communication vers les autres ZBS (de TAR différentes) qu'elle a de ZCS différentes la partitionnant. Ainsi, et comme illustré sur figure 4.4, les entités rattachées aux zones ZBS1, ZBS2 et ZBS3 auront respectivement 3,4, et 3 ports de communication. Il est à noter aussi, que chaque entité ainsi définie aura au plus un port de communication vers chacune des autres entités des autres ZBS.

Les leviers de notre approche pour la Gestion des ressources radio dans un RSMH sont: l'adoption du handover vertical en tant qu'outil de base, l'utilisation des zones de chevauchement de plusieurs TARs comme sources (et puits) des ressources radio, ainsi que l'introduction du concept de "capteurs de contexte". La GRR elle-même fonctionne selon les principes décrits ci-après :

L'idée est de mobiliser à tout instant le volume de ressources radio laissées "oisives", dans chaque ZBS, suite aux fluctuations dans les degrés d'utilisation de ces ressources au cours du temps, ainsi que les décalages temporels qui caractérisent ces fluctuations dans les diverses ZBS du réseau. Chacune des ZBS du réseau essaye de réserver (ou dégager) un certain volume de ressources radio qu'elle partage, périodiquement, avec les ZBS des autres TAR avec lesquelles elle a des ZCS communes. Tant que la ZBS en question n'est pas surchargée, elle participe activement à pourvoir les autres ZBS en ressources dont elles ont besoin, à diverses fins. Dès que le taux d'utilisation des RR dans la ZBS pourvoyeuse atteint un certain seuil (à définir) un processus de recouvrement de la bande de ressources pourvue est déclenché, si ces ressources ne sont pas déjà reconstituées. De même, les zones qui absorbent l'excès de ressources des zones pourvoyeuses ne doivent pas le faire au-delà d'un certain seuil, sous peine de freiner le processus. Le but étant, que chaque zone impliquée dans ce scénario offre ce qu'elle a de ressources radio libres à qui voudrait bien en profiter, en exigeant en contre partie de pouvoir reconstituer au besoin, sa réserve de ressources mobilisées (peut être même plus). Il importe peu que les zones sources, de reconstitution des réserves d'une zone, soient celles qui ont été pourvues par cette dernière. Notre souci principal est de "traquer" toutes les ressources radio momentanément libres, de certaines TAR, dans l'éventualité de les

exploiter pour secourir les autres TARs. Bien entendu, ce scénario se déroule dans tout le réseau hétérogène, où il est caractérisé par une certaine localité. En effet, nous supposons que dans chaque ensemble de ZBS se partageant des ZCS communes (région) certaines de ces zones sont des pourvoyeuses de ressources radio tandis que d'autres en sont des demandeuses (le cas de forte surcharge du RSMH sera traité plus loin). Bien qu'elle s'apparente à de l'équilibrage de charge, ce n'est pas le souci premier de notre approche. Notre souci est de pouvoir considérer qu'on est souvent en présence d'une source de RR quasi permanente, à gérer. C'est-à-dire que nous prôtons une politique agressive de réutilisation des ressources radio disponibles dans le RSMH. Dans ce cadre, une zone pourrait avoir besoin de ressources radio pour plusieurs raisons dont nous citons ci-après quelques unes, à titre d'exemples :

- Répondre à un besoin de "Load Balacing" si son propre réseau particulier et les mécanismes y afférant dans celui-ci se révèlent insuffisants,
- La zone elle-même est surchargée et a besoin, par conséquent d'être aérée,
- Se doter d'une capacité virtuelle, supérieure à la sienne réelle. Le souci ici n'est plus de se décharger mais, de libérer plus de ressources pour être en mesure de prendre en charge plus d'appels,
- Améliorer la QoS en ramenant les degrés d'interférences, les puissances d'émission, les taux d'erreurs, ...etc, à des valeurs plus favorables, ou bien, en réduisant artificiellement la zone de couverture,
- Reconstituer sa réserve de ressources mobilisées auparavant. Cet aspect est vital, comme déjà signalé, car on n'accepte de partager ses ressources que si on est sûr de la solidarité des autres dans les moments de "disette".

Pour décrire l'activité du réseau hétérogène, on pourrait dégager trois situations limites, représentatives du champ des états du système. La première des ces situations est caractéristique du fonctionnement "normal", dans lequel le système peut être partitionné en zones sous chargées (underloaded) et zones surchargées (overloaded). La deuxième situation est celle où un ensemble de zones sont en possession d'un volume de ressources libres qu'elles pourraient partager, mais sous réserve de pouvoir en disposer à plus ou moins court terme, alors que d'autres zones ont en besoin immédiatement, escomptant en contre partie, une libération de ressources imminente mais non à temps. La troisième situation enfin, est caractéristique des états de fortes surcharges, où la majorité des zones ont des marges très réduites de manœuvre en termes de partage de ressources. Dans pareils cas, chaque zone ne consent à pourvoir que si elle est en retour pourvue en ressources radio; c'est en somme, une sorte de troc qui s'établi. De ce qui précède nous pouvons identifier les états possibles de chaque zone du réseau, que ce soit dans son statut de pourvoyeuse de ressources ou de celui de consommatrice (la table 4.1 résume ces états).

Zone ZBS pourvoyeuse	Excès de ressources libres mobilisables plus ou moins à moyen terme	Une marge de ressources libres mobilisables à court terme	Une marge de ressources libres échangeables seulement (troc)
Zone ZBS déficitaire	Déficit de ressources libérables plus ou moins à moyen terme	Déficit de ressources libérables à court terme	déficit de ressources échangeables seulement (troc)

Table 4.1 : Etats possibles des ZBS

L'état global du réseau pourrait donc être décrit par une combinaison des états des zones repris dans la table 4.1. Le comportement de chacune des ZBS sera, bien sûr, dicté par son état ainsi que par les états des ZBS formant avec elle une région. Le transfert des ressources s'effectue grâce au mécanisme du handover vertical (HOV). Ainsi, une zone pourvoit une autre en réservant des ressources à des stations, se trouvant dans la ZCS commune, éligibles à un HOV dont le déclenchement réalisera le transfert des RR proprement dit,. Le sens du HOV détermine le sens inverse du mouvement des ressources.

4.3.2 Problématique de la mise en œuvre du schéma

La présente section est le lieu où va être présenté notre vision du schéma de mise en œuvre des principes qui viennent d'être décrits. La problématique à laquelle on se trouve confronté dans pareille mise en œuvre, se pose en des termes que le commentaire qui suit, celui de l'exemple illustré sur la figure 4.5, tente de cerner. Tout d'abord, des précisions préliminaires s'imposent, pour la clarté de l'exposé. La figure 4.5 considère une région d'un RSMH dont deux zones de base, ZBS2 et ZBS3, ont respectivement un besoin de ressources de 22 et 15 unités. Les autres zones de base, de la région, en état de participer à pourvoir à ces besoins, sont aussi représentées, mention étant faite pour chacune du nombre de ressources qu'elle est en mesure de mettre en jeu. La figure 4.5 indique que la zone ZBS1, par exemple, dispose de 9 unités qu'elle doit répartir entre les zones ZBS2 et ZBS3.

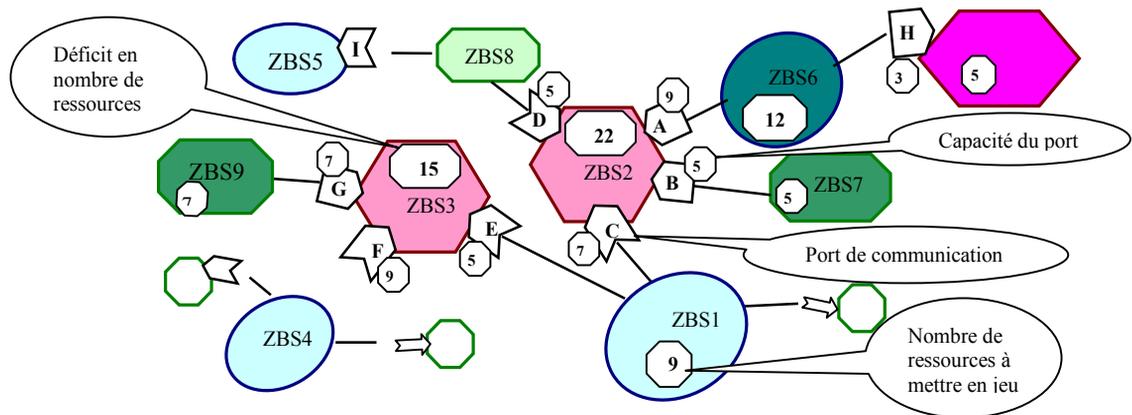


Figure 4.5.Exemple d'un modèle de RSHM

L'existence de zones de chevauchement, entre les zones pourvoyeuses de ressources et celles à pourvoir, est indiquée sur la figure par des liens (A,B,C,...) représentant les ports de communication entre ces dernières. A chacun de ces ports de communication est associé un nombre, qui traduit la capacité de la zone de chevauchement, c.a.d le nombre maximum de ressources que la zone dans le besoin, relative à la zone de chevauchement, puisse espérer obtenir, en l'état actuel des choses. Cette limite est dictée par le nombre de stations présentes dans la ZCS, et connectées à la zone déficitaire. A la lumière de ce qui précède on peut lire sur la figure 4.5 que la zone ZBS1 dégage un excédent de 9 unités, qu'elle renferme trois ZCS avec trois ZBS différentes dont deux, ZBS2 et ZBS3, expriment des besoins respectifs de 21 et 15 unités. On peut lire aussi, qu'indépendamment du nombre de ressources libres dont dispose ZBS1, elle ne doit fournir, dans la situation actuelle, à ZBS2 et ZBS3 qu'au maximum 7 et 4 unités respectivement (capacités des ports C et E). D'un autre côté, on constate que la zone ZBS2 peut s'adresser à quatre ZBS différentes, à travers les ports A,B,C et D, pour combler son déficit de 22 unités. On distingue aussi deux types de ZBS pourvoyeuses, selon que leur apport, en nombre, est certain ou incertain. Dans notre cas l'apport des ZBS rattachées aux ports A et B est certain, tandis que celui des ZBS rattachées aux ports C et D est incertain. Cela provient du fait que dans le premier cas, le nombre de ressources libres dont dispose la ZBS est supérieur (ou égal) à la somme des capacités des ports de communication à travers lesquels des demandes de ressources sont enregistrées, alors que dans le second cas, l'inéquation est inversée.

De manière symétrique, deux types de ZBS déficitaires sont à distinguer, selon que leur déficit en nombre de ressources a ou non des chances d'être totalement comblé. Ainsi, un déficit supérieur à la somme des capacités des ports de communication avec les zones pourvoyeuses ne pourra être que partiellement comblé, contrairement au cas où ce déficit serait inférieur à la somme de ces capacités.

Ces précisions préliminaires étant apportées, nous pouvons maintenant aborder la problématique de mise en œuvre, dont il est question de dans cette section. Essentiellement, il s'agit d'une part, pour une zone déficitaire telle que ZBS2, de décider d'une bonne répartition de la requête de 22 unités parmi les zones ZBS1,ZBS6,ZBS7,ZBS8, et d'autre part, pour une zone pourvoyeuse telle que ZBS1, de décider d'une bonne répartition de son apport de 9 unités parmi les zones ZBS2,ZBS3. Par bonne répartition, nous entendons une répartition équitable, assurant le plus d'apport, ou de contribution, possibles pour chaque ZBS, tout en garantissant une sorte d'équilibre de pareto. Dans le présent contexte, un équilibre de pareto

désigne une répartition telle, qu'il n'existe pas d'autres répartitions possibles dont résulterait une amélioration dans l'affectation pour une ou plusieurs ZBS qui ne soit pas aux dépends d'autres ZBS du RSMH.

La situation représentée sur la figure 4.5 va nous donner matière à développer ce que nous avons résumé. Ainsi, un scénario possible pourrait être celui où les zones ZBS2 et ZBS3 choisiraient d'émettre, à destinations des zones qui leurs sont connectées à travers les ports A,B,C,D et E,F,G respectivement, les requêtes suivantes:

$$\begin{aligned} & \text{requête}_{ZBS2}([A:9],[B:5],[C:7],[D:5]) \text{ et} \\ & \text{requête}_{ZBS3}([E:9],[F:5],[G:7]) \end{aligned}$$

Il est aisé de remarquer que le nombre total d'unités demandé dans chacune des requêtes (26 et 21) dépasse celui du besoin exprimé (22 et 15). Cela provient du fait que, dans leurs quête du maximum possible d'unités, les zones ne doivent pas perdre de vue l'incertitude qui caractérise, généralement, les apports escomptés; d'où le principe retenu (pour les zones déficitaires) de toujours relever le plafond des requêtes à hauteur des capacités des ports. Une conséquence immédiate de ce choix est, qu'en l'absence de concertation entre les zones pourvoyeuses, la distribution des ressources générera vraisemblablement des situations de déséquilibre qu'il faudra rétablir. Dans notre exemple, un scénario produisant pareille situation pourrait être celui représenté sur la figure 4.6, où les zones ZBS2 et ZBS3 reçoivent des ports A,B,D et F,G respectivement les apports : [A:9],[B:5],[D:4] et [F:3],[G:7] et où la zone ZBS1 répond aux requêtes reçues par l'apport $\text{app}_{ZBS1}([C:7],[E:2])$.

$\text{requête}_{ZBS2}([A:9],[B:5],[C:7],[D:5])$	$\text{apport}_{ZBS6}([A:9],[H:3]),$
	$\text{apport}_{ZBS7}([B:5]),$
	$\text{apport}_{ZBS8}([D:4],[I:1]),$
	$\text{apport}_{ZBS1}([C:7],[E:2]),$
$\text{requête}_{ZBS3}([E:9],[F:5],[G:7])$	$\text{apport}_{ZBS4}([F:3]),$
	$\text{apport}_{ZBS9}([G:7]),$

Figure 4.6. Spécification des messages transmis entre entités du réseau.

