

Méthodes d'optimisation de contrôleurs de logique floue pour le paramétrage automatique des réseaux mobiles UMTS

Hervé Dubreil

▶ **To cite this version:**

Hervé Dubreil. Méthodes d'optimisation de contrôleurs de logique floue pour le paramétrage automatique des réseaux mobiles UMTS. domain_other. Télécom ParisTech, 2005. English. NNT : . pastel-00001822

HAL Id: pastel-00001822
https://pastel.hal.science/pastel-00001822

Submitted on 30 Jun 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ECOLE NATIONALE SUPERIEURE DES TELECOMMUNICATIONS

T H E S E

pour obtenir le grade de

DOCTEUR DE L'ECOLE NATIONALE SUPERIEURE DES TELECOMMUNICATIONS

Discipline : Électronique et communications

présentée et soutenue

par

Hervé DUBREIL

Le 27 juin 2005,

Le chapitre 5.2 ayant été soutenu à huis clos.

Titre :

Méthodes d'optimisation de contrôleurs de logique floue pour le paramétrage automatique des réseaux mobiles UMTS

———

Directeurs de thèse :

Pr. Jean-Claude Belfiore

Dr. Arturo Ortega-Molina

———

JURY

M. Guy Pujolle	, Rapporteur
M. Djamal Zeghlache	, Rapporteur
M. Christian Bonnet	, Examinateur
M. Armand Lévy	, Examinateur

Remerciements

Je souhaiterais témoigner ma reconnaissance à l'ensemble des personnes qui m'ont permis de mener à terme mon projet de thèse.

Tout d'abord, je tiens à remercier Monsieur Armand Lévy qui accepta à mon arrivée dans son laboratoire que je puisse effectuer cette thèse en parallèle de mes activités professionnelles, ainsi que pour sa participation au jury de ma thèse, bien qu'il ait quitté entre-temps France Télécom Recherche & Développement.

Je tiens aussi à exprimer ma gratitude à Messieurs Arturo Ortega et Jean-Claude Belfiore, Professeur à l'Ecole National Supérieure des Télécommunications Paris, pour avoir accepté de co-encadrer cette thèse ainsi que pour leurs conseils et suivis avisés.

Mes plus vifs remerciements vont également à Zwi Altman pour son aide, ses nombreux conseils scientifiques et relectures qui m'ont été des plus précieux pour l'aboutissement de cette thèse.

Ma gratitude va aussi aux collègues et personnes qui par leurs idées et soutiens m'ont aidé dans mon travail, en particulier Christophe Gay, Ridha Nasri, Jean-Marc Picard, Benoît Fourestié, Vincent Diascorn, Frédéric Garabedian et Sana Ben Jemaa, ainsi que Mikaël Salaun pour ses relectures rigoureuses et ses avis éclairés.

Je souhaiterais remercier Monsieur Christian Bonnet, Professeur à Eurécom, qui accepta de présider le jury de ma thèse, ainsi que Messieurs Guy Pujolle, Professeur à l'Université Paris VI, et Djamal Zeghlache, Professeur à l'Institut National des Télécommunications, qui consentirent à en être les rapporteurs.

Et je n'oublie pas évidemment les encouragements et réconforts amicaux à divers titres de Fabrice, Laurent, Stéphane, Pierre, Javier, Frédéric, Olfa, Alban et Anne.

Enfin, ma reconnaissance et mon affection vont à mes chers parents, ainsi qu'à Julie, qui m'a soutenu et encouragé sans relâche au cours des périodes difficiles de ce projet.

Table des matières

1. Introduction à l'ingénierie automatique 7

2. Algorithmes RRM et performances des réseaux UMTS 14
2.1. Deux algorithmes RRM à optimiser 14
2.1.1. L'algorithme de contrôle d'admission 14
2.1.2. L'algorithme de macro-diversité 18
2.1.3. Autres paramètres à optimiser dans les réseaux mobiles 22
2.2. Etude des performances d'un réseau UMTS 24
2.2.1. Formulation statique du problème de capacité 24
2.2.2. Etude de la couverture 29
2.2.3. Evaluation statistique des performances des réseaux 30
2.3. Conclusions 30

3. Contrôle des paramètres RRM par logique floue 31
3.1. Modélisation par un système d'inférence floue 32
3.1.1. Caractéristiques d'un système d'inférence floue 32
3.1.2. Cas des réseaux mobiles UMTS 32
3.2. Définition d'une architecture de contrôle du paramétrage automatique 35
3.2.1. Description du contrôle automatique par logique floue 36
3.2.2. Types d'architectures réalistes envisageables 37
3.3. Description des contrôleurs de logique floue modélisé par un processus FKBC 42
3.3.1. Procédure de *fuzzification* 44
3.3.2. Module de prise de décisions 48
3.3.3. Procédure de defuzzification 66
3.4. Conclusions 67

4. Méthodes d'optimisation des contrôleurs de logique floue 69
4.1. Optimisation par essaim de particules 70
4.1.1. Introduction à l'optimisation par essaim de particules 70
4.1.2. Description de l'algorithme d'optimisation par essaim de particules 73
4.1.3. Application de l'optimisation par essaim de particules aux CLF 76
4.1.4. Schéma d'utilisation pour une optimisation du paramétrage des CLF 79
4.2. Optimisation par apprentissage par renforcement 81
4.2.1. Introduction à l'apprentissage par renforcement et choix du Q-learning 81
4.2.2. Définitions des fonctions utilisées dans les méthodes d'apprentissage par renforcement 84
4.2.3. Les équations de Bellman et notion de politique optimale 86
4.2.4. Les algorithmes d'apprentissage 88
4.2.5. Adaptation du Q-learning aux systèmes d'inférence floue 91
4.2.6. Mise en œuvre du Q-learning dans le cas de l'optimisation des CLF 98
4.2.7. Possibilités d'extensions de l'algorithme de Q-learning 102
4.2.8. Exploitation de systèmes d'inférence floue avec le Q-learning 104
4.3. Conclusions sur les méthodes d'optimisation 105

5. Fonctionnement en modes *off-line* / *on-line* 107
5.1. Le mode on-line en UMTS 108
5.1.1. Contraintes liées au mode *on-line* 108
5.1.2. Application de l'ingénierie automatique 113

5.1.3. Exemples de contrôle de paramètres RRM en mode *on-line* 114
5.2. Le mode off-line en UMTS 116
5.2.1. Objectifs du mode *off-line* 116
5.2.2. Avec l'optimisation par essaim de particules 118
5.2.3. Avec l'apprentissage par renforcement 119
5.2.4. Le mode intermédiaire "semi-dynamique" avec l'apprentissage par renforcement 123
5.3. Conclusions 124