En effet, au terme du premier cycle (requête/apport), ZBS2 aura acquis 25 unités, contre 12 unités acquises par ZBS3. Le déséquilibre est clair puisque ZBS2 reçoit un surplus de 3 unités alors que ZBS3 à encore un manque de 3 unités par rapport aux besoins respectifs exprimés. Cette distribution doit être corrigée pour, d'une part, ce qu'elle est, c'est-à-dire un déséquilibre, et d'autre part, pour ce qu'on doit préserver les capacités d'apports des zones, en n'attribuant que le strict nécessaire. A cet effet, d'autres cycles encore devront être enclenchés entraînant une redistribution, des surplus dégagés, vers les zones pourvoyeuses. Mais ce retour des apports devra être pensé, de manière à le diriger vers les zones où ces ressources seront les plus utiles. Ce serait, peut être, une mauvaise idée si (**Cas 1**) nous décidions, dans notre exemple, de retourner le surplus dans ZBS2 (de 3 unités) vers le port A, car de la sorte nous aurions seulement déplacé le surplus sans réduire le déficit (celui de ZBS3). Par contre (**Cas 2**), retourner le même surplus vers ZBS1 pourrait servir à combler le déficit affiché dans ZBS3 (dans un cycle ultérieur). Dans un cas comme dans un autre, la qualité de la décision dépend des décisions prises, en même temps, dans les autres zones. Ainsi, si ultérieurement (**Cas 3**), ZBS4 venait à combler le déficit de ZBS3, la décision considérée au premier cas ne paraît plus aussi mauvaise, par contre, la décision de retourner les 3 unités à ZBS1 n'apparaît plus aussi bonne, puisque qu'elle est de nature à générer un autre surplus, dans ZBS3 cette fois-ci, qu'il faudra à nouveau corriger à travers d'autres cycles encore. L'occurrence des déficits pourrait être telle, qu'un nombre important de cycles est nécessaire pour aboutir à un équilibre de pareto. En fait, le nombre de cycles n'est pas le seul facteur qui augmente, la complexité du "comportement" des zones augmente aussi. Pour s'en apercevoir, il faut remarquer le changement de "comportement" de ZBS1 (et ZBS2) entre le premier cycle (requête/apport), où elle devait attribuer des ressources (recevoir des ressources pour ZBS2) à des zones qui lui étaient "directement" connectées, le cycle où il était question de jouer le rôle d'intermédiaire en récupérant des ressources de ZBS2 pour les

attribuer à ZBS3 et, les cycles où ZBS1 (et ZBS2) pourrait constituer un maillon d'une chaîne d'intermédiaires acheminant le surplus de ressources d'une source vers un puit. C'est à une situation similaire que nous pourrions aboutir si, partant du cas précédent (Cas 3) avec surplus retourné à ZBS1), nous supposons que ZBS5 enregistre aussi un déficit. Il s'agirait alors, dans ce cas de figure, d'acheminer le surplus de ZBS3 vers ZBS5, par l'entremise de ZBS1 et ZBS2, au prix plusieurs cycles (requête/apport).

Ainsi donc, il apparaît, suite au développement qui précède, que fixer aux entités du RSMH des "comportements" permettant une redistribution coopérative des ressources radio, en l'absence d'informations mutuelles suffisantes, n'est pas une tâche évidente. En effet, en l'absence d'autres informations que leurs profils, un ZBS en déficit, telle que ZBS1 (voir figure 4.7), doit quand même décider d'une part, des zones pourvoyeuses qu'elle estime à même de contribuer à combler son déficit, et d'autre part, des niveaux des contributions dont elle les sollicitera (dans les limites des capacités respectives des ZCS associées). Elle pourrait choisir, dans une première option, de miser sur certains ports

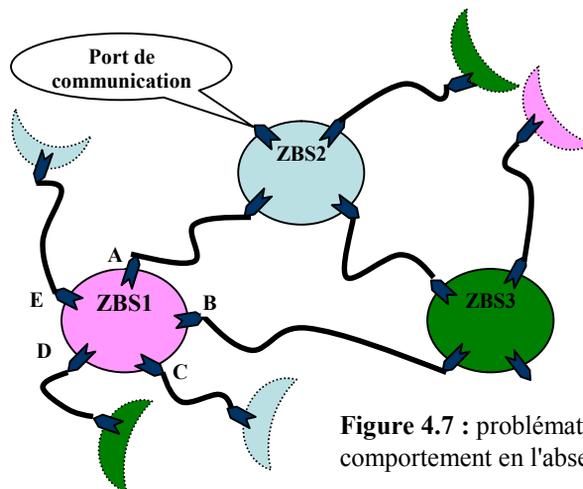


Figure 4.7 : problématique de définition d'un comportement en l'absence d'informations mutuelles.

plutôt que d'autres pour l'extinction de son déficit, par exemple, placer des niveaux de contributions à demander relativement élevés sur les ports A et B par rapport au reste des ports (C,D,E). Le taux de réussite aléatoire mis à part, ce choix risque, en cas d'échec (apports récoltés en deca des espoirs), de laisser se dissiper des ressources qui étaient disponibles sur les ports négligés par la mise. Une autre tentative est toujours possible (lors d'un deuxième cycle) pour obtenir plus de ressources, mais en période de disette il y a de fortes chances qu'au premier cycle déjà le plus gros soit distribué à d'autres zones. Dans une deuxième option, la zone pourrait bien demander le maximum de ressources possible sur chacun des ports, espérant ainsi en recueillir assez pour résorber son déficit. En cas de réussite cette option nous ramène au cas de l'accumulation de surplus dans des zones en contre partie de la persistance de besoins dans d'autres zones (cas sus discuté). Une troisième option, semblable à la première par ses choix mais différente dans son résultat, est aussi possible. Le fait est que cette alternative illustre un type de "comportements" en désaccord avec le principe de la coopération, puisqu'elle est de nature à générer des déséquilibres nuisibles à la bonne répartition recherchée. Cela provient de ce qu'on a, peut être, misé sur les ports où il y a une forte concurrence, et relativement peu de ressources, en négligeant ceux où il y a assez de ressources. Du coup nous privons les autres zones des ressources qu'on a reçu des ces ports alors qu'on aurait pu les obtenir autrement.

De manière symétrique, et dans les mêmes conditions, une zone en possession d'un excédent de ressources, telle que ZBS2 (voir figure 4.7), doit aussi décider, d'une part, des zones déficitaires qu'elle estime mériter le plus son concours, et d'autre part, des niveaux des dits concours. Comme pour son homologue des zones dans le besoin, plusieurs options s'offrent à notre zone en question. Selon une première alternative, elle pourrait miser sur certains ports au détriment d'autres en matière d'apports en ressources. La question qui se poserait alors serait la suivante : "n'a-t-on pas secouru des zones ayant recueilli assez de ressources autrement qu'à travers notre apport, et laissé d'autres, dont le seul recours de l'instant, ou presque, était notre zone ?". Dans l'affirmative nous aurons alors agi en désaccord avec le

principe de coopération nécessaire dans pareil contexte qu'est le notre. Une variante de la précédente option est celle où l'apport générerait des surplus dans les zones des ports soutenus en contre partie de déficits non résorbés dans les zones des ports faiblement soutenus (cas sus discuté). Une autre possibilité serait d'envisager le cas où, par un engagement en déca de ce qui est possible, sur certains ports, la zone oblige les zones connectées à ces derniers d'entamer les ressources d'autres zones, très sollicitées, alors qu'un engagement plus important aurait beaucoup amélioré la situation.

Dans les deux paragraphes précédents nous avons, à dessein, omis d'analyser le fonctionnement conjoint des deux zones de profils différents, tout comme nous avons évité de considérer l'autre option, qu'est celle adoptée dans notre travail, en tant qu'alternative de celles examinées. La raison en est qu'il s'agissait avant tout d'examiner la problématique liée à la mise en œuvre du schéma de redistribution, et de mettre en évidence la nécessité de concevoir un mécanisme de distribution des ressources qui obéit aux critères d'équité, du plus d'apport ou de contribution possibles, d'équilibre de pareto, de flexibilité, d'extensibilité et d'efficacité, dans le contexte de la gestion des ressources radio, dans réseau sans fil et mobile hétérogène, tel que défini dans la présente section. Ce mécanisme sera à la base du processus de mise en œuvre des principes retenus en la matière (GRR dans un RSMH).

4.3.3 Schéma de base pour la GRR

Le principe sous-tendant notre approche en matière de GRR, ainsi que la problématique relative à la mise en œuvre de son schéma de base étant présentés, nous pouvons à présent, décrire proprement le schéma de mise en œuvre. Pour la clarté de l'exposé nous allons passer en revue deux versions de ce schéma, l'une préliminaire, l'autre "finale". La notation adoptée utilise les éléments de la figure 4.8.

- $i \in \{1, 2, \dots, N\}$ N : nombre de technologies d'accès radio (*TAR*) présentes dans le *RSMH*,
- Z_{ij} : zone de base de couverture sans fil (*ZBS*) numéro j de la *TAR* i ,
- Z_{ij}^k : port de communication k de la zone Z_{ij} ,
- c_{ij}^k : capacité du port Z_{ij}^k ,
- r_{ij}^k : nombre d'unités de ressources demandées sur le port Z_{ij}^k ,
- a_{ij}^k : nombre d'unités de ressources reçues (ou notifiées en tant qu'apport) sur le port Z_{ij}^k ,
- $Ports(Z_{ij})$: ensemble des ports de communication de la *ZBS* Z_{ij} ,
- $Déficit(Z_{ij})$: déficit en nombre d'unités de ressources de la *ZBS* Z_{ij} ,
- $Excès(Z_{ij})$: excédent en nombre d'unités de ressources de la *ZBS* Z_{ij} ,
- $Résidu(K, R, expr)$: ensemble $\{i_1, i_2, \dots, i_R\}$ tel que $i_k \in \{1, 2, \dots, K\}$ et $expr = Vrai$,
- $Profil(Z_{ij})$: profil de la *ZBS* Z_{ij} (*Excédent*, *Déficit*, *Port_Virtuel*),
- $Etat_{pori}[Z_{ij}^k]$: profil de la *ZBS* connectée à Z_{ij} via Z_{ij}^k ,
- $Voisinage(Z_{ij})$: ensemble des *ZBS* connectées à Z_{ij} .

Figure 4.8 Notation adoptée dans la description du schéma proposé

4.3.3.1 Version Préliminaire

Le processus se déclenche périodiquement, au rythme d'un cycle principal et de cycles secondaires (figure 4.9), et commence par une phase d'initialisation au cours de laquelle chaque *ZBS* (Z_{ij}) s'adresse à son capteur de contexte local pour lui fournir un état du contexte (figure 4.13). Cet état reprend des informations importantes relatives à la situation en terme de ressources et aux capacités des *ZCS* de la *ZBS* en question. A la sortie de cette phase, on doit être fixé sur l'état des ressources (déficit, excédent, ou équilibre) de la zone Z_{ij} , pour décider de son comportement ultérieur. Si la zone Z_{ij} décide de participer au processus de redistribution des ressources, elle le fait savoir en communiquant son état aux autres *ZBS* qui lui sont reliée par des ports de communication (zones avec lesquelles elle partage des *ZCS* communes). De son côté, elle reste à l'écoute de ces mêmes ports pour détecter lesquelles, parmi les zones

voisines, sont aussi engagées dans le processus, et avec quels profils. La suite du comportement de chaque zone dépend de son profil.

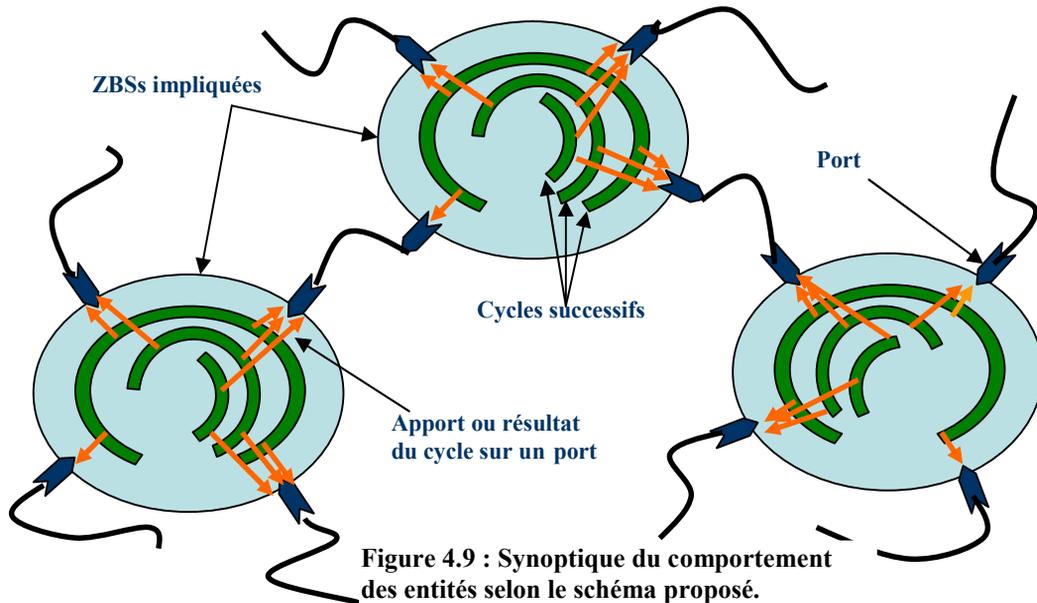


Figure 4.9 : Synoptique du comportement des entités selon le schéma proposé.

Zone en déficit

Ainsi, une zone $Z_{i,j}$ en déficit de ressources procèdera d'abord à la détermination des ports de $Pourv(Z_{i,j})$, ainsi que leurs capacités, débouchant sur des zones en excédent de ressources. Cet ensemble ($Pourv(Z_{i,j})$) permettra à $Z_{i,j}$ de s'adresser aux zones susceptibles de la pourvoir en ressources. Les capacités des ports ($c^k_{i,j}$), quant à elles, indiquent les limites des nombres de ressources que $Z_{i,j}$ est capable de recevoir à travers chacun des ports de ($Pourv(Z_{i,j})$). En fait, combinées à $Déficit(Z_{i,j})$, ces capacités vont permettre de calculer le nombre de ressources ($R^T_{i,j}$) que $Z_{i,j}$ va effectivement demander aux zones pourvoyeuses. Le calcul de $R^T_{i,j}$ traduit simplement le fait, qu'on ne peut espérer obtenir plus de ressources que ne le permettent les capacités des ports, d'où la formule :

$$R^T_{i,j} = \min(Déficit(Z_{i,j}), \sum c^k_{i,j}) : Z^k_{i,j} \in Pourv(Z_{i,j})$$

Fonction Répartir (Ress: real, Prt : ensemble de ports): ensemble d'entiers $[x^k_{i,j}]$

1. **Var**
2. m, R : real ;
3. **begin**
4. $m:=0 ; R:=0;$
5. $m := Ress / Card(\{Z^k_{i,j} \in Prt : indic^k_{i,j} = true \}) ;$
6. $x^k_{i,j} := \min(m, c^k_{i,j}) : Z^k_{i,j} \in Prt \text{ and } indic^k_{i,j} = true ;$
7. $R := R + c^k_{i,j} : Z^k_{i,j} \in Prt \text{ and } x^k_{i,j} = c^k_{i,j} \text{ and } indic^k_{i,j} = true ;$
8. $indic^k_{i,j} := false : Z^k_{i,j} \in Prt \text{ and } x^k_{i,j} = c^k_{i,j} \text{ and } indic^k_{i,j} = true ;$
9. **if exist**($indic^k_{i,j} = true : Z^k_{i,j} \in Prt$) and $R > 0$ **then**
10. **Répartir** ($Ress-R, Prt$) ;
11. **Ajuster** ($x^k_{i,j} : Z^k_{i,j} \in Prt$) ; / répartir le résidu des ressources /
12. **retrun** [$x^k_{i,j}$] : $Z^k_{i,j} \in Prt$;
13. **end**

Figure 4.10 Algorithme de la fonction de répartition des requêtes.

Une fois $R^T_{i,j}$ calculé, il s'agira alors de répartir équitablement sa valeur entre les différents ports de $Pourv(Z_{i,j})$, en tenant compte, toutefois, des limites imposées par les capacités ($c^k_{i,j}$) de ces derniers. Cette répartition est réalisée suivant la logique décrite par le fonction "Répartir" sur la figure 4.10, selon le

modèle : *commande* : *spécification(s)*, c'est à dire que *commande* est exécutée pour les ports spécifiés par *spécification(s)*. Il faut signaler, que cette logique de répartition reflète, en fait, notre vision de l'alternative "comportementale" évoquée dans la section précédente, et qui se veut une réponse aux préoccupations soulevées à cet endroit. Deux principes fondamentaux guident cette réponse; l'un a trait à la symétrie des "comportements" des zones de profils différents, l'autre au partage "équitable". Le premier de ces deux principes est traduit par l'utilisation d'une même fonction sous des appellations différentes ("*Répartir*" et "*Allouer*") dans la description des "comportements" des entités associées aussi bien aux zones déficitaires que celles associées à des zones ayant un excès (Figure 4.13). De la sorte, le très peu d'informations que chaque entité, d'un profil donné, a sur ses homologues, de l'autre profil, s'enrichi de sa prévision de leurs réactions lorsque, ensemble, elles sont engagées dans un processus de redistribution. Un exemple de l'utilisation d'une telle connaissance est donné par la phase de mise à jour (abordée plus loin dans cette section). Le second principe instaure l'égalité dans le traitement des ports en matière de niveaux des contributions sollicitées ou d'apports fournis. L'origine de son adoption est à rechercher dans les critères qui ont structuré le cadre dans lequel notre approche fut élaborée, à savoir :

- Préférer la simplicité tant qu'elle n'altère pas les aspects essentiels de l'approche : ne pas perdre de vue que les "comportements" ainsi que les interactions à concevoir sont celles d'un mécanisme de base (qui sera, par conséquent, intégré à un système multi-agents, un système "policy-based" ou même un système programmé). La simplicité, conditionnée par la non altération des autres aspects essentiels, faciliterait l'intégration et l'adaptation à ces systèmes hôte, et leurs donnerait plus de liberté dans l'utilisation de ce qui leurs est adjoint.
- Rechercher l'équité : dans la section précédente nous avons montré les conséquences des décisions non équitables.
- Surmonter le manque de critères objectifs et calculables de prise de décision: comme déjà signaler, les entités associées aux ZBS du réseau n'ont à la base de leurs prises de décisions que très peu d'informations. Cela rend ardue la définition de critères d'optimisation objectifs et calculables; d'où l'obligation de recourir à des critères de remplacement. Pour notre part l'équilibre de pareto nous semble être un bon compromis.
- Penser la qualité du schéma : les choix de conception retenus, doivent adresser des aspects tels que l'efficacité dans la gestion des ressources, la flexibilité, l'extensibilité, la charge de signalisation nécessaire, ... etc.

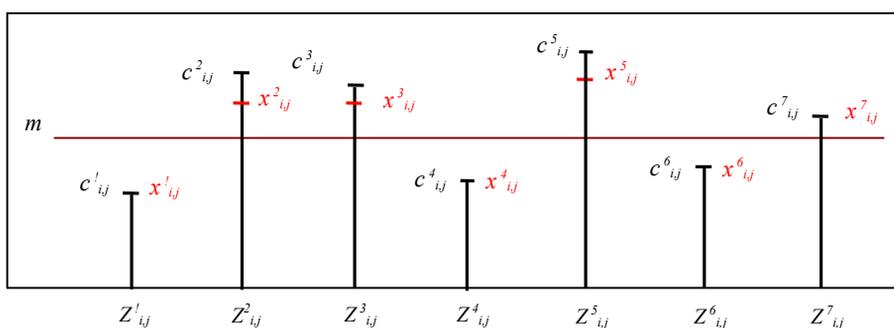


Figure 4.11 Illustration des différentes variables utilisées dans la fonction "Répartir".

Pour reprendre la description de notre schéma, à la phase répartition, signalons que la figure 4.11, illustre les différentes variables utilisées dans la fonction "*Répartir*", ainsi que leurs inter relations.

La phase qui suit (*Envoyer*) construira les *requêtes*, dont les paramètres viennent d'être fixés, puis les enverra à travers les ports correspondants aux zones associées. C'est alors (*Recevoir*) qu'on se mettra dans l'attente des *notifications* de ressources qui parviendraient des zones sollicitées. A ce stade l'exécution pourrait se terminer si tous les ports se fermaient signalant, ainsi, les retraits des zones pourvoyeuses correspondantes de l'opération. Autrement, c'est la phase mise à jour qui prendra le relais et qui effectuera, en préparation d'un éventuel nouveau cycle, les actions suivantes :

- Tenir à jour les niveaux des apports en ressources obtenus (r^k_{ij}) sur les divers ports ainsi que les capacités de ces derniers (c^k_{ij}),
- Eliminer les ports à "potentiel nul" en termes de nouveaux apports. L'état d'un port est déclaré à "potentiel nul" dans l'un des cas suivants :
 1. le port est fermé,
 2. le niveau des apports reçus atteint la capacité du port,
 3. la dernière notification reçue est inférieure à la valeur sollicitée par la requête correspondante.
- Déterminer la nouvelle valeur temporaire du déficit ($(Déficit(Z_{ij}))$) suite aux nouvelles notifications reçues,
- Recalculer la nouvelle valeur de R^T_{ij} , c'est-à-dire, le nombre d'unités qu'on essayera de récolter au prochain cycle.

Il peut arriver qu'on ne puisse pas, en un seul cycle, éponger le déficit et que cela exigera, au contraire, d'autres cycles encore. La poursuite des traitements s'averra nécessaire si les deux conditions suivantes sont satisfaites :

1. $R^T_{ij} \neq 0$: c.a.d, le déficit n'a pas encore été épongé
2. $Pourv(Z_{ij}) \text{ non vide}$: c.a.d, des zones prêtes à apporter leurs contributions existent.

L'arrêt du traitement est décrété, les ports de la zones fermés et le résultat de l'opération renvoyé (cumuls des notifications reçues à travers chacun des ports), si les ressources qu'il était possible d'avoir sont recueillies ou bien si tous les ports sont à "potentiel nul" en termes de nouveaux apports.

Zone pourvoyeuse

Le "comportement" d'une zone Z_{ij} en pourvoyeuse de ressources adopte une attitude presque symétrique à celle du "comportement" de son homologue, de profil opposé, sus décrit. D'abord, il faut évidemment déterminer les ports de $Ports(Z_{ij})$ associés aux zones qui se sont déclarées déficitaires, puis initialiser les notifications de chacun des ports à zéro (0). On se mettra ensuite dans l'attente d'éventuelles requêtes sur l'ensemble des ces ports ($Solli(Z_{ij})$). A moins que ces derniers se soient fermés (cas où le traitement est déclaré terminé), il faudra connaître le nombre de ressources (R^T_{ij}) que la zone Z_{ij} est effectivement en mesure de pourvoir. Le calcul de R^T_{ij} traduit simplement le fait qu'on ne pourvoira jamais plus de ressources qu'il nous est demandé, d'où la formule suivante :

$$R^T_{ij} = \min(Excédent(Z_{ij}), \sum r^k_{ij}) : Z^k_{ij} \in Solli(Z_{ij})$$

Une fois le nombre R^T_{ij} calculé, on procèdera à sa répartition sur l'ensemble des ports de $Solli(Z_{ij})$, dans les limites des valeurs requises (r^k_{ij}) sur ces derniers. Cette répartition est effectuée selon la même logique décrite par la fonction "Répartir", et représentée dans la figure 4.12 par la fonction "Allouer" (cette fonction étant similaire à la fonction "Repartir", déjà assez détaillée, nous ne la développerons pas plus).

A cet instant du déroulement des opérations, les parts (de l'apport) consenties à chacun des ports de $Solli(Z_{ij})$, au titre du cycle en cours, sont bien connues (x^k_{ij}). Les choses continueront leur cours à travers les actions suivantes :

- réserver dans la zone Z_{ij} un nombre de ressources égal à $\sum x^k_{ij}$,
- mettre à jour les cumuls des notifications (a^k_{ij}), pour chacun des ports de $Solli(Z_{ij})$,
- notifier effectivement les parts consenties (x^k_{ij}) à chacun des port de $Solli(Z_{ij})$.

L'effet de ces actions sur le contexte est reflété, à la phase suivante, à travers la mise à jour (diminution) de l'excédent de ressources de la zone courante ($Excédent(Z_{i,j})$), et l'élimination des ports fermés de l'ensemble $Solli(Z_{i,j})$.

```

Fonction Allouer (Ress: real, Prt : ensemble de ports): ensemble d'entiers [  $x^k_{i,j}$  ]
14. Var
15. m , R : real ;
16. begin
17. m:=0 ; R:=0;
18. m := Ress / Card( {  $Z^k_{i,j} \in Prt : indic^k_{i,j} = true$  } ) ;
19.  $x^k_{i,j} := \min(m, r^k_{i,j}) : Z^k_{i,j} \in Prt \text{ and } indic^k_{i,j} = true$  ;
20.  $R := R + r^k_{i,j} : Z^k_{i,j} \in Prt \text{ and } x^k_{i,j} = r^k_{i,j} \text{ and } indic^k_{i,j} = true$  ;
21.  $indic^k_{i,j} := false : Z^k_{i,j} \in Prt \text{ and } x^k_{i,j} = r^k_{i,j} \text{ and } indic^k_{i,j} = true$  ;
22. if exist(  $indic^k_{i,j} = true : Z^k_{i,j} \in Prt$  ) and  $R > 0$  then
23. Répartir (Ress-R, Prt ) ;
24. Ajuster ( $x^k_{i,j} : Z^k_{i,j} \in Prt$ ) ; / répartir le résidu des ressources /
25. retrun [  $x^k_{i,j}$  ] :  $Z^k_{i,j} \in Prt$  ;
26. end

```

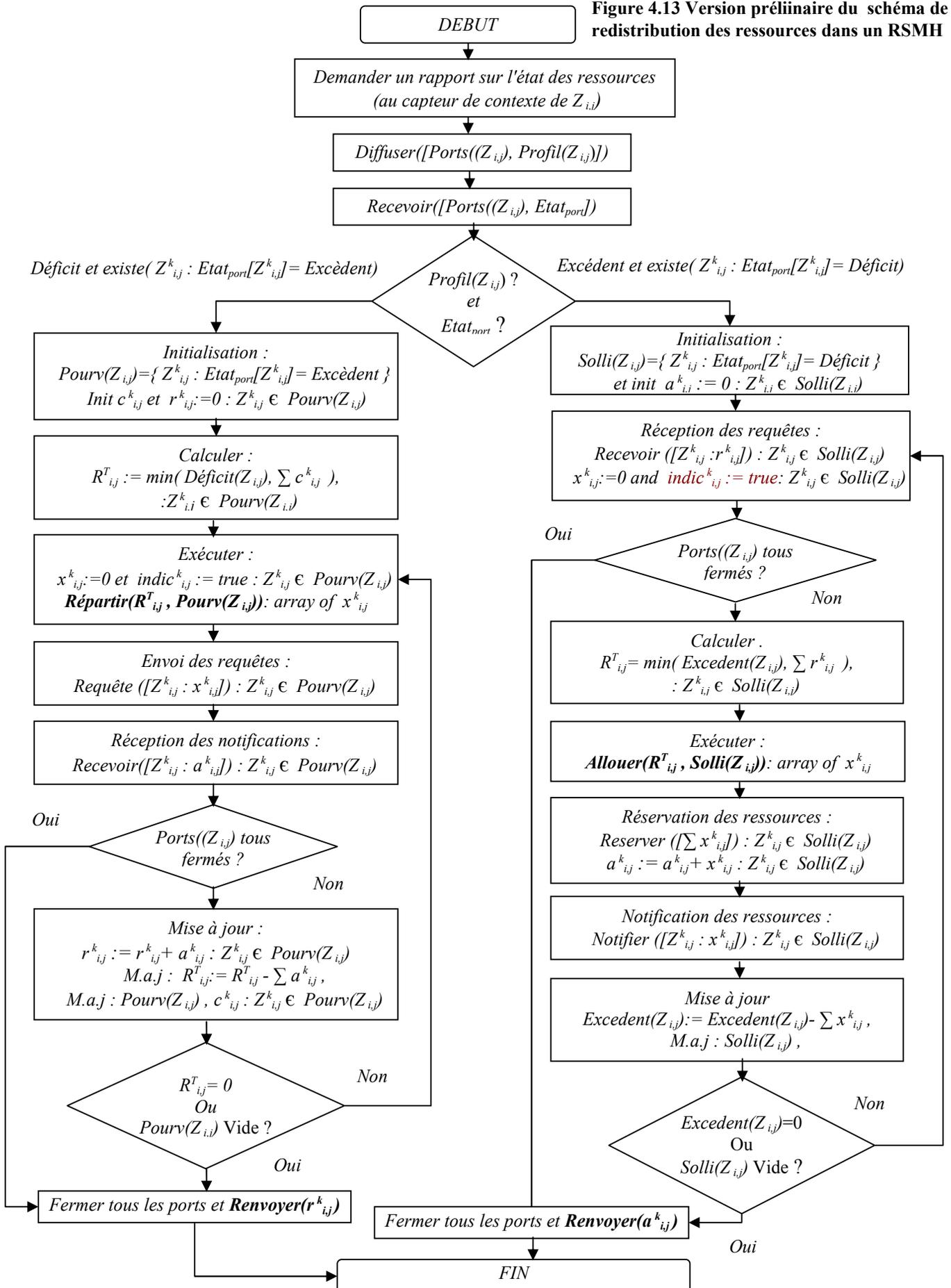
Figure 4.12 Algorithme de la fonction de répartition des apports.