6. Résultats obtenus avec l'ingénierie automatique en mode *on-line* 126
6.1. Premiers résultats d'optimisation de l'ingénierie dynamique par OEP 127
6.1.1. Détermination des fronts de paramétrage uniforme 127
6.1.2. Optimisation de l'ingénierie dynamique pour le contrôle d'admission 128
6.1.3. Optimisation de l'ingénierie dynamique pour la macro-diversité 130
6.2. Applications des politiques optimisées par OEP à un réseau différent 131
6.2.1. Applications des politiques optimisées pour le contrôle d'admission 131
6.2.2. Applications des politiques optimisées pour le contrôle d'admission et la macro-diversité 133
6.2.3. Variations de l'efficacité des contrôleurs en fonction de la densité de trafic . 134
6.2.4. Performances d'une politique directement optimisée sur le réseau de la ville 2 136
6.3. Analyse des premiers résultats d'optimisation des CLF 140
6.4. Méthode d'OEP en mode on-line 142
6.4.1. Rappel de la méthode 142
6.4.2. Mises en applications sur le réseau de la ville 1 142
6.4.3. Comparaison des résultats de l'OEP en mono et multi-objectifs 149
6.5. Premiers résultats d'optimisation de l'ingénierie dynamique avec l'algorithme du Q-learning 153
6.5.1. Mises en applications sur une zone réduite de la ville 2 153
6.5.2. Comparaison avec l'optimisation par essaim de particules 155
6.6. Conclusions 156

7. Résultats obtenus avec l'ingénierie automatique en mode *off-line* 158
7.1. Analyse de l'optimisation par essaim de particules en mode off-line 158
7.1.1. Méthode d'optimisation en mode *off-line* 158
7.1.2. Mises en applications sur le réseau de la ville 1 158
7.2. Comparaison des résultats de l'OEP en mode off-line en mono et multi-objectif . 161
7.3. Conclusions 163

8. Conclusions et perspectives 164

9. Références 166

10. Annexes 171
10.1. Architecture des réseaux mobiles UMTS 171
10.2. Différentes représentations paramétriques des matrices de décisions 172

11. Glossaire 174

1. Introduction à l'ingénierie automatique

A ce jour, il existe plusieurs outils de planification automatique ou ACP (*Automatic Cell Planner*) qui permettent aux opérateurs de réaliser un design efficace mono-système de leurs différents réseaux d'accès radio GSM/GPRS (*Global System for Mobile communications / General Packet Radio Service*), UMTS (*Universal Mobile Telecommunications System*) ou CDMA 2000 (*Code Division Multiple Access* : le standard 3G américain) en termes de localisation des sites, de paramétrage antennaire (azimuts et tilts), et de puissances des canaux communs (en utilisant par exemple des techniques d'optimisation par algorithmes génétiques) [Ben05] [Pic05]. Ces outils fournissent en sortie un ou plusieurs (selon le fonctionnement mono ou multi-objectifs) jeux des meilleurs sites à déployer (parmi une liste de sites possibles), et pour chaque site un ou plusieurs jeux de paramètres efficaces (azimut, tilt, puissance du canal pilote). L'optimisation du choix et du paramétrage des sites a pour objectifs d'améliorer principalement la couverture, la capacité et la qualité de service (*Quality of Service* : QoS) de chaque réseau d'accès radio.

Cependant malgré l'efficacité de ces outils, les opérateurs sont aussi contraints dans la sélection des sites à déployer par des facteurs extérieurs : contraintes environnementales (écoles, hôpitaux, bâtiments publics), politiques appliquées par les collectivités locales, régionales et nationales, négociations avec les associations de consommateurs et propriétaires de logements, sites négociés antérieurement par des opérateurs concurrents… Cela conduit souvent l'opérateur d'une part à réutiliser en partie les sites d'un réseau en service (GSM par exemple) pour le déploiement d'un nouveau réseau (UMTS récemment) et d'autre part à se contenter des sites "sous-optimaux" en termes de design. En conséquence, les outils de type ACP se limitent souvent à une optimisation des azimuts et tilts des antennes et des puissances des canaux communs. Une solution pour améliorer un tel design pourrait être de compenser ces choix de sites sous-optimaux par l'adaptation des paramètres des algorithmes de gestion de la ressource radio aux contraintes locales d'environnement et de trafic, afin d'améliorer la capacité et la couverture, mais aussi la qualité de service perçue par l'utilisateur. Par exemple, pour le réseau UMTS, les algorithmes RRM (*Radio Ressource Management*) dont on pourrait améliorer efficacement le réglage des paramètres comprennent la macro-diversité, le contrôle d'admission, le contrôle de charge, le *packet scheduling* (algorithme de gestion du mode paquet de transfert de données), voire le contrôle de puissance… Mais la difficulté principale est qu'il n'existe pas encore d'outil permettant une optimisation locale (au niveau de la station de base ou *node B*[*]) du réglage des paramètres de gestion de la ressource radio.

Par ailleurs, à l'ouverture des réseaux UMTS, le paramétrage des algorithmes RRM a été fixé constant sur des plaques de stations (se conformant dans la plupart des cas aux paramétrages par défaut fournis par les constructeurs), qui pouvait distinguer par exemple les zones urbaines, suburbaines et rurales. Dans un premier temps, ce paramétrage sous-optimal et non adapté localement n'a pas une incidence trop importante sur les performances du réseau, car la demande en trafic est encore faible. Le nombre de clients de troisième génération va ensuite s'accroître et la demande en trafic, bas débit (voix, SMS) comme haut débit (*streaming video*, visiophonie, téléchargement de fichiers), augmenter sensiblement. Il va alors s'agir de garantir voire améliorer la qualité de service et les performances du réseau. Concernant la disponibilité du service, les objectifs seront de diminuer les taux de coupure et de blocage des appels, et d'améliorer la couverture du réseau. Pour améliorer la capacité et la qualité de service du réseau, il sera étudié la possibilité de maximiser le gain de recombinaison (*MRC gain*), de

[*] Station de base dans les réseaux UMTS.

réduire la puissance allouée par lien, d'optimiser l'allocation des débits de transfert de données…

Pour pouvoir répondre de manière satisfaisante à l'augmentation du nombre de clients et du trafic, il va être nécessaire d'adapter le paramétrage des algorithmes RRM localement, c'est à dire au niveau des stations de base (ou *nodes B* dans un réseau UMTS, cf. annexes 10.1) :

- Adéquation à la demande locale en trafic (*hot spots*, zones commerciales), aux conditions particulières d'interférences (topographie de la zone), à la distance inter-site de design…
- Adaptation aux variations temporelles du trafic aussi bien sur de longues périodes qu'à l'échelle de la journée. Ainsi, on peut imaginer des mécanismes de réglage automatique des paramètres RRM en fonction de l'heure de la journée, du jour de la semaine, des jours fériés ou vacances scolaires, du trafic saisonnier.

Par ailleurs, en tenant compte du développement récent des réseaux d'accès radio-mobile, il apparaît que la taille et la complexité des réseaux cellulaires s'accroissent rapidement. L'optimisation manuelle du paramétrage devient un processus de plus en plus compliqué et coûteux. Par conséquent, les coûts des opérations de surveillance et d'administration (coûts d'OPEX) augmentent corrélativement pour les opérateurs. Il est donc essentiel de simplifier et d'automatiser ces tâches, ce qui permettra de réduire les moyens consacrés à l'optimisation manuelle des réseaux. De plus, en optimisant ainsi de manière automatique les performances et l'exploitation des réseaux mobiles déployés, il sera possible de retarder les opérations de sectorisation et densification du réseau et l'acquisition de nouveaux sites. L'ingénierie automatique permettra donc aussi d'étaler voire de réduire les investissements et les coûts de maintenance du réseau (coûts de CAPEX).