A présent que la fin du cycle est atteinte, faudra t-il le terminer ou, au contraire, d'autres cycles encore sont nécessaires?, car, il n'est pas sûr, en effet, que l'excédent de ressources s'épuise en un seul cycle. La terminaison des traitements s'averra nécessaire si les deux conditions suivantes sont satisfaites :

1. $Excédent(Z_{i,j}) \neq 0$: c.a.d, l'excédent de ressources de la zone ne s'est pas encore épuisé,
2. $Solli(Z_{i,j}) \text{ non vide}$: c.a.d, des zones voisines dans le besoin existent encore.

L'arrêt du processus est décrété, les ports de la zones fermés et le résultat de l'opération renvoyé(cumuls des notifications envoyées à travers chacun des ports), si les ressources qu'il était possible de distribuer ont bien été fournies ou bien si tous les ports se sont fermés.

Figure 4.13 Version préliminaire du schéma de redistribution des ressources dans un RSMH



4.3.3.2 Commentaires sur la version préliminaire du schéma

En termes d'efficacité on peut montrer, d'une part, qu'aucune zone déficitaire ne reçoit plus de ressources qu'elle n'en a sollicité, d'autre part, qu'aucune zone excédentaire ne laisse de ressources sans les distribuer alors que des zones voisines sont encore en déficit, et que leurs capacités sont positives. Cela provient du principe d'équité adopté et qui consiste à augmenter les chances, des zones déficitaires, d'obtenir les ressources nécessaires en s'adressant à toutes les zones excédentaires voisines sans distinction aucune, et à impliquer les zones excédentaires dans l'approvisionnement de toutes les zones déficitaires voisines sans distinction aucune aussi. Les itérations successives permettront d'orienter progressivement les demandes vers les zones à haut "potentiel" tout en préservant l'égalité des chances.

En termes des cycles nécessaires à l'achèvement de la distribution selon le schéma en question, nous pouvons montrer qu'ils dépendent, pour une entité engagée dans le processus de distribution, de son profile et du nombre de ses ports impliqués au démarrage de ce processus. Ces résultats sont obtenus de la manière suivante:

- ❑ Soit R_{ij}^k et n_{ij}^k respectivement le volume de ressources et le nombre de ports impliqués pendant le cycle k dans une ZBS déficitaire Z_{ij} . Il est possible, alors, de représenter l'état de cette zone pendant le cycle k par le couple (R_{ij}^k, n_{ij}^k) .
- ❑ L'évolution de l'état de la ZBS déficitaire Z_{ij} au cours des cycles sera représenté par la suite des états :

$$(R_{ij}^0, n_{ij}^0) \rightarrow (R_{ij}^1, n_{ij}^1) \rightarrow \dots \rightarrow (R_{ij}^k, n_{ij}^k) \rightarrow \dots \rightarrow (R_{ij}^w, n_{ij}^w).$$

Le premier cycle est le cycle 0 et le dernier cycle est le cycle w .

L'analyse de la logique de comportement des entités déficitaires permet de déduire les relations suivantes :

- Si $R_{ij}^k \geq n_{ij}^k \rightarrow R_{ij}^{k+1} < R_{ij}^k$ et $n_{ij}^{k+1} < n_{ij}^k$.
ou
 $\rightarrow R_{ij}^{k+1} = R_{ij}^k$ et $n_{ij}^{k+1} = 0$ (l'entité s'arrete)
- Si $R_{ij}^k < n_{ij}^k \rightarrow R_{ij}^{k+1} < R_{ij}^k$ et $n_{ij}^{k+1} < n_{ij}^k$.
ou
 $\rightarrow R_{ij}^{k+1} = R_{ij}^k$ et $n_{ij}^{k+1} < n_{ij}^k$.

Ces relations mettent en évidence d'une part, le caractère décroissant de la suite des états et d'autre part que, le nombre maximum de cycles w (nombre de cycles(Z_{ij})) est tel que :

$$w \leq n_{ij}^0$$

- ❑ Soient Z_{ij} une ZBS pourvoyeuse et $Solli(Z_{ij})$ l'ensemble des zones déficitaires qui sollicitent Z_{ij}
- ❑ L'évolution de l'état de la ZBS déficitaire Z_{ij} au cours des cycles sera représenté par la suite des états :

$$(R_{ij}^0, n_{ij}^0) \rightarrow (R_{ij}^1, n_{ij}^1) \rightarrow \dots \rightarrow (R_{ij}^k, n_{ij}^k) \rightarrow \dots \rightarrow (R_{ij}^w, n_{ij}^w).$$

- ❑ On montre que le nombre de cycles w est tel que :

$$w = \max(\text{nombre de cycles}(Z_{u,v}) : Z_{u,v} \in Solli(Z_{u,v})).$$

- ❑ On montre aussi que :
si pour une ZBS Z_{ij} déficitaire, $w = n_{ij}^0$
alors au premier cycle déjà le nombre de ressources obtenues est $R_{ij}^0 / (n_{ij}^0 - 1)$

L'examen des conditions d'arrêt du schéma va nous donner une idée précise de la configuration résultante en termes d'états des ressources des zones engagées dans le processus de redistribution. Pour se faire, nous allons commencer par établir quelques assertions caractéristiques de la configuration, des zones engagées, au début et à la fin du déroulement des étapes du schéma (voire figure 4.14)

Début :

- *Déficit Résorbable* ($DefRes_{ij}$) = R_{ij}^T :
 $R_{ij}^T \leq \sum c_{ij}^k : Z_{ij}^k \in Pourv(Z_{ij})$ et $Profil(Z_{ij}) = \text{Déficit}$,
- *Voisinage* (Z_{ij}) = $Pourv(Z_{ij})$
- *Excédent Distribuible* ($ExDist_{ij}$) = $Excedent(Z_{ij})$:
 $Excedent(Z_{ij}) > 0 : Profil(Z_{ij}) = \text{Excédent}$.
- *Voisinage* (Z_{ij}) = $Solli(Z_{ij})$

Après :

- **Si** $\sum r_{ij}^k = DefRes_{ij}$ **Alors** $R_{ij}^T = 0$ et $R_{k,l}^T \geq 0 : Z_{k,l} \in Voisinage(Z_{ij})$
Sinon ($\sum r_{ij}^k < DefRes_{ij}$) $R_{ij}^T < 0$ et $R_{k,l}^T = 0 : Z_{k,l} \in Voisinage(Z_{ij})$ et $c_{ij}^p > 0$
 c_{ij}^p reliant Z_{ij} à $Z_{k,l}$
- **Si** $\sum a_{ij}^k = Excedent(Z_{ij})$ **Alors** $Excedent(Z_{ij}) = 0$ et $R_{k,l}^T \leq 0 : Z_{k,l} \in Voisinage(Z_{ij})$
Sinon ($\sum a_{ij}^k < Excedent(Z_{ij})$) $Excedent(Z_{ij}) > 0$ et $R_{k,l}^T = 0 : Z_{k,l} \in Voisinage(Z_{ij})$
et $c_{k,l}^p > 0$ $c_{k,l}^p$ reliant $Z_{k,l}$ à Z_{ij}

Figure 4.14 Assertions spécifiant l'état de sortie du schéma. Dans sa version préliminaire.

Traduites autrement, ces assertions signifient qu'au terme du processus de distribution, et une fois les ports de capacités nulles désactivés, mis à part ceux reliant des zones d'excédent et de déficit nuls, les états des ZBS sont tels que (voire figure 4.15) :

- Chaque zone encore excédentaire est reliée à des zones dont le déficit est nul,
- Chaque zone encore déficitaire est reliée à des zones dont l'excédent est nul,
- Les zones d'excédent ou de déficit nul viennent en quelque sorte s'intercaler (logiquement) entre les zones des types précédents.

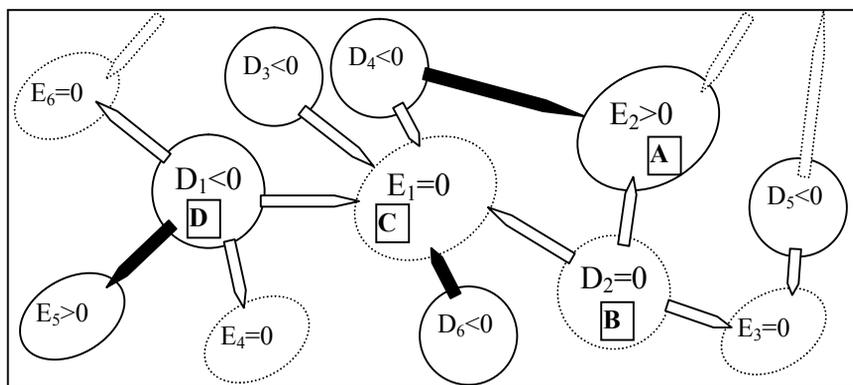


Figure 4.15 : Représentation graphique des états des entités réseau après exécution du schéma.

Cette figure a ceci de particulier, qu'elle met en évidence le fait que la distribution effectuée ne réalise pas un équilibre de pareto. Rappelons qu'un équilibre de pareto désigne, dans notre contexte, une répartition telle, qu'il n'existe pas d'autres répartitions possibles dont résulterait une amélioration dans l'affectation pour une ou plusieurs ZBS, qui ne soit pas aux dépens d'autres ZBS du RSMH. Par "amélioration dans l'affectation" nous entendons une augmentation du nombre de ressources pourvues dans les limites aussi bien des capacités des ports que du déficit enregistré. Par "au dépens de" nous entendons une diminution du nombre de ressources affectées. Il est aisé de remarquer que si la zone A avait donnée plus de ressources à la zone B, celle ci aurait moins exigé de la zone C qui aurait à son tour donné plus de

ressources à la zone D. De la sorte on aurait abouti à une distribution réalisant une amélioration dans les affectations à D sans que cela soit au dépend d'une quelconque autre zone. La version "finale" du schéma tente justement d'exploiter ce genre de configuration pour réaliser un équilibre de pareto. L'idée, est de transformer les zones à excédent (ou déficit) nul, en une sorte de ports virtuels qui relieraient les zones encore déficitaires à celles encore excédentaires. A travers ces ports virtuels des réajustements des affectations initiales seront opérés jusqu'à l'instauration de l'équilibre recherché.

4.3.3.3 Version "finale"

Ce que la version "finale" (figure 4.17) a de particulier par rapport à la version préliminaire est, d'une part, l'existence d'une étape de synchronisation, et d'autre part, la prise en compte du "comportement" relatif au nouveau état désigné par "Port_Virtuel_D" ou "Port_Virtuel_E", selon le profile d'origine de la zone concernée.

Etape de synchronisation

L'étape de synchronisation (figure 4.16) vient terminer le cycle principal et préparer chacun des éventuels cycles qui lui sont postérieurs. Le résultat majeur de cette étape de synchronisation est la reconstruction de nouveaux chemins reliant les zones encore excédentaires à celles encore en déficit, par des ports virtuels interposés. D'abord, les zones dans les états "Excédent" et "Port_Virtuel_E" mettront leurs ports à l'un des deux états "Synchro_E" ou "Synchro_V". Elles attendront ensuite de connaître les nouvelles configurations de leurs voisinages. Ainsi, les zones à état "Excédent" chercherons à déterminer les ports vers les zones, nécessairement "Port_Virtuel_D", qui constitueront les ensembles $Solli(Z_{i,j})$ potentiels.

```

Procedure Synchronisation ( $Z_{i,j}$ )
begin
  Initialiser Profil( $Z_{i,j}$ ) ;
  case Profil( $Z_{i,j}$ ) of

    Excédent :
      Set ( $\text{Etat}_{\text{port}}[Z_{i,j}^k] = \text{Synchro\_E}$ ) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $Z_{i,j}^k$  non Fermé ;
      Wait ( $\text{Etat}_{\text{port}}[Z_{i,j}^k] = \text{Port\_Virtuel}$ ) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $Z_{i,j}^k$  non Fermé ; /  $Solli(Z_{i,j})$  /
      Diffuser (Profil( $Z_{i,j}$ ),  $Z_{i,j}^k$ ) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $\text{Etat}_{\text{port}}[Z_{i,j}^k] = \text{Port\_Virtuel}$  ;

    Port_Virtuel_D :
      Wait ( $\text{Etat}_{\text{port}}[Z_{i,j}^k]$  in ( $\text{Synchro\_E}$ ,  $\text{Synchro\_V}$ )) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $Z_{i,j}^k$  non Fermé ;
      Si non Vide{  $\text{Etat}_{\text{port}}[Z_{i,j}^k] = \text{Synchro\_E}$  :  $Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $Z_{i,j}^k$  non Fermé } Alors /  $\text{Pourv}(Z_{i,j})$  /
        Set ( $\text{Etat}_{\text{port}}[Z_{i,j}^k] = \text{Port\_Virtuel}$ ) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $\text{Etat}_{\text{port}}[Z_{i,j}^k] = \text{Synchro\_E}$  ;
        Recevoir ( $\text{Etat}_{\text{port}}[Z_{i,j}^k] = \text{Excédent}$ ) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $Z_{i,j}^k$  non Fermé ;
        Diffuser ("Excédent",  $Z_{i,j}^k$ ) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$ ,  $\text{Etat}_{\text{port}}[Z_{i,j}^k] \neq \text{Excédent}$  (si non vide) ;
          /  $Solli(Z_{i,j})$  /

    Port_Virtuel_E :
      Set ( $\text{Etat}_{\text{port}}[Z_{i,j}^k] = \text{Synchro\_V}$ ) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $Z_{i,j}^k$  non Fermé ;
      Recevoir ("Excédent",  $Z_{i,j}^k$ ) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $Z_{i,j}^k$  non Fermé ; /  $\text{Pourv}(Z_{i,j})$  /
      Diffuser (Profil( $Z_{i,j}$ ),  $Z_{i,j}^k$ ) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $\text{Etat}_{\text{port}}[Z_{i,j}^k] \neq \text{Excédent}$  ; /  $Solli(Z_{i,j})$  /

    Déficit :
      Wait ( $\text{Etat}_{\text{port}}[Z_{i,j}^k] = \text{Synchro\_V}$ ) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $Z_{i,j}^k$  non Fermé ;
      Recevoir ( $\text{Etat}_{\text{port}}[Z_{i,j}^k] = \text{Excédent}$ ) :  $q.q Z_{i,j}^k \in \text{Voisinage}(Z_{i,j})$  et  $Z_{i,j}^k$  non Fermé ; /  $\text{Pourv}(Z_{i,j})$  /

  end ;
  Set  $\text{Pourv}(Z_{i,j})$  et  $Solli(Z_{i,j})$  ;
end .