L'objectif global de cette thèse est de développer des mécanismes de paramétrage automatique (ou *self-tuning*) des algorithmes RRM dans les réseaux mobiles UMTS. Ainsi définie, l'ingénierie automatique est un processus qui ajustera certains paramètres des algorithmes RRM afin d'adapter le réseau mobile aux fluctuations de trafic, et d'améliorer la qualité de service perçue par les utilisateurs ainsi que les performances du réseau (en capacité, couverture…). Le fonctionnement général de l'ingénierie automatique est basé sur une boucle de régulation pilotée par un contrôleur. Ce contrôleur reçoit en entrées des indicateurs de qualité filtrés et produit en sortie un jeu de corrections de certains paramètres RRM (voir figure 1.1). Ces corrections sont appliquées aux nodes B concernés via des outils de paramétrage du réseau (comme le module RNO, *Radio Network Optimisation*, de l'OMC de Alcatel) :

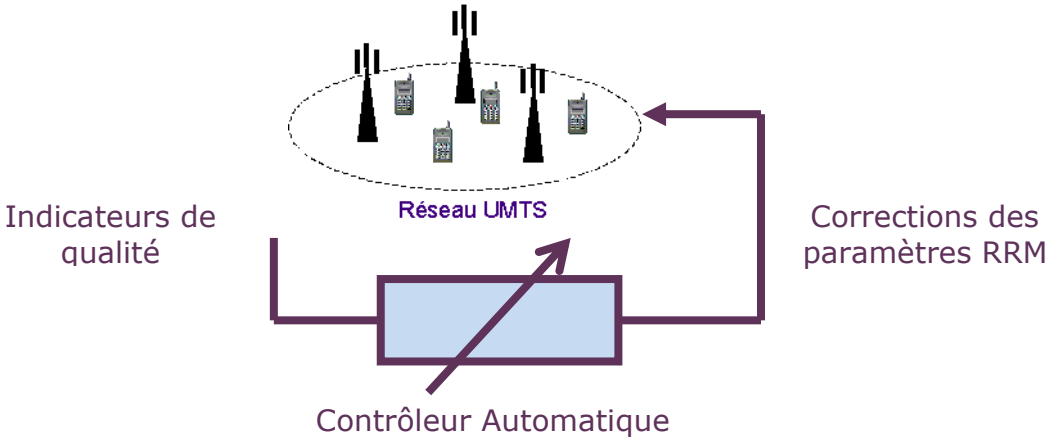


Figure 1.1 : Contrôleur automatique des paramètres RRM

Certaines méthodes décrites dans ce mémoire pourront ensuite être adaptées (plus ou moins directement) à d'autres techniques d'accès radio (CDMA*, GPRS/EDGE, voire WiFI™**). Dans un premier temps, nous avons choisi de faire la démonstration de ces concepts d'ingénierie automatique avec deux algorithmes RRM fondamentaux au fonctionnement des réseaux mobiles UMTS : le contrôle d'admission et la macro-diversité. Le chapitre 2 résumera les caractéristiques principales de ces deux algorithmes et explicitera les paramètres RRM et les indicateurs de performances que nous avons identifiés et sélectionnés dans une optique de paramétrage automatique du réseau.

Etapes vers l'ingénierie automatique des réseaux UMTS

La figure 1.1 résume la problématique de l'ingénierie automatique dans les réseaux mobiles : à partir d'indicateurs de qualité filtrés extraits du réseau, le contrôleur automatique doit déterminer des corrections à appliquer aux paramètres RRM. Les intervalles temporels de remontées de mesures et de modifications des paramètres, ainsi que la taille des filtres*** moyennant les mesures radio vont dépendre évidement des contraintes de finesse des équipements, mais surtout du type de fonctionnement souhaité par l'opérateur : ingénierie *on-line* (contrôle dynamique des paramètres) ou ingénierie *off-line* (paramétrage adapté localement mais modifié sur des échelles de temps beaucoup plus longues).

Au chapitre 2 l'analyse des algorithmes RRM de contrôle d'admission et de macro-diversité a permis d'identifier pour chaque algorithme les indicateurs significatifs de l'état des cellules ainsi que les paramètres dont une correction adéquate améliorera les performances du réseau. Par ailleurs, à l'aide de formulations analytiques des performances des réseaux UMTS, les impacts des principaux algorithmes RRM sur la capacité, la couverture et la probabilité de dysfonctionnements (ou *outage probability*) dans le réseau seront aussi étudiées rapidement. Cette approche quantitative sommaire soulignera l'influence d'un paramétrage correct des algorithmes de gestion des ressources radios sur les performances du réseau.

Dans le chapitre 3, nous allons tout d'abord modéliser le problème d'ingénierie automatique par une structure de Système d'Inférence Floue (SIF). Un tel système est composé par une base de règles de décisions qui fixent les règles de contrôle. Dans un système d'inférence floue, les mesures en entrées (réelles continues) sont transformées en valeurs floues (à l'aide de fonctions d'appartenance aux labels****) auxquelles vont s'appliquer les règles de décisions pour déterminer les valeurs floues en sorties. Dans la dernière étape, les valeurs floues de sortie sont transformées en actions réelles (discrète ou continue) que l'on appliquera ensuite au paramétrage du réseau. L'intérêt d'une telle modélisation est que notre problème d'ingénierie automatique entre parfaitement dans le cadre d'un SIF : les entrées et sorties du contrôleur sont réelles continues (métriques du réseau et corrections à appliquer aux paramètres RRM), et on ramène ainsi le problème du contrôle à un ensemble de règles de décisions élémentaires qui peuvent être déduites de l'expérience ou de l'analyse des experts radio. De plus, en énonçant notre problème de contrôle automatique sous la forme d'un SIF, cela permettra ensuite de lui appliquer des méthodes d'optimisation pour améliorer les performances des contrôleurs.

Dans la suite du chapitre 3, à partir du cadre général d'un SIF, nous allons définir la structure de nos contrôleurs sous la forme d'un processus *Fuzzy Knowledge Based Controller* ou FKBC.

---

* *Code Division Multiple Access* : technologie d'accès multiples en division par code.

** WiFI™ : contraction de *Wireless Fidelity*. Un réseau WiFI™ est un réseau répondant à la norme 802.11.

*** Filtre de moyenne glissante.

**** Un label est un terme qualitatif rendant compte de l'état d'un indicateur : par exemple *faible*, *moyen* et *fort* sont des labels qui permettront de qualifier le taux de blocage dans le réseau.

Les composant d'un processus FKBC comprennent un *Module de fuzzification* transformant les variables d'entrées continues en des labels linguistiques, une *Base de connaissance* qui se décompose en une *base de données*[*] et une *base de règles*[**], un *Moteur d'inférence* permettant de déterminer la valeur floue globale de sortie à partir des contributions individuelles de chaque règle élémentaire, et un *Module de defuzzification* transformant la valeur floue de correction en une valeur déterministe. Suite à cette modélisation de nos contrôleurs, nous les désignerons indifféremment dans l'ensemble du mémoire sous le terme de contrôleur automatique ou contrôleur de logique floue (CLF).