```

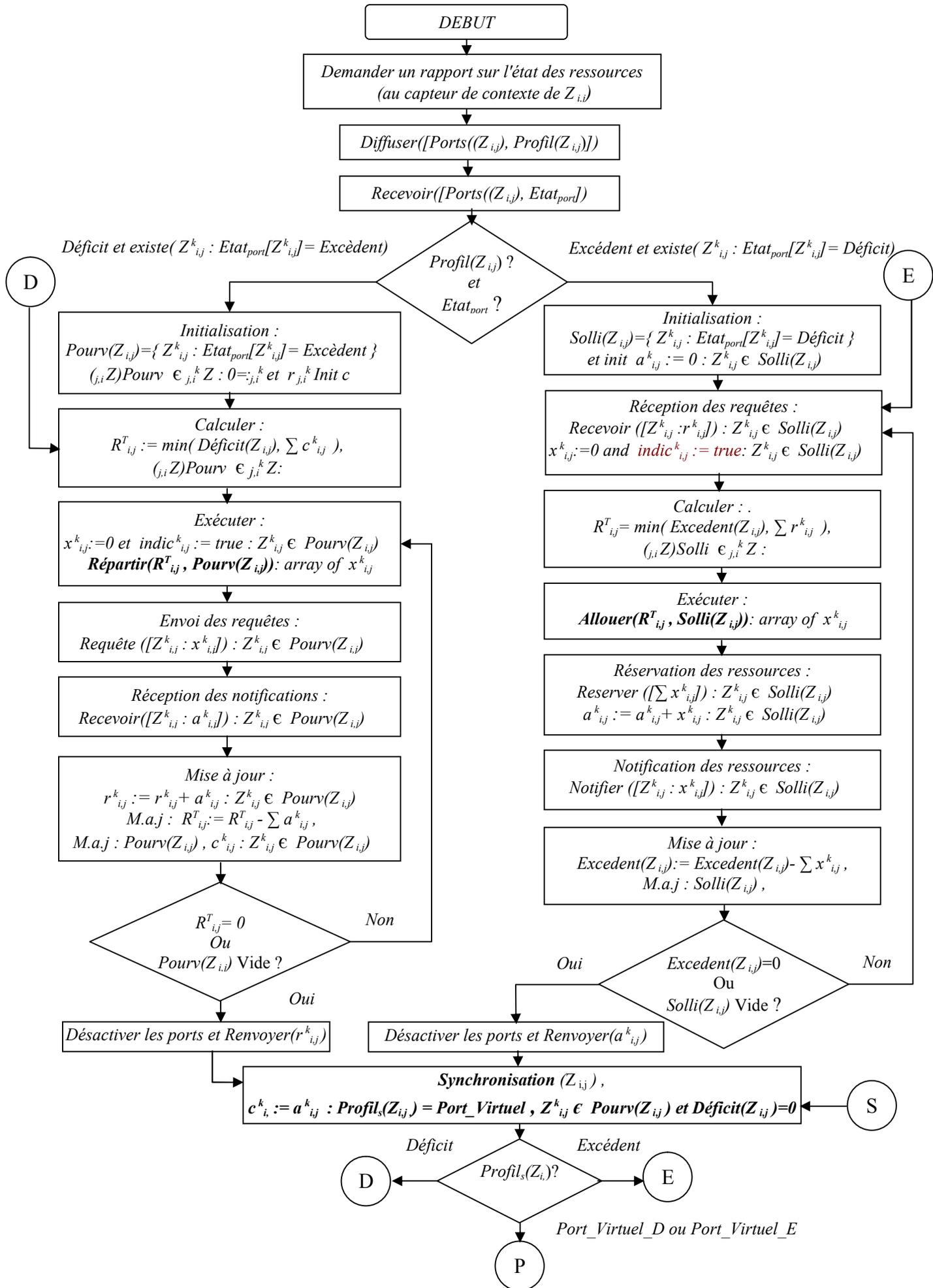
Figure 4.16 Algorithme de l'étape de synchronisation.

Pour les zones à état "Port_Virtuel_E", la question est de savoir quels sont les ports qui débouchent sur des zones "Excédent", et quels sont ceux qui débouchent sur des zones "Déficit", éventuellement à travers des zones "Port_Virtuel_D". Les réponses à ces questions seront données par leurs interactions

avec les zones à états "*Déficit*" et "*Port_Virtuel_D*". A leur tour, les zones "*Déficit*" et "*Port_Virtuel_D*" voudront aussi connaître les nouvelles configurations de leurs voisinages. Les zones à état "*Déficit*" chercheront à déterminer les ports vers les zones, nécessairement "*Port_Virtuel_E*", qui constitueront les ensembles $Pourv(Z_{i,j})$ potentiels. Pour les zones à état "*Port_Virtuel_D*", la question est de savoir quels sont les ports qui débouchent sur des zones "*Port_Virtuel_E*", et quels sont ceux qui débouchent sur des zones "*Excédent*". La connaissance, pour chaque zone, de la configuration de son voisinage, tel que présenté ci-dessus, permettra de relayer les indications de présence de zones encore excédentaires jusqu'aux zones encore en déficit à travers une chaîne de zones "*Port_Virtuel*", qui joueront le rôle des relais dans cette opération, recevant les indications arrivant sur certains ports et les retransmettant sur le reste des autres ports. Ces interactions de signalisation conduiront à la formation d'un nouveau réseau de liens, entre les différentes zones, qui supportera le déroulement des traitements du cycle suivant.

Le "Comportement" du Port Virtuel

Sous ce vocable nous désignons des zones qui soit, ont distribuer tout leurs excédents de ressources ("*Port_Virtuel_E*"), soit, ont résorber tout leurs déficits de ressources ("*Port_Virtuel_D*"), mais qui continuent de participer au processus de redistribution. Leur présence pallie l'absence de liens directs (d'où le terme de "*port virtuel*") entre les zones excédentaires et celles déficitaires, qui devront pouvoir interagir encore selon les mêmes "comportements" (adoptés en présence de liens directs), sans changements aucuns. Ainsi donc, ces "*port virtuel*" devront être transparents, et devront aussi agir en accord avec les principes qui ont guidés la construction du schéma. Pour s'acquitter de ces contraintes, le "comportement" des zones "*port virtuel*" a été conçu à l'image d'un "comportement" de zones déficitaires alternant avec celui des zones excédentaires. C'est en fait, d'un fonctionnement à quatre temps dont nous parlons ici. Dans le premier temps, le "*port virtuel*" collecte les demandes de ressources de son voisinage (peut être indirect) puis, dans le deuxième temps, fera siennes ces demandes en adoptant un "comportement" de zone déficitaire cherchant à combler son déficit. Dans le troisième temps, il constatera le nombre de ressources qu'il a réussi à obtenir, pour les redistribuer, dans le quatrième temps, à travers les ports d'où les demandes sont parvenues, en adoptant un "comportement" de zone excédentaire pourvoyant les ressources qui lui sont sollicitées. Rappelons à cet endroit, que chaque zone "*port virtuel*" distingue deux groupes de ports (identifiés lors de l'étape de synchronisation), l'un constituant son ensemble $Solli(Z_{i,j})$, l'autre son ensemble $Pourv(Z_{i,j})$, et qu'à ce titre elle doit adopter les "comportements" qui conviennent vis-à-vis de chacun de ces deux groupes de ports.



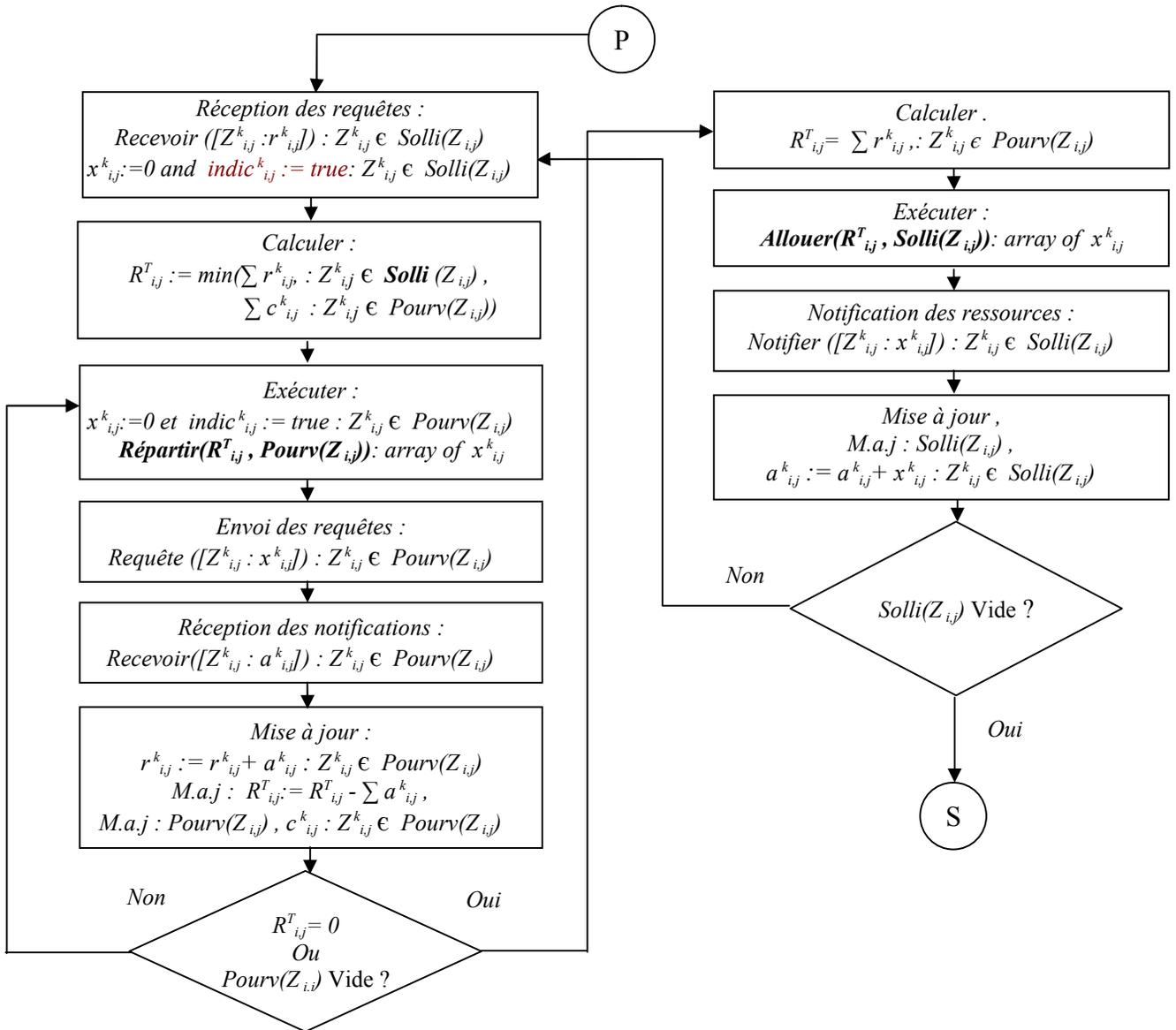


Figure 4.17 Schéma de redistribution des RR dans un RSMH, dans sa version finale.

4.3.4 Résultats des Simulations

Dans cette partie de notre travail, nous présentons et commentons, les résultats des simulations que nous avons conduit pour évaluer l'impact du schéma proposé, sur les performances d'un réseau sans fil mobile hétérogène auquel il viendrait à être intégré. Notre objectif premier était surtout de nous assurer de l'apport, que pouvait procurer l'adoption d'un tel schéma dans un réseau sans fil mobile hétérogène. Pour se faire, nous avons choisi d'analyser le fonctionnement d'un tel réseau, sous le contrôle d'un mécanisme à base de Bandes de Garde, avec et sans intégration du schéma que nous avons proposé. Les mécanismes à base de bandes de garde ont fait l'objet, dans le domaine des réseaux sans fil mobiles, de plusieurs travaux de recherche, sous diverses variantes. D'autres mécanismes existent aussi mais, nous signalons que pour notre part, il ne s'agit pas de les comparer, ni de vanter les mérites de ce que nous proposons par rapport à ces mécanismes. Nous voulons juste valider l'idée que l'implication de notre schéma dans la gestion des ressources radio d'un réseau sans fil mobile hétérogène est de nature à améliorer, en la matière, les performances de celui ci. Nous avons adopté une démarche qui consiste à se placer dans un cas limite, pour lequel le mécanisme à bandes de garde n'assure plus les performances souhaitées et, de montrer ensuite que l'implication de notre schéma arrive bien à redresser la situation. Il faudra, bien sûr, vérifier que ses résultats sont bien le fait du schéma proposé. Cette vérification, nous l'avons menée à travers trois

axes, caractérisant le degré de mobilité dans le réseau, le niveau d'intervention du schéma dans le processus de gestion des ressources et, le taux d'utilisation de ces dernières. Trois métriques sont aussi utilisées, il s'agit de la probabilité de coupure d'un appel(Dropping probability), de la probabilité de blocage d'un appel(Blocking probability) et, de l'intensité d'utilisation des ressources.