Simultanément à la définition de la structure des CLF, nous engagerons une réflexion sur deux aspects essentiels de l'introduction des contrôleurs automatiques dans les réseaux mobiles UMTS. Le premier point concerne les différentes échelles de temps relatives à un tel processus de contrôle. Le second point fondamental dans l'optique d'une mise en oeuvre opérationnelle est la définition d'une architecture permettant le contrôle automatique des paramètres RRM. Ainsi dans le chapitre 3, nous introduirons deux types d'architectures de paramétrage automatique des réseaux mobiles : un modèle centralisé et un modèle par agents distribués. Dans le premier modèle, les applications de surveillance et d'administration du réseau sont effectuées au niveau d'un nœud central de management. Une telle architecture est a priori plus adaptée à un fonctionnement de type *off-line*. Dans le second modèle, les applications de surveillance et de gestion des paramètres sont éclatées en sous-applications élémentaires qui sont réparties entre des nœuds actifs (ou agents) du réseau. Une telle architecture permettrait d'envisager un fonctionnement *on-line* du paramétrage automatique. Ces deux types d'architecture ne sont pas exclusives l'une de l'autre mais semblent plutôt correspondre à des calendriers différents de déploiement opérationnel : court ou moyen terme pour le modèle centralisé, long terme pour le modèle par agents distribués.

Dans le chapitre 4, nous aborderons la question de l'optimisation des contrôleurs de logique floue définis pour l'ingénierie automatique. En effet, l'adaptation du contrôleur est requise du fait de la variabilité des conditions d'application (mode *off-line*) ou de fonctionnement (mode *on-line*) du CLF, i.e. différents types d'environnements (conditions d'interférences, modèle de canal de propagation multi-trajets, distance inter-site), et de demande en trafic (conditions locales particulières, différentes échelles de variations temporelles du trafic heure / jour / mois), pour lesquelles un contrôleur peut s'avérer sous-optimal. Deux méthodes d'optimisation automatique ont été étudiées, adaptées et testées dans le contexte de l'ingénierie automatique des réseaux UMTS :

- L'optimisation par essaim de particules (OEP) appartient à la famille des méthodes d'optimisation itératives stochastiques. Avec cette méthode, la position de chaque particule correspond à un paramétrage particulier du CLF à optimiser. A l'initialisation, on définit un ensemble de particules (l'essaim) qui vont évoluer de façon partiellement aléatoire dans l'espace de recherche (ensemble des paramétrages possibles du CLF). Les paramètres de l'algorithme (la taille de l'essaim, le nombre de particules informatrices, les coefficients de confiance) vont être choisis ([Cle02]) afin de permettre la convergence en un temps de calcul acceptable et la sélection de contrôleurs efficaces et adaptés à un environnement particulier. La comparaison de l'efficacité des contrôleurs est faite à l'aide de fonctions coûts globales qui synthétisent les indicateurs de performances élémentaires des nodes B de la zone considérée.

[*] Contenant l'ensemble des informations relatives à l'environnement du contrôleur ainsi qu'aux variables et ensembles flous.

[**] Contenant l'ensemble des règles de paramétrage élémentaires énoncées sous la forme d'une proposition comportant un antécédent concernant l'état du système et une conséquence sur les paramètres RRM de la cellule.

- L'optimisation par apprentissage par renforcement (OAR) est fondée sur l'attribution de récompenses ou punitions en fonction des états atteints par le système sur lequel on applique le contrôle automatique. Un algorithme d'apprentissage typique est le *Q-learning*, mais il ne s'applique en principe que dans le cas de processus de décisions Markoviens (l'ensemble *S* des états du processus devant être discret). Nous avons donc adapté l'algorithme de *Q-learning* au système d'inférence floue introduit au chapitre 3. Le SIF permet ainsi de transformer les entrées continues du contrôleur en des valeurs floues discrétisées sur lesquelles on appliquera le *Q-learning*. Cet algorithme se décompose en deux phases principales : une phase d'exploration et une phase d'exploitation. Durant la phase d'exploration, l'algorithme va prendre des décisions (de corrections des paramètres RRM) partiellement aléatoires qui vont ensuite, à l'aide d'une boucle de rétroaction et des valeurs de renforcements, lui permettre de distinguer les corrections efficaces des corrections inutiles en fonction de l'état de la cellule et de ses voisines. Une fois que l'on considère que l'algorithme a suffisamment visité les différents états du système, on passe à la phase d'exploitation. Dès lors, à chaque itération, l'algorithme va sélectionner systématiquement la "meilleure" correction à appliquer au paramétrage du node B. Contrairement à l'algorithme d'OEP où on agrégeait les indicateurs des différentes cellules dans les fonctions coûts en n'exploitant pas une partie importante de l'information, l'algorithme de *Q-learning* va tenir compte, à chaque itération, de l'état, du renforcement reçu et de la correction appliquée au niveau de chaque cellule de la zone considérée : on peut dire que le contrôleur va apprendre simultanément sur l'ensemble des nodes B.

Ces deux méthodes d'optimisation ont été développées et testées numériquement à l'aide d'un simulateur semi-dynamique de réseaux mobiles, développé en interne par France Télécom Recherche & Développement, et de deux modules additionnels de paramétrage automatique et d'optimisation.

Dans le chapitre 5, nous décrirons les deux types de fonctionnement envisagés pour l'ingénierie automatique : les modes *off-line* et *on-line*.

Le mode *off-line* (hors-ligne) permettra le réglage local (au niveau du node B) des paramètres RRM. En effet, à l'aide d'un simulateur de réseau UMTS et d'un module logiciel de paramétrage RRM des stations, il sera possible de déterminer les modifications de paramétrage RRM les plus efficaces pour une demande en trafic donnée sur la zone géographique considérée (homogène comme inhomogène). On pourra ensuite vérifier sur le terrain l'efficacité des corrections effectuées en termes de qualité de service perçue et de performances du réseau (capacité, couverture…), en comparant les métriques radio avant et après modifications des paramètres RRM pour des périodes correspondant à des niveaux et types de trafic équivalents : cela constitue une boucle de validation de l'approche *off-line*. Ce mode de fonctionnement aura l'avantage d'être envisageable dans un futur proche, en modifiant le paramétrage RRM au niveau de l'OMC (*Operations & Management Controller*) en UMTS, et pourra aussi être appliqué à d'autres réseaux d'accès radio existants (GSM / GPRS / EDGE, CDMA). Enfin avec le mode *off-line*, il sera aussi possible de prévoir la détermination (par des simulations validées ensuite par des tests) et l'utilisation de plusieurs jeux de paramètres RRM typiques, node B par node B, en fonction du jour de la semaine, de la période de l'année (vacances), voire d'évènements exceptionnels mais récurrents (salons, grandes manifestations musicales ou sportives…).

A plus long terme, on envisage un deuxième type de fonctionnement de l'ingénierie automatique; le mode *on-line* ou ingénierie dynamique. Comme décrit au chapitre 3, la solution de paramétrage dynamique considérée est fondée a priori sur l'utilisation de modules

de contrôle de logique floue (CLF) intégrés à des agents distribués dans le réseau. Ce mode de fonctionnement permettra d'ajuster dynamiquement certains paramètres* contrôlant les algorithmes RRM, afin d'améliorer la qualité de service et les performances du réseau mobile et de l'adapter aux fluctuations de trafic. Le processus d'ingénierie dynamique est basé sur une boucle de régulation qui est dirigée par le contrôleur de logique floue. Ce contrôleur reçoit dynamiquement en entrées les indicateurs de qualité de la cellule et de ses proches voisines, et produit en sortie un jeu de corrections des paramètres RRM de la station centrale (voir figure 1.2). On envisage deux stratégies de modification des paramètres RRM, à chaque itération : a) modifier simultanément toutes les stations du réseau, b) modifier une seule station par groupe d'interférence de stations (ou *clusters*). Au chapitre 5, nous verrons comment définir des groupes d'interférence disjoints à partir de données d'interférence et d'une méthode systématique d'agrégation de sous-groupes (ou *clustering*) [Ben05].