Nous signalons enfin, que dans nos simulations le rôle primordial des capteurs de contextes n'a pas été intégré, pour des raisons de simplicité, et vue que cet aspect mérite, à lui seul, un travail à part. Ceci nous conduit à dire que dans nos analyse des résultats obtenus nous ne devons pas perdre de vue cette absence, qui autrement aurait "boosté" ces résultats.

4.3.4.1 Cadre des simulations

La simulation est conduite dans le cadre du réseau représenté sur la figure 4.18, qui abrite trois technologies d'accès différentes (trois couleurs différentes), réparties en quinze (15) zones, numérotées de 0 à 14. Les données de configuration des zones sont représentées dans la table 4.2. La bande de ressources comprise entre le seuil de surcharge et la capacité de la zone est réservée exclusivement pour les

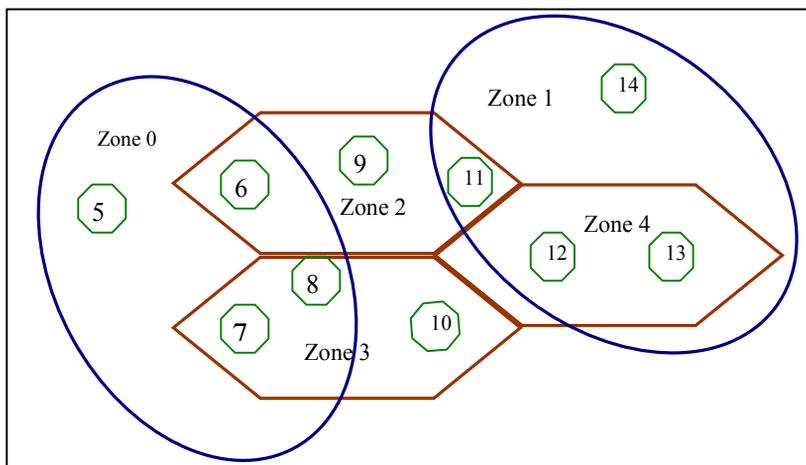


Figure 4.18 : Topologie du réseau cadre de la simulation.

handovers verticaux. La durée de la simulation est l'équivalent de 20 heures de fonctionnement, dont le début est fixé à 1800 s à partir du début de fonctionnement (aléatoirement choisi) de chaque zone. Les simulations sont faites pour des degrés de mobilité différents, un paramètre caractérisant celle-ci est utilisé et varie de un à trois (1 à 3). La valeur un (1) pour ce paramètre dénote une faible mobilité, alors

Zone	Capacité	Seuil de surcharge	Seuil de sous utilisation	Arrivées/seconde	Durée moyenne(/s)
0	26	23	15	012.00	180.00
1	25	23	15	012.00	180.00
2	22	20	15	012.00	150.00
3	22	20	15	012.00	150.00
4	22	20	15	012.00	150.00
5	26	24	20	009.57	240.00
6	26	24	20	009.57	240.00
7	26	24	20	009.57	240.00
8	26	24	20	009.57	240.00
9	26	24	20	009.57	240.00
10	26	24	20	009.57	240.00
11	26	24	20	009.57	240.00
12	26	24	20	009.57	240.00
13	26	24	20	009.57	240.00
14	26	24	20	009.57	240.00

Table 4.2 données de configuration des zones du réseau.

qu' une valeur de trois (3) est utilisée pour le cas d'une très grande mobilité. Trois scénarios sont simulés à chaque fois. Le premier (B.G) représente le cas où seul un mécanisme à base de bandes de garde est à l'œuvre, sans intervention de notre schéma de redistribution des ressources. Le deuxième scénario est celui de l'intervention du schéma mais, dans 70% des cas seulement (B.G + Sch à 70%). Enfin le dernier scénario, qui est celui de l'intervention du schéma à l'occurrence de chaque situation de déséquilibre (B.G + Sch à 100%). La prise de décision relative au choix des zones de destination lors des handovers est basée sur les pourcentages de recouvrement des zones entre elles. A défaut, l'appel est réinjecté dans le réseau, mais à partir du bord opposé.

Pour interagir entre elles les zones constituent un réseau de ports tel que celui repris sur la figure 4.19. La première métrique utilisée dans la simulation est la probabilité de coupure, qui représente la fréquence des coupures d'appels en cours de session (à l'occasion des handovers verticaux) pour manque de ressources radio. Elle est calculée comme étant le rapport: nombre d'échecs de handovers/nombre totale de

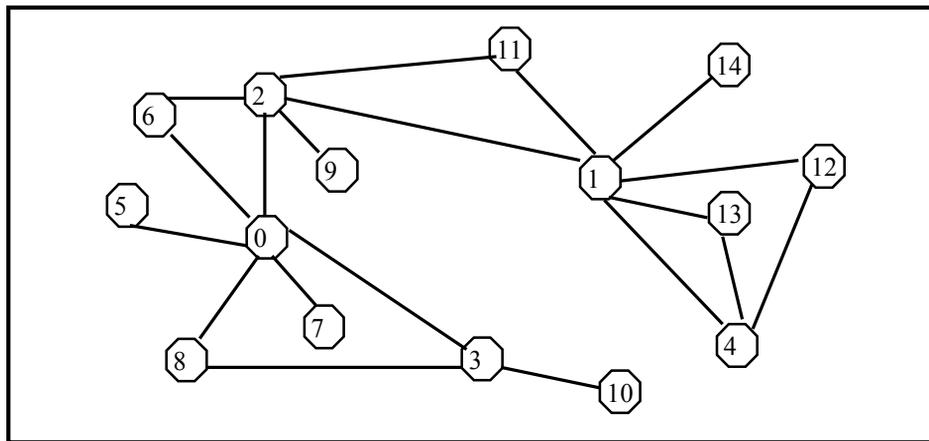


Figure 4.19: Réseau formé par l'interconnexion des ports des zones.

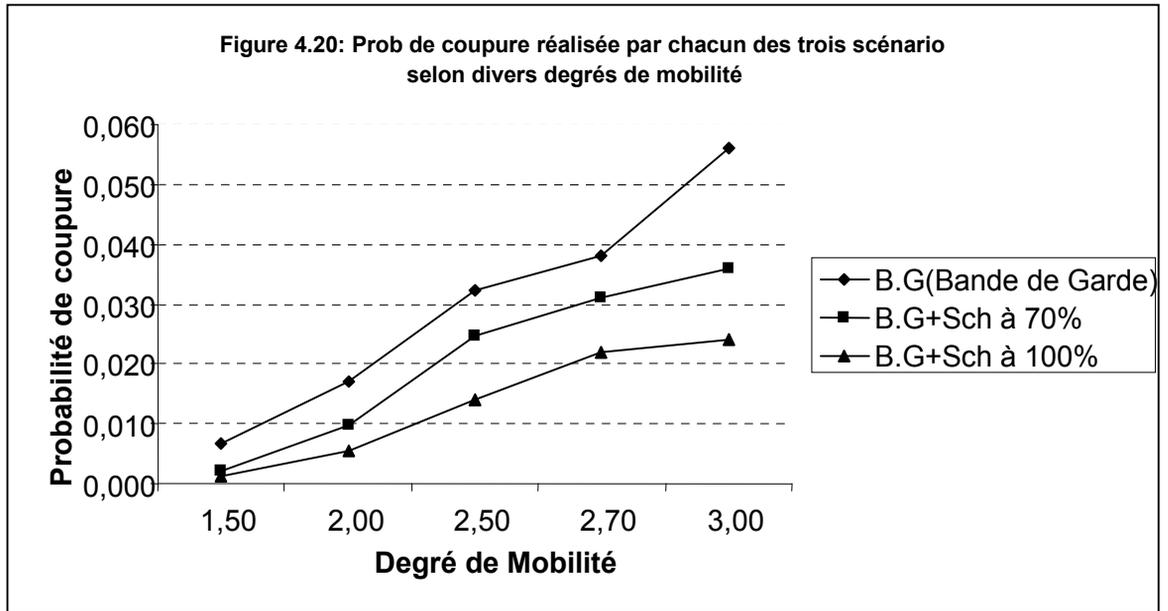
handovers. On admet en général que cette grandeur doit rester en dessous de 2%. La deuxième métrique est la probabilité de blocage d'appel, qui représente la fréquence des nouveaux appels rejetés pour insuffisance de ressources radio. On calcul cette valeur comme étant le rapport : nombre d'échecs de nouveaux appels/nombre total d'appels. Enfin, le taux d'utilisation des ressources est calculé comme la valeur moyenne, sur la durée de la simulation, des quantités : $(s_i \times t_i)$, où s_i représente un volume de ressources radio d'une zone donnée, et t_i la durée pendant laquelle ce volume est resté inchangé. Cette valeur est un indicateur de l'efficacité dans l'utilisation des ressources radio dans le réseau.

4.3.4.2 Résultats des Simulations

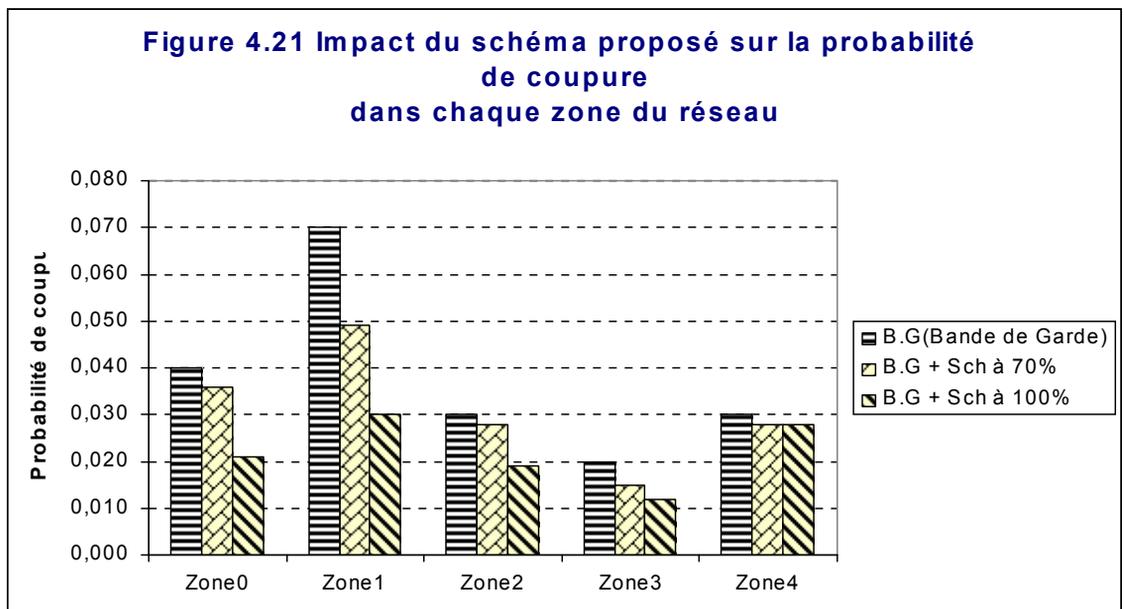
Nous allons dans ce qui suit passer en revue les résultats délivrés par nos simulations.

Probabilité de Coupure

Le premier, et le plus important, de ces résultats est celui retraçant l'impact de l'implication de notre schéma sur un critère important de la QoS, à savoir, la probabilité de coupure d'appel. Comme illustré sur la figure 4.20, on peut constater que l'intervention de notre schéma améliore sensiblement la moyenne des valeurs de cette métrique. Cette constatation reste valable même lorsque le schéma n'est impliqué qu'avec une probabilité de 0.7, c'est-à-dire qu'il n'intervienne que dans 70% des cas (B.G + Sch à 70%). Plus encore, nos résultats montrent aussi que la valeur de la probabilité de coupure est ramenée sous la barre des 2%, pour des degrés de mobilité assez élevés (entre 2.5 et 2.7). En effet, nous nous sommes placés dans



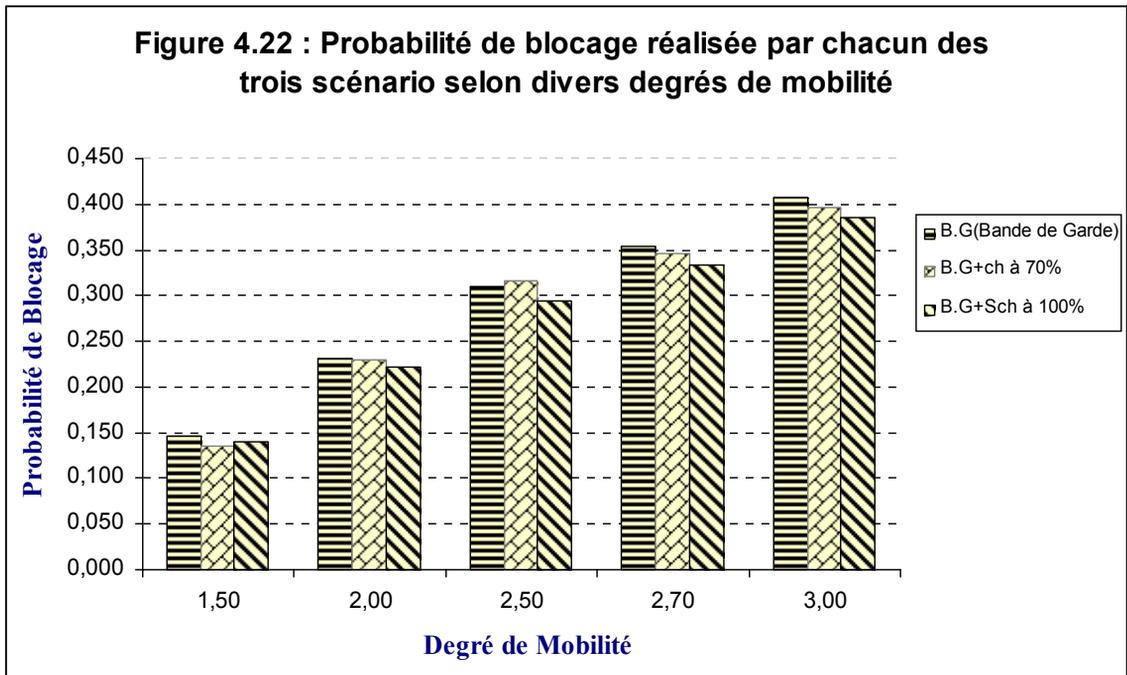
un cas limite où les bandes de garde, seules, ne suffisent plus à maintenir une QoS acceptable, à moins de d'augmenter la capacité du réseau par l'injection de plus de ressources radio. Notre schéma vient lever cette contrainte en évitant un investissement supplémentaire en ressources radio. Nous pouvons donc affirmer que le schéma proposé réalise effectivement une meilleure efficacité dans la gestion des ces ressources rares, que sont les ressources radio.