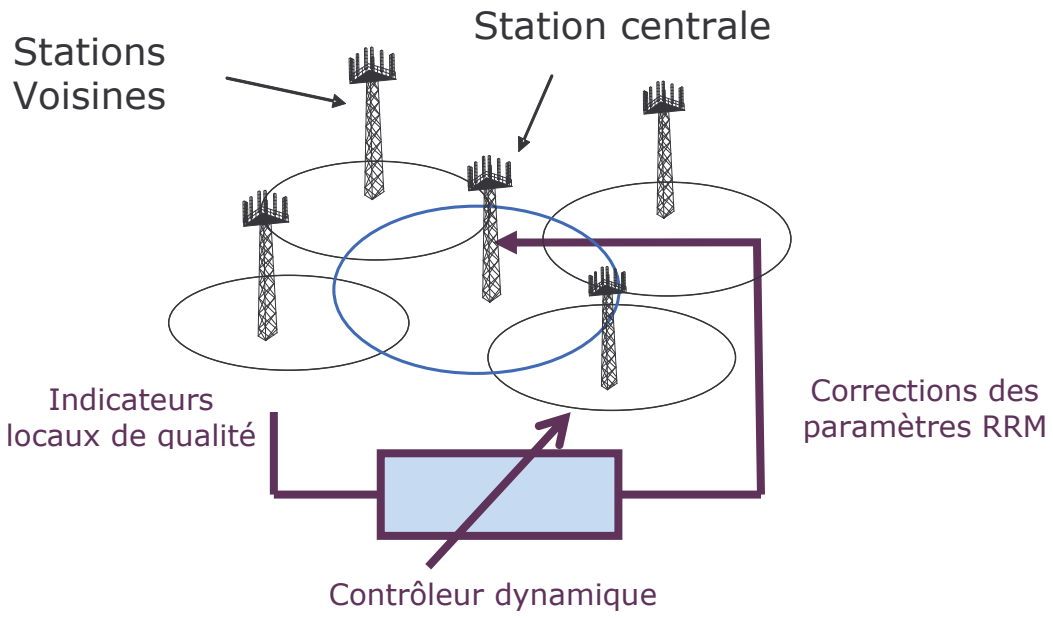


Figure 1.2 : Contrôleur dynamique de logique floue des paramètres RRM de la station centrale**.

Enfin, dans le cas des contrôleurs de logique floue utilisant l'apprentissage par renforcement, nous avons imaginé un mode de fonctionnement intermédiaire – "semi *on-line*". Nous supposons que nous avons calibré (phase d'exploration) à l'aide de simulations logicielles un contrôleur fondé sur l'algorithme de *Q-learning* comme décrit au chapitre 4. Un contrôleur de ce type est introduit au niveau de l'OMC pour administrer un ensemble de nodes B. La fréquence des remontées de mesures ainsi que des corrections des paramètres RRM vont dépendre des contraintes liées aux équipements du réseau, mais aussi de la politique de l'opérateur (demi-heure, heure, demi-journée, journée). La nouveauté est qu'au lieu d'appliquer le *Q-learning* dans son mode d'exploitation, on continue de l'utiliser en mode d'exploration (avec une probabilité ε faible de ne pas choisir l'action optimale, et un contrôle éventuel de l'action choisie aléatoirement). Ainsi à chaque itération, le module va recevoir pour l'ensemble des nodes B le renforcement correspondant à l'état de la cellule et la correction appliquée (à l'itération précédente). Ainsi, le module pourra mettre à jour les

* Comme le seuil de charge cible *downlink* (lien radio descendant, c'est à dire de la station vers le mobile) pour le contrôle d'admission ou les fenêtres d'ajout / de retrait de liens de MD pour la macro-diversité

** La station centrale est la station dont on veut modifier le paramétrage, en tenant compte de l'état des stations qui lui sont proches (ses voisines). Chacune des stations dans la zone à optimiser est tour à tour considérée comme station centrale.

fonctions de qualité correspondant aux situations des différents nodes B. Cette méthode permettra donc d'adapter à un environnement réel des contrôleurs calibrés seulement à partir de simulations logicielles.

Dans les chapitres 6 et 7, nous présenterons les principaux résultats de contrôle automatique, en modes *on-line* et *off-line*, obtenus sur des environnements de simulations réalistes[*]. A partir d'un paramétrage uniforme et statique des stations, nous montrerons comment les contrôleurs de logique floue améliorent les performances du réseau en adaptant les paramètres des algorithmes d'admission et de macro-diversité. Les performances données par le paramétrage initial des stations seront représentées sous formes de fronts de paramétrage uniforme[**]. Pour compléter cette analyse des performances effectuée dans le plan des indicateurs de performance, les causes de blocage et de coupure des appels seront identifiées, analysées et comparées avec et sans activation des contrôleurs, en modes *on-line* et *off-line*. Il sera ainsi plus facile de mettre en évidence les avantages et défauts des différentes politiques de contrôle.

Nous produirons ensuite des exemples d'améliorations de performances des contrôleurs obtenus respectivement par l'optimisation par essaim de particules et par l'apprentissage par renforcement. Pour l'algorithme de *Q-learning*, nous avons seulement testé l'utilisation d'un SIF[***] unique dans les phases d'exploration et d'exploitation. En revanche avec l'OEP, quand les CLF contrôlent plusieurs algorithmes RRM, l'optimisation des contrôleurs de chaque paramètre RRM pourra se faire soit séquentiellement[****] soit simultanément. L'optimisation par essaim de particules sera testée dans les modes mono et multi-objectif. Le front optimal du nuage des solutions explorées en multi-objectif sera ainsi comparé avec les meilleures solutions obtenues en mono-objectif. Enfin, l'efficacité et les spécificités des deux méthodes d'optimisation décrites dans ce mémoire – par essaim de particules et par l'algorithme de *Q-learning* – seront testées et comparées sur deux réseaux de taille et de caractéristiques différentes respectivement dans les villes 1 et 2. Nous présenterons et comparerons des exemples de politiques de contrôle optimisées soit par essaim de particules soit par l'algorithme de *Q-learning*.

[*] Sur les zones de la ville 1 (réseau de 21 secteurs) et la ville 2 (réseaux de 109 et 32 secteurs).
[**] Dans le plan taux de coupure – taux de blocage par exemple.
[***] Un système d'inférence floue (SIF) contrôlant un seul paramètre RRM.
[****] Optimisation d'un seul contrôleur de paramètre RRM par simulation.

2. Algorithmes RRM et performances des réseaux UMTS

Dans l'optique de développer et introduire des méthodes de contrôle automatique dans les réseaux UMTS, en mode *on-line* et *off-line*, deux algorithmes de gestion de la ressource radio ou RRM, essentiels dans la technologie W-CDMA vont être analysés dans ce chapitre. Tout d'abord, on considère le contrôle d'admission qui gère l'acceptation ou le blocage des nouveaux appels dans le réseau UMTS. C'est le principal algorithme avec le contrôle de charge (*Congestion Control*) et le *Packet Scheduling* qui permet de contrôler le niveau de charge dans le réseau [Hol02]. Puis, nous allons étudier l'algorithme de macro-diversité qui gère la mobilité des terminaux dans le réseau, c'est-à-dire la prise en charge des communications par les stations de la zone du réseau, afin d'éviter la coupure des appels. Pour chaque algorithme, il s'agit de déterminer des indicateurs significatifs de l'état des cellules et d'identifier les paramètres RRM que l'on pourra modifier afin d'améliorer l'état des indicateurs sélectionnés et plus généralement les performances du réseau. De plus, nous allons mentionner les autres paramètres radio du réseau UMTS auxquels on pourrait appliquer des méthodes de contrôle automatique. Et à plus long terme, il est aussi possible d'élargir ces méthodes au cas multi-système, en contrôlant par exemple les algorithmes de sélection, resélection ou de *load balancing* ([Tol02a]) dans une approche intégrant les systèmes UMTS, GSM/GPRS et WLAN. Dans la suite du chapitre, nous allons introduire des descriptions analytiques des problèmes de capacité et couverture dans les réseaux UMTS. Ces formulations analytiques vont nous permettre d'induire certaines conséquences qualitatives du paramétrage RRM sur les performances du réseau UMTS.

2.1. Deux algorithmes RRM à optimiser

Pour développer et tester nos méthodes d'ingénierie automatique des réseaux UMTS, deux algorithmes RRM de gestion de trafic ont tout d'abord été choisis et étudiés : le contrôle d'admission et la macro-diversité. En effet, ces deux algorithmes sont relativement stables et indépendants des constructeurs, contrairement par exemple à l'algorithme de contrôle de charge, et constituent des éléments essentiels de la technologie W-CDMA : gestion de l'accès au service et du niveau de charge au niveau des nodes B pour le contrôle d'admission, gestion de la mobilité et de l'interférence entre les voisines pour la macro-diversité. Les performances des réseaux UMTS seront ainsi très sensibles au paramétrage de ces algorithmes. Dans le cadre de ce mémoire, nous souhaitons faire la preuve que l'ingénierie automatique permet d'adapter efficacement le paramétrage de la macro-diversité et du contrôle d'admission aux fluctuations du trafic et aux contraintes locales d'environnement. Ce processus devrait ainsi améliorer la qualité de service et les performances des réseaux mobiles UMTS. Par la suite, il sera important d'étendre cette approche aux autres algorithmes RRM de gestion de trafic [Sal03], notamment aux algorithmes A-RRM (*Advanced Ressource Radio Management*) en cours de développement dans la communauté scientifique.

2.1.1. L'algorithme de contrôle d'admission

Le contrôle d'admission est un algorithme fondamental de la gestion de trafic dans les réseaux UMTS : pour chaque nouvel appel (voix ou données), cet algorithme détermine en fonction du type de client, du service demandé et des ressources radio disponibles sur le lien radio montant (ou *uplink* : du mobile vers la station), et le lien descendant (ou *downlink* : de la station vers le mobile), si cet appel doit être admis dans le réseau ou bloqué. Un paramétrage efficace et adapté localement de cet algorithme est donc essentiel, car il détermine l'accès ou

non des utilisateurs au service et conditionne la perception par les clients de la disponibilité du service pour le réseau mobile. Dans une zone géographique soumise à des contraintes de propagation radio particulières ou ayant une densité d'utilisateurs élevée, les clients n'admettront pas d'avoir un taux d'accès au service beaucoup plus faible que les autres utilisateurs du réseau. Ainsi dans le but d'optimiser, station par station, l'algorithme de contrôle d'admission des appels[*], plusieurs stratégies ont déjà été proposées : prise en compte du niveau d'interférence dans le CAC [Cap01], établissements de priorité entre les services à l'admission des appels [Per02], modélisation du problème d'admission par un système d'inférence floue fondé sur des observables du système [Din04] (congestion de la cellule, niveau d'interférence) ou sur des statistiques des appels [Cha04] (durée d'un appel, temps moyen passé dans une cellule).

Paramètres du contrôle d'admission

L'algorithme de contrôle d'admission bloque ou accepte dans le réseau UMTS les nouveaux appels en fonction du type de client, du service demandé et des ressources radio disponibles. Cet algorithme conditionne aussi le maintien ou la coupure des appels en communication. Sur le lien descendant, une première version du contrôle d'admission utilise deux paramètres principaux :

- Le seuil de charge cible *downlink* (ou *load target threshold*) : ce paramètre est défini au niveau de la cellule et détermine l'admission de nouveaux appels dans le réseau UMTS. Quand la charge du node B est en dessous de ce seuil, tout nouvel appel est admis dans le réseau. Au-dessus de ce seuil, les nouveaux appels sont bloqués à l'admission (blocking).
- La charge maximum *downlink* acceptable (ou *maximum acceptable load*) : ce paramètre est défini au niveau de la cellule et détermine le maintien des appels en communication dans le réseau UMTS. Quand la charge du node B est en dessous de ce seuil, aucun appel n'est coupé. Au-dessus de ce seuil, certains appels en communication sont coupés (*dropping*) dans le réseau. Quand la charge *downlink* est au-dessus de la charge cible mais en dessous de la charge maximum, tous les nouveaux appels sont bloqués à l'admission, mais aucun appel en communication n'est coupé dans le réseau.

Modification des paramètres de l'algorithme

Les impacts d'une modification du paramètre charge cible *downlink* sont (voir figure 2.1) :

- Augmenter la charge cible *downlink* d'une cellule permet d'augmenter le nombre de nouveaux appels admis dans le réseau et donc de diminuer le taux de blocage à l'admission.
- Diminuer la charge cible *downlink* d'une cellule permet de diminuer la charge de la cellule (en augmentant le taux de blocage à l'admission) mais favorise le maintien des appels en communication et donc réduit le taux de coupure.

Les impacts d'une modification du paramètre de charge maximum *downlink* acceptable sont :

- Augmenter ce paramètre permet de réduire dans un premier temps le taux de coupure dans la cellule. Mais le risque est important d'atteindre plus fréquemment des cas de saturation du lien descendant (100% de charge au node B) ce qui aurait pour conséquences une hausse brutale du nombre d'appels bloqués et coupés.

[*] ou CAC : *Call Admission Control.*

- Diminuer ce paramètre permet de réduire la probabilité d'une saturation du lien descendant dans la cellule mais risque d'en augmenter le taux de coupure.