A ce stade une question se pose. Ne se peut-il pas que cette amélioration soit le fait de l'amélioration enregistrée dans seule zone, qui vient rééquilibrer la valeur moyenne du réseau? La réponse à cette question est justement contenue dans les résultats illustrés sur la figure 4.21. Les résultats reportés à cet endroit montrent que l'impact est au niveau de toutes les zones du réseau. Comme on peut le constater les valeurs de la probabilité de coupure s'améliorent, par l'entremise de notre schéma, dans chaque zone et non pas pour quelques unes seulement. Ceci nous conforte dans notre conviction que l'augmentation des performances est bien le fruit de la mise en œuvre du schéma dont il est question dans e travail.

Probabilité de Blocage

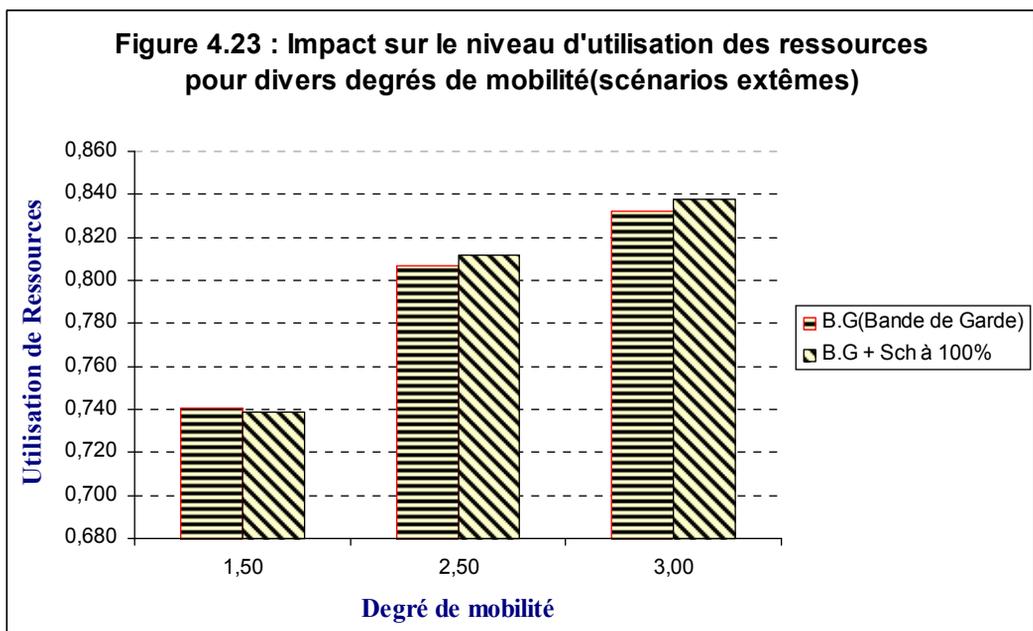
Les résultats évoqués précédemment n'ont pas dissipés tous les doutes. En effet, il se pourrait que derrière cette apparence se cache une détérioration d'un autre critère de QoS, certes moins important que le premier, mais très important tout de même. Il s'agit de la probabilité de blocage d'appel. La figure 4.22,



montre des résultats qui dissipent cette inquiétude. On peut voir qu'au contraire une très légère amélioration de cette probabilité est perceptible, et que les résultats de la probabilité de coupure sont vraiment à chercher dans l'action du schéma de redistribution des ressources, qui vient compléter celle du mécanisme en place, c'est-à-dire celui des bandes de garde. On peut remarquer aussi que, même lorsque le schéma intervient avec une probabilité de 0.7 la probabilité de blocage n'est pas pour autant détériorée. Avec ces résultats on est presque sûr que l'impact du schéma proposé est bien réel.

Niveau d'Utilisation des ressources

Si les faits sont bien, tels que nous les avons présentés, nous devrions nous attendre à ce que les niveaux



d'utilisation des ressources radio (tels que les avons définis plus haut) enregistrent aussi une amélioration. Qu'en est-il réellement des résultats relatifs à ce volet. Ces résultats, reportés sur la figure 4.23, semblent être, à première vue, un peu décevant. Cela est dû au fait que l'amélioration dans l'utilisation des ressources n'est pas très importante. En fait, cette modestie dans l'amélioration de l'utilisation des ressources radio n'est qu'apparente, car derrière, se cache un impact bien réel. Cet impact, nous le découvrons à travers l'analyse de l'évolution, des niveaux des ressources, induite par la mise en œuvre du schéma de redistribution. En effet, les valeurs, pour les niveaux d'utilisation des ressources, calculées avant la mise en œuvre, sur le compte des zones surchargées, le sont en partie calculées, sur le compte des zones sous utilisées après la mise en œuvre. Par le jeu de l'alternance des états de surcharge et de sous utilisation ces valeurs se compensent entre elles et n'apparaissent pas dans les résultats en tant qu'augmentations dans des niveaux d'utilisation des ressources radio. Donc, au lieu de constituer une contradiction, les résultats de la figure 4.23 confirment le fonctionnement que nous attendions de la mise en œuvre du schéma.

Comme nous venons de le montrer, chacun des trois axes de la simulation que nous venons de présenter, consolide les autres axes. Cela nous donne un ensemble cohérent, de résultats, qui nous permet de conclure au bien fondé des idées qui sous tendent notre proposition d'un schéma de redistribution des ressources radio dans un réseau sans fil hétérogène, en tant que complément important à une gestion efficace de ces ressources rares que sont les ondes radio. Ceci est d'autant plus vrai que l'apport des capteurs de contextes n'a pas pu être intégré, comme déjà signalé à l'introduction.

CONCLUSION

La rareté des ressources radio dont souffrent les réseaux sans fil actuels, et dont ils continueront d'en souffrir à l'avenir, appelle des techniques sophistiquées pour la gestion de ces dites ressources. Des travaux de recherches nombreux, ont examinés la question et ont débouchés sur des approches multiples. Il semble qu'aucune approche ne suffit, à elle seule, pour relever tous les défis qui se posent aux concepteurs, tant le problème est complexe et ses aspects divers. Nous sommes arrivés à conclure que l'implication de plusieurs composants spécialisés est une bonne piste, pourvu qu'il y ait des mécanismes efficaces qui permettent l'intégration et la coopération entre ces composants. Le travail qui vient d'être présenté tente d'envisager une telle approche. Nous y avons discuté un schéma de redistribution des ressources dans un réseau sans fil et de mobiles hétérogène (RSMH) qui met en œuvre une politique agressive d'utilisation des ressources radio "oisives", conséquence de la disparité spatiale et temporelle, des taux d'utilisation de ces ressources dans les diverses régions du réseau. Nous avons aussi exploité l'hétérogénéité de ces réseaux et les chevauchements de couverture sans fil pour dresser une sorte de carte dynamique de répartition des ressources radio, utilisée en tant qu'outil de prévision. Nous pensons que notre schéma doit être analysé en tant que point d'articulation des divers composants impliqués dans la GRR. Dans la conception de ce schéma lui-même nous avons, à dessein, choisi un fonctionnement distribué et parallèle où les entités impliquées peuvent se contenter d'un minimum de connaissances mutuelles. Nous pensons avoir pris le parti de la simplicité, tout comme nous garantissons l'équité de la distribution et sa réalisation d'un équilibre de pareto. Une étude préliminaire nous laisse penser que notre approche est adaptée à la gestion de la mobilité et la prise en charge de la QoS. Cette question pourrait faire l'objet de travaux de recherches futurs en vue de proposer une architecture complète d'un RSMH mettant en œuvre cette approche. Des questions connexes pourraient aussi faire l'objet de recherches plus poussées telles que la faisabilité du schéma de prévision proposé, le fonctionnement dans les conditions de fortes charges (contrôle de congestion), et la prise en charge de la QoS.

REFERENCES

- [1] POLY JL Langlois Chap 2.1.
- [2] Cours B11 – *Transmission Des Télécommunications* – CNAM Partie 2 - Chapitre 8.
- [3] Jean Philippe Muller. *Structure et Fonctionnement du Réseau GSM*. Techno Assistance Formation, <http://www.ta-formation.com>, 2000.
- [4] Hakim Badis. *Les Réseaux Mobiles et Sans Fil Réseaux Cellulaires*. IGM, Université de Paris-Est Marne-la-Vallée 2007/2008.
- [4] Marc Kahabka. *Pocket Guide for Fundamentals and GSM Testing*, Publisher: Wandel & Goltermann Germany.
- [5] Michel Terré. *Radio Communications*. Conservatoire des Arts et Métiers ELE208 première partie 2008-2009.
- [6] Stéphane Ubéda & Fabrice Valois. *Télécommunications Services & Usages Chapitre 3*. INSA-INRIA 2007.
- [7] C. Demoulin, M. Van Droogenbroeck. *Principes de Base du Fonctionnement du Réseau GSM*. *Revue de l'AIM*, pages 3-18, N04, 2004.
- [18] *Agilent Understanding General Packet Radio Service (GPRS) Application Note 1377*, Agilent Technology 2001.
- [9] H.kan Granbohm and Joakim Wiklund. *GPRS—General packet radio service*. Ericsson Review No. 2, 1999.
- [10] Carlo Copp. *An Introduction to Spread Spectrum Techniques*. Dr Carlo Kopp's Industry Publications <http://www.csse.monash.edu.au/~carlo/SYSTEMS/Spread-Spectrum-0597.htm>.
- [11] Randy Robert. The ABC of Spread Spectrum, a Tutorial. RF Spread Spectrum Consulting. <http://sss-mag.com/ss.html#tutorial>.
- [12] Michael Hendry. Introduction to CDMA. <http://www.bee.net/mhendry/vrml/library/cdma/cdma.htm>.
- [13] Al Agha Pujolle Vivier. *Réseaux de Mobiles & Réseaux Sans Fil*. Editions EYROLLES, 2001.
- [14] Charan Lagton. *Intuitive Guide to Principles of Communications*, 2002. www.complextoreal.com/chapters/fm.pdf.
- [15] Rackley. *Wireless Networking Technology From Principles to Successful Implementation*. Newnes 2007.
- [16] P. Nicopolitidis, M. S. Obaidat, G. I. Papadimitriou, A. S. Pomportsis. *Wireless Networks*. JOHN WILEY & SONS, 2003.
- [17] *Introduction to cdma2000 Standards for Spread Spectrum*. 3GPP2 V.1.0 September 2009.
- [18] Serge Willenegger. *cdma2000 Physical Layer: An Overview*. QUALCOMM International, Switzerland 2000.
- [19] Faruque, Saleh. (1996). *Cellular Mobile Systems Engineering*. Norwood, MA: Artech House.
- [20] P. Nicopolitidis, M. S. Obaidat, G. I. Papadimitriou, A. S. Pomportsis. *Wireless Networks*. John Wiley & Sons 2003.
- [21] Mohamed Moussaoui. *L'UMTS et les Technologies Emergentes*. ENSA Tanger, Maroc 2009.
- [22] Harri Holma, Antti Toskala. *WCDMA for UMTS: HSPA evolution and LTE*. 2007 John Wiley & Sons.
- [23] Keiji Tachikawa. *W-CDMA Mobile Communications System*. Wiley and Maruzen 2002.
- [24] Livre Pujolle
- [25] Sassan Ahmadi. *Introduction to mobile WiMAX Radio Access Technology: PHY and MAC Architecture*. UCSB Presentation 2006. http://www.mat.ucsb.edu/~ggroup/ahmadiUCSB_slides_Dec7.pdf.
- [26] National Instruments Tutorial. *OFDM and Multi-Channel Communication Systems*. 2007. <http://zone.ni.com/devzone/cda/tut/p/id/3740>
- [27] Per Hjalmar Lehne, Frode Bøhagen. *OFDM(A) for Wireless Communication*. research report telenor 2008.