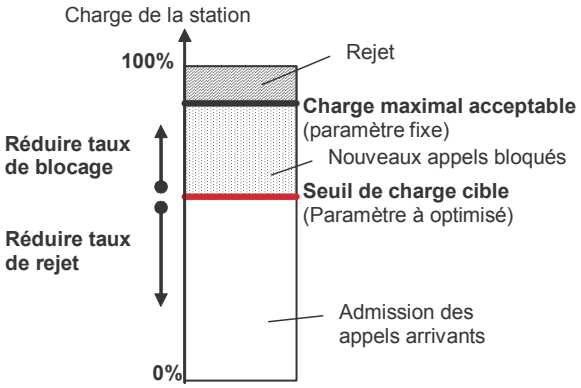


Figure 2.1 : Conséquences d'une modification de la charge cible *downlink* sur le taux de coupure et le taux de blocage

Choix du paramètre à modifier dans le cadre de l'ingénierie automatique

L'objectif de l'ingénierie automatique est d'améliorer la qualité de service et les performances du réseau en appliquant une politique de l'opérateur du réseau mobile permettant d'atteindre un compromis optimal en termes de blocage et de coupure. Sur ce point, les opérateurs préfèrent généralement bloquer des appels à l'admission plutôt que couper des appels en communication. On retient le paramètre de seuil de charge cible *downlink* parmi les paramètres modifiables en sortie des contrôleurs automatiques. En effet comme décrit sur la figure 2.1, la valeur de ce paramètre a une incidence directe sur les taux de blocage et de coupure dans le réseau. Par ailleurs, le paramètre de charge maximum *downlink* acceptable sera fixé pour l'instant à une valeur typique (correspondant au paramétrage a priori des constructeurs et confirmée par des tests sur plate-forme expérimentale). Par la suite on pourra envisager d'optimiser simultanément, node B par node B, les deux paramètres de seuil de charge cible et de charge maximum *downlink*.

Indicateurs de qualité pour l'algorithme de contrôle d'admission

Comme décrit sur la figure 1.1, le contrôleur automatique doit recevoir en entrées des indicateurs de qualité filtrés. Il est donc nécessaire d'établir pour chaque algorithme RRM la liste des indicateurs de qualité pertinents permettant un contrôle efficace de son paramétrage. Nous indiquons ci-dessous une liste de différents indicateurs qui pourraient s'avérer utiles pour l'étude de l'impact du paramétrage automatique de l'algorithme de contrôle d'admission sur les performances du réseau. Toutes les métriques énumérées ci-dessous sont des indicateurs calculés au niveau de la cellule :

- Charge *downlink* : pourcentage de puissance émise par le node B.
- Charge *uplink* : calculée à partir de l'interférence totale UL* reçue à la station [Sha94] [Hol02] :

$$X_{UL} = \frac{I_{Total}^{UL} - N_0}{I_{Total}^{UL}} \qquad (2.1)$$

- Taux de coupure *uplink* (UL *dropping rate* ou *outage*).

* UL : *Uplink* ou lien montant, c'est à dire du mobile vers la station.

- Taux de coupure *downlink* (DL[*] *dropping rate* ou *outage*).
- Taux de coupure global par node B (tenant compte des coupures DL et UL).
- Taux de coupure CM couverture UL, taux de coupure CM couverture DL, taux de coupure CM charge UL, taux de coupure CM charge DL (CM signifiant contrôle du maintien) : les taux de coupure son classés selon les différentes causes identifiées. Par exemple, une coupure de type CM couverture a pour cause une perte de couverture causée par exemple par un CIR cible (*Carrier to Interference Ratio*) non atteint pour le service considéré ou par un signal reçu du canal pilote CPICH[**] trop faible. Une coupure de type CM charge a pour cause une surcharge du lien lors du contrôle de charge.
- Taux de blocage *uplink* à l'admission de nouveaux appels (ne tient pas compte des blocages à la création de liens de macro-diversité).
- Taux de blocage *downlink* à l'admission de nouveaux appels (ne tient pas compte des blocages à la création de liens de macro-diversité).
- Taux de blocage global par node B (tenant compte des blocages DL et UL).

Le compromis entre coupure et blocage

Une stratégie d'optimisation des paramètres de l'algorithme de contrôle d'admission devra se fonder sur un compromis entre les phénomènes de coupure et de blocage. Il est donc essentiel de compter à la fois des métriques de coupure et de blocage parmi les indicateurs de QoS en entrées des contrôleurs automatiques. Dans cette optique, nous avons sélectionné les indicateurs de taux de coupure et de taux de blocage par station. Ces métriques permettent d'estimer l'état de la cellule centrale et de ses voisines. Ces indicateurs seront fournis en entrées du contrôleur automatique qui ajustera le seuil de charge cible DL.

Exemple de l'impact du contrôle automatique du paramètre de seuil de charge cible

Sur la figure 2.2 ci-dessous, sont indiquées les variations des taux de blocage (en jaune) et taux de coupure (en rose) du node B en fonction du seuil de charge cible DL (en bleu), en supposant un contrôle *on-line* de ce paramètre.

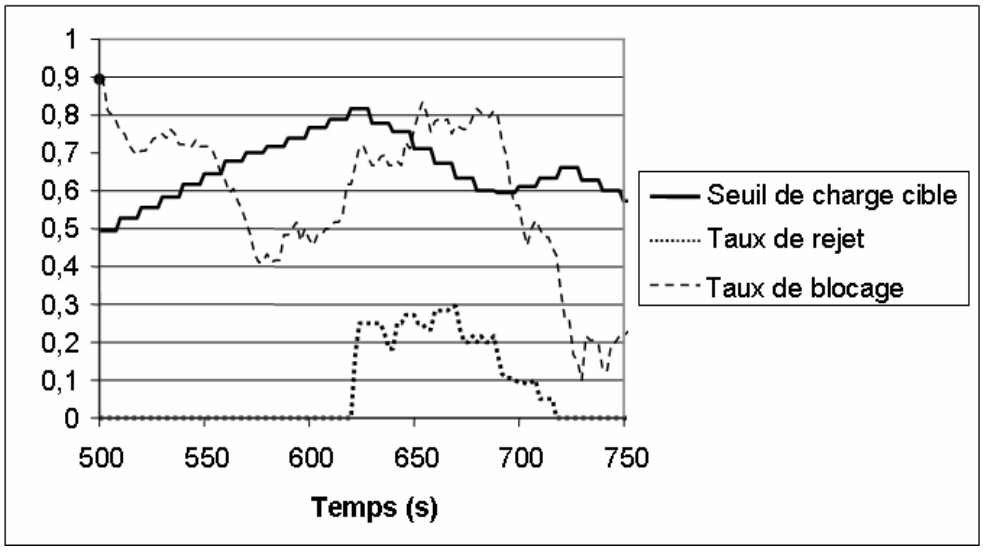


Figure 2.2 : Variations du seuil de charge cible et impacts sur les taux de blocage et taux de coupure du node B.

[*] DL : *Downlink* ou lien montant, c'est à dire de la station vers le mobile.
[**] *Common Pilot Channel* : Canal commun pilote, sur le lien descendant.

Dans la première partie de la courbe, on assiste à un fonctionnement “normal” du paramétrage automatique : le taux de blocage diminue avec l’augmentation du seuil de charge cible *downlink*. Etant donné que nous avons accepté trop de mobiles dans la cellule (entre les temps t_0=500 s et t_1=620 s), la charge du node B est devenue supérieure au paramètre de charge maximum *downlink* acceptable. Dés lors, le contrôle d’admission commence à couper des communications, ainsi qu’à bloquer certains nouveaux appels (en dépit de la charge cible très élevée). A ce moment-là, le contrôleur de logique floue décide de réduire la charge cible *downlink* afin, dans un premier temps, de diminuer le taux de coupure (quitte à augmenter temporairement le taux de blocage). Une fois le taux de coupure revenu sous un seuil acceptable, le contrôleur augmente à nouveau le seuil de charge cible pour diminuer le taux de blocage.