- [28] X. Gelabert et al., *Radio Resource Management in Heterogeneous Networks. HET-NETs 2005*, West Yorkshire, UK, July 2005.
- [29] Roberto Verdone, Alberto Zanella. *Radio Resource Management in Mobile Radio Networks*, TeleDoc at <http://docs.google.com/viewer?a=v&q=cache:8XvhyHDJaU8J:www-site.deis.unibo.it/htdida/sistra>
- [30] C. M. Chin, C. M. Tan and M. L. Sim. *Future Trends in Radio Resource Management for Wireless Communications*, BT Technology Journal Vol 24 No 2 April 2006.
- [31] Z. Liu and M. El Zarki, "SIR-based Call Admission Control for DS-CDMA Cellular Systems," *IEEE JSAC*, vol.12, no. 4, May 1994, pp. 638–44.
- [32] T. Shu and Z. Niu, "Call Admission Control using Differentiated Outage Probabilities in Multimedia DS-CDMA Networks with Imperfect Power Control," *Proc. 11th Int'l. Conf. Comp. Commun. and Net.*, 2002, Oct. 2002, pp. 336–41.
- [33] Z. Dziong, M. Jia, and P. Mermelstein, "Adaptive Traffic Admission for Integrated Services in CDMA Wirelessaccess Networks," *IEEE JSAC*, vol. 14, no. 9, Dec. 1996, pp. 1737–47.
- [34] T. Liu and J. Silvester, "Joint Admission/Congestion Control for Wireless CDMA Systems Supporting Integrated Services," *IEEE JSAC*, vol. 16, no. 6, Aug. 1998, pp. 845–57.
- [35] J. Wu, "Performance Analysis of QoS-Based Voice/Data CDMA Systems," *Wireless Pers. Commun.*, June 2000, pp. 223–36.
- [36] Y. Guo and B. Aazhong, "Call Admission Control in Multiclass Traffic CDMA Cellular System Using Multiuser Antenna Array Receiver," *Proc. IEEE Vehic. Tech. Conf. (VTC '00-Spring)*, May 2000, vol. 1, pp. 365–69.
- [37] M. Andersin, Z. Rosberg, and J. Zander, "Soft and Safe Admission Control in Cellular Networks," *IEEE Trans. Net.*, vol. 5, no. 2, Apr. 1997, pp. 255–65.
- [11] C. Huang and R. Yates, "Call Admission in Power Controlled CDMA Systems," *Proc. IEEE 46th Vehic. Tech. Conf. (VTC'96)*, 1996, vol.3, pp. 1665–69.
- [38] M. Park and S. Oh, "Dynamic Power and Resource Allocation for Wireless Multimedia Traffics in Reverse Link of WCDMA Systems," *Proc. Wireless '99*, 1999, pp. 85–91.
- [39] C. Ho et al., "Impact of the Cell Size on the Cell's Erlang Capacity and Call Admission Control in the DS/CDMA Cellular Networks," *Proc. IEEE Vehic. Tech. Conf. (VT'00-Spring)*, Tokyo, May 2000, vol. 1, pp. 385–89.
- [40] W. Yang and E. Geraniotis, "Admission Policies for Integrated Voice and Data Traffic in CDMA Packet Radio Networks," *IEEE JSAC*, vol. 12, no. 4, May 1994, pp. 654–64.
- [41] D. Hong and S. Rappaport, "Traffic Model and Performance Analysis for Cellular Mobile Radio Telephone Systems with Prioritized and Nonprioritized Handoff Procedures," *IEEE Trans. Vehic. Tech.*, vol. 35, no. 3, Aug. 1986, pp. 77–92.
- [42] M. Naghshineh and M. Schwartz, "Distributed Call Admission Control in Mobile/Wireless Networks Personal, Indoor and Mobile Radio Communications," *Proc. 6th IEEE Int'l. Symp. Pers., Indoor and Mobile Radio Commun. (PIMRC'95)*, 1995, vol. 1, pp. 289–93.
- [43] M. Naghshineh and M. Schwartz, "Distributed Call Admission Control in Mobile/Wireless Networks," *IEEE JSAC*, vol. 14, no. 4, May 1996, pp. 711–17.
- [44] C. Oliviera, J. Kim, and T. Suda, "An Adaptive Bandwidth Reservation Scheme for High-Speed Multimedia Wireless Net- Works," *IEEE JSAC*, vol. 16, no. 6, Aug. 1998, pp. 858–72
- [45] K. Begain, G. Bolch, and M. Telek, "Scalable Schemes for Call Admission and Handover in Cellular Networks with Multiple Services," *Wireless Pers. Commun.*, Nov. 2000, pp. 125–44.
- [46] D. Levine, I. Akyildiz, and M. Naghshineh, "A Resource Estimation and Call Admission Algorithm for Wireless Multimedia Networks using the Shadow Cluster Concept," *IEEE/ACM Trans. Net.*, vol. 5, no. 1, Feb. 1997, pp. 1–12.
- [47] M. Naghshineh and A. Acampora, "QOS Provisioning in Microcellular Networks Supporting Multimedia Traffic," *Proc. 14th Annual Joint Conf. IEEE Comp. and Commun. Societies (INFOCOM '95)*, 1995, vol. 3, pp. 1075–84.

- [48] M. Naghshineh and A. Acampora, "Design and Control of Micro-cellular Networks with QoS Provisioning for Data Traffic," *Wireless Networks*, vol. 3, no. 4, 1997, pp. 249–56.
- [49] J. Choi *et al.*, "Call Admission Control for Multimedia Services in Mobile Cellular Networks: A Markov Decision Approach," *Proc. 5th IEEE Symp. Comp. and Commun. (ISCC 2000)*, 2000, pp. 594–99.
- [50] S. Kim, T. Kwon, and Y. Choi, "Call Admission Control for Prioritized Adaptive Multimedia Services in Wireless/Mobile Networks," *Proc. IEEE Vehic. Tech. Conf. (VTC '00-Spring)*, Tokyo, 2000, vol. 2, pp. 1536–40.
- [51] T. Kwon *et al.*, "Threshold-type Call Admission Control in Wireless/Mobile Multimedia Networks using Prioritized Adaptive Framework," *IEEE Electronics Letters*, vol. 36, no. 9, Apr. 2000, pp. 852–54.
- [52] T. Kwon and M. Naghshineh, "Optimal Distributed Call Admission Control for Multimedia Services in Mobile Cellular Network," *Proc. Mobile Multimedia (MoMu '98)*, 1998.
- [53] Qing-An Zeng, Dharma P. Agrawal. *Handbook of Wireless Networks and Mobile Computing*, Edited by Ivan Stojmenovic. ISBN 0-471-41902-8 © 2002 John Wiley & Sons, Inc.
- [54] Yue Chen. *Soft Handover Issues in Radio Resource Management for 3G WCDMA Networks*. Thesis Submitted for the degree of Doctor of Philosophy Department of Electronic Engineering Queen Mary, University of London September 2003.
- [55] Yaxin Cao, Victor O. K. Li. *Scheduling Algorithms in Broad-Band Wireless Networks*. IEEE Proceedings of The IEEE, Vol. 89, NO. 1, January 2001
- [56] P. Bhagwat, A. Krishna, S. Tripathi, *Enhancing throughput over wireless LAN's using channel state dependent packet scheduling*, in *Proc. INFOCOM96*, Mar. 1996, pp. 1133–1140.
- [57] C. Fragouli, V. Sivaraman, and M. Srivastava, *Controlled multimedia wireless link sharing via enhanced class-based queuing with channel-state dependent packet scheduling*, in *Proc. INFOCOM' 98*, vol. 2, Mar. 1998, pp. 572–580.
- [58] Oriol Sallent, QoS and RRM in 3G and Beyond systems. *2nd Conference on Next Generation Internet Design and Engineering València*, April 3- 5 2006.
- [59] Kandaraj Piamrat, César Viho, Adlen Ksentini, Jean-Marie Bonnin, Resource Management in Mobile Heterogeneous Networks: State of the Art and Challenges. *IRISA Publication Interne N° 1889*, February 2008.
- [60] G. Koundourakis, D.I. Axiotis, and M. Theologou. Network-Based Access Selection in Composite Radio Environments. *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE*, pages 3877–3883, 11-15 March 2007.
- [61] O. Ormond, P. Perry, and J. Murphy. Network Selection Decision in Wireless Heterogeneous Networks. *Personal, Indoor and Mobile Radio Communications, 2005. PIMRC 2005. IEEE 16th International Symposium on*, 4:2680–2684 Vol. 4, 11-14 Sept. 2005.
- [62] Xiaoshan Liu, Victor O. K. Li, and Ping Zhang. Nxg04-4: Joint Radio Resource Management through Vertical Handoffs in 4g networks. *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*, pages 1–5, Nov. 2006.
- [63] Xu Yang, J. Bigham, and L. Cuthbert. Resource management for service providers in heterogeneous wireless networks. *Wireless Communications and Networking Conference, 2005 IEEE*, 3:1305–1310 Vol. 3, 13-17 March 2005.
- [64] A.-E.M. Taha, H.S. Hassanein, and H.T. Mouftah. On Robust Allocation Policies in Wireless Heterogeneous Networks. *Quality of Service in Heterogeneous Wired/Wireless Networks, 2004. QSHINE 2004. First International Conference on*, pages 198–205, 18- 20 Oct. 2004.
- [65] Lorenza Giupponi, Ramon Agustí, A Novel Approach for Joint Radio Resource Management Based on Fuzzy Neural Methodology. *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*, Vol. 57, No. 3, MAY 2008.
- [66] Suliman IM, Pomalaza-Ráez C, Oppermann I., Lehtomäki J. Radio Resource Allocation in Heterogeneous Wireless Networks using Cooperative Games. *Nordic Radio Symposium 2004 / Finnish Wireless Communications Workshop 2004 (NRS/FWCW 2004)*, August 2004.

- [67] W. F. Bialas. Cooperative N-person Stackelberg games. In *Proceedings of the 28th IEEE Conference on Decision and Control*, volume 3, pages 2439– 2444, 1989.
- [68] Qingyang Song and A. Jamalipour. A Network Selection Mechanism for Next Generation Networks. *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, 2:1418–1422 Vol. 2, 16-20 May 2
- [69] Abd-Elhamid M. Taha, Hossam S. Hassanein, Hissein T. Mouftah, Vertical handoffs as radio resource management tool in: *Computer Communications* 31 (2008) 950-961.
- [70] Salman AlQahtani, Ashraf S. Mahmoud, Asrar U. Sheikh, Adaptive radio resource with borrowing for multi-operators 3G+ wireless networks with heterogeneous traffic in: *Computer Communicatios* 29 (2006) 2945-2951.
- [71] Meriem Kassar, Brigitte Kervella, Guy Pujolle, An overview of vertical handover decision strategies in heterogeneous networks in: *Computer Communicatios* 31 (2008) 2607-2620.
- [72] Alexandros Kaloxylos, George Lampropoulos, Nikos Passas, Lazaros Merakos, A flexible handover mechanism for seamless service continuity in heterogeneous environments in: *Cmputer Communications* 29 (2006) 717-729.
- [73] Quoc-Thanh Nguyen-Vuong, Nazim Agoulmine, Yacine-Ghamri-Doudane, Auser-centric and context-aware solution to interface management and access network selection in heterogeneous wireless environments in: *Computers Netwroks* 52 (2008) 3358-3372
- [74] Wei Song, Hai Jiang, Weihua Zhuang, and Xuemin (Sherman) Shen, Resource Management for QoS Support in Cellular/WLAN Interworking in: *IEEE Network* • September/October 2005 12-18.
- [75] Roel Ocampo Lawrence Cheng Kerry Jean Alberto Gonzalez Prieto, Alex Galis, Zhaohong Laim Towards a Context Monitoring System for Ambient Networks in: *IEEE CHINACOM* (2006) 16-19.
- [76] L.Wang (1), H.Aghvami (1), Nima Nafisi(1), J. Pérez-Romero (2), O. Sallent (2), R. Agustí (2). Coverage-based Common Radio Resource Management in heterogeneous CDMA/TDMA Cellular Systems, available at: www.aroma-ist.upc.edu/publicdocuments/conferences/C24.pdf
- [77] Wei Shenm Qing-And Zeng, Resource Allocation Schemes in Integrated Heterogeneous Wireless and Mobile Networks in: *Journal of Netzorks*, VOL. 2, NO. 5, September 2007.
- [78] Mohamed Hossam Ahmed, MEMORIAL UNIVERSITY OF NEW FOUND LAND Call Admission Control in Wireless Networks: A Comprehensive Survey in: *IEEE Communications Surveys & Tutorials* • First Quarter 2005.
- [79] J.Perez-Romero, O.Sallent, A.Barbaresi, R.Ljung, A.Serrador, L.M Correa, Advanced Radio Resource Management Solutions in AROMA in: *Workshop Trends in Radio Resource Management (3rd Edition) Barcelona 21st November 2007*.
- [80] Chadi Tarhini, Tijani Chahedm, Density-based admission control in IEEE802.16e Mobile WiMAX .
- [13] R. Jain, D.M. Chiu and W. Hawe, "A Quantative Measure of fairness and Discrimination for Resource Allocation in Shared Systems," Technical Report, Digital Equipment Corporation, DEC-TR-301, 1984.
- [81] Hyun-Ho Choi and Dong-Ho Cho, TAKEOVER: A New Vertical Handover Concept for Next-Generation Heterogeneous Networks in: *IEEE* 2005.
- [82] Natasa Vulic, Sonia Heemstra de Groot and Ignas Niemegeers Common radio resource management for WLAN-UMTS integration at radio access level.
- [83] Yang Guang, Chen Jie, Yu Kai, Zhang Ping, Victor O.K. Li, Joint Radio Resource Management based on the Species Competition Model in: *IEEE* 2006.
- [84] ALEX HILLS, CARNEGIE MELLON UNIVERSITY BOB FRIDAY, AIRESpace, Radio Resource Management in Wireless LANs in: *IEEE Radio Communications* • December 2004.
- [85] KANDARAJ PIAMRAT, CÉSAR VIHO, ADLEN KSENTINI, JEAN-MARIE BONNIN, Resource Management in Mobile Heterogeneous Networks: State Of The Art and Challenges, IRISA Publication Interne N° 1889 February 2008.
- [86] X. Gelibert, J. Pérez-Romero, O. Sallent, R. Agustí On the Suitability of Load Balancing Principles in Heterogeneous Wireless Access Networks in: *IWS 2005 / WPMC'05 Aalborg. Denmark*.

- [87] Ruggiero Cavallo, Optimal Decision Making With Minimal Waste: Strategyproof Redistribution of VCG Payments in: *AAMAS'06* May 8–12 2006, Hakodate, Hokkaido, Japan.
- [88] Abraham, D.J. and Cechlarova, K. and Manlove, D.F. and Mehlhorn, K. (2004) Pareto optimality in house allocation problems. In, Fleischer, R. and Trippen, G., Eds. *Proceedings of ISAAC 2004: the 15th Annual International Symposium on Algorithms and Computation, 20-22 December, 2004* Lecture Notes in Computer Science Vol 3341, pages 3-15, Hong Kong.
- [89] Sungwook Kim and Pramod K. Varshney, An Integrated Adaptive Bandwidth-Management Framework for QoS-Sensitive Multimedia Cellular Networks in: *IEEE Transactions On Vehicular Technology*, VOL. 53, NO. 3, MAY 2004 835