2.1.2. L’algorithme de macro-diversité

La macro-diversité est un second algorithme fondamental de la gestion du trafic dans les réseaux UMTS : pour chaque utilisateur se déplaçant d’une cellule vers une de ses voisines ou immobile en bordure de cellule, l'algorithme détermine régulièrement les liens à créer entre le mobile et les stations avoisinantes. Un paramétrage efficace et adapté localement de cet algorithme est donc essentiel, car il détermine la continuité du service pour les mobiles se déplaçant sur plusieurs cellules. ***En outre, cela permet d’améliorer la couverture des zones problématiques, et ne détériore pas la capacité radio de la zone (contrairement à un mauvais paramétrage de la macro-diversité qui peut permettre la création de liens inutiles et consommateurs en puissance DL, notamment pour les mobiles hauts débits nécessitant des puissances de transmission importantes au node B)***. Ainsi dans le but d'optimiser l'algorithme de macro-diversité, plusieurs stratégies ont déjà été proposées, comme la modélisation du problème de macro-diversité par un système d'inférence floue fondé sur des observables du système (charge des cellules, niveau d'interférence) et sur des données par mobile (nombre de pilotes reçus). Le système d'inférence floue adapte les paramètres de taille de fenêtre d'ajout d'une station à l'*Active Set* d'un mobile – *Add_win* – et de taille de fenêtre de retrait d'une station à l'*Active Set – Drop_win* [Hom00] [Hom04].

Paramètres de l’algorithme de macro-diversité

Les paramètres contrôlant l’algorithme de macro-diversité sont essentiels dans une optique d’ingénierie automatique car la gestion du *hand-over* a un impact important sur les performances des réseaux UMTS. En effet, le *hand-over* améliore la couverture en bordure de cellules et accroît la capacité *uplink*, en permettant la diminution du niveau de signal UL émis par le mobile et donc du niveau global d'interférence. Dans la plupart des cas, le *hand-over* améliore aussi la qualité du lien *downlink* grâce à la recombinaison des signaux (MRC ou *Maximum Ratio Combining*). Le paramétrage du *hand-over* a aussi des conséquences sur la capacité radio cellulaire car cet algorithme contrôle l’ajout et la suppression des liens radio dans les cellules (ainsi que le contrôle d’admission et le contrôle de charge). En particulier, le *soft hand-over*[*] (ou SHO) impacte directement la capacité *downlink*: en effet, les interférences sur le lien *downlink* augmentent avec le nombre moyen de liens par utilisateur, et par conséquent trop de mobiles en situation de SHO feront diminuer la capacité du lien *downlink*.

Le paramétrage des algorithmes de *hand-over* devra donc être fixé avec précaution car il a des effets importants sur les performances du réseau. Plusieurs paramètres de cet algorithme peuvent être réglés pour modifier par exemple le nombre d’utilisateurs en SHO. Il est

[*] Situation où deux liens DL sont créés par deux nodes B voisins vers le même mobile et transportant la même information.

important de noter que ces paramètres dépendent de l'implémentation des constructeurs, car les algorithmes de *hand-over* sont propriétaires. Cependant, les algorithmes proposés par les équipementiers dans leurs premières versions gérant la macro-diversité sont très similaires et se fondent en général sur les évènements *1A* (addition d'un nouveau lien dans l'*Active Set*), *1B* (suppression d'un lien dans l'*Active Set*) et *1C* (remplacement du plus mauvais lien de l'*Active Set* par un nouveau lien) définis dans la norme. Ces évènements peuvent être basés sur les mesures de CPICH Ec/Io (Energie reçue du canal pilote CPICH par chip divisée par la densité de puissance dans la bande [3GPP3]) ou de CPICH RSCP (*Received Signal Code Power* – puissance reçue sur le code du canal pilote CPICH [3GPP3]). Par conséquent, nous indiquons ci-dessous les paramètres qui, d'après nous, pourraient être réglés de manière à optimiser les taux de *hand-over* dans le réseau :

- Le paramètre *Add_win* est la taille de la fenêtre d'ajout d'une station à l'*Active Set* d'un mobile. Ce seuil est utilisé pour l'évènement *1A* sur la différence en décibels entre la métrique choisie (Ec/Io ou RSCP) de la meilleure serveuse et la métrique de la station considérée, en tenant compte pour celle-ci du paramètre d'*offset* par cellule, CIO[*]:

$$\left(Ec/Io\right)_{CPICH}^{Best\ station} - \left(\left(Ec/Io\right)_{CPICH}^{Station} + CIO^{Station}\right) \leq Add_win \qquad (2.2)$$

- Le paramètre *Drop_win* est la taille de la fenêtre de retrait d'une station de l'*Active Set* d'un mobile. Ce seuil est utilisé pour l'évènement *1B* sur la différence en décibels entre la métrique choisie (Ec/Io ou RSCP) de la meilleure serveuse et la métrique de la station considérée, en tenant compte pour celle-ci du paramètre d'*offset* par cellule, CIO :

$$\left(Ec/Io\right)_{CPICH}^{Best\ station} - \left(\left(Ec/Io\right)_{CPICH}^{Station} + CIO^{Station}\right) \geq Drop_win \qquad (2.3)$$

- Le paramètre *Rep_win* est la taille de la fenêtre de remplacement d'une station de l'*Active Set* d'un mobile. Ce seuil est utilisé pour l'évènement *1C* sur la différence en décibels entre la métrique choisie (Ec/Io ou RSCP) de la meilleure serveuse et la métrique de la station considérée, en tenant compte pour celle-ci du paramètre d'*offset* par cellule, CIO.

$$\left(\left(Ec/Io\right)_{CPICH}^{Station} + CIO^{Station}\right) - \left(Ec/Io\right)_{CPICH}^{In\ AS} \geq Rep_win \qquad (2.4)$$

- La taille maximum de l'*Active Set* (liste des nodes B en macro-diversité avec le terminal).
- Le *timer* de retrait: quand il est activé, ce *timer* permet de déclencher l'évènement *1B*.
- Le coefficient de pondération pour l'*Active Set* (paramètre *W*) : quand *W=0* (valeur par défaut), l'*Active Set* est mis à jour par comparaison à la meilleure cellule de l'*Active Set* (équations 2.2, 2.3 et 2.4). Sinon ce coefficient pondère la prise en compte de la valeur en dB de la meilleure cellule et de la valeur en dB de la somme des métriques des stations de l'*Active Set*: $10 \times \log\left(\sum_{1}^{N_{AS}} M_i\right)$, où N_{AS} est la taille de l'*Active Set*.
- L'*offset* individuel par cellule (*CIO* ou *Cell Individual Offset*) : ce paramètre permet de moduler cellule par cellule la prise en compte des seuils *Add_win* et *Drop_win*.
- La liste des cellules voisines…

<u>Remarques sur le réglage de ces différents paramètres de macro-diversité:</u>

- Pour réduire ou augmenter le nombre de mobiles en *hand-over* par cellule, il est plus efficace de modifier à la fois les fenêtres d'ajout et de retrait, de telle façon que les modifications concernent aussi bien les nouveaux liens créés que les liens existants présents dans l'*Active Set*.

[*] *Cell Individual Offset* : *Offset* individuel par cellule.

- Conformément à la norme, les paramètres impliqués dans l'algorithme de macro-diversité sont envoyés à chaque terminal par un message dédié de *measurement control* ou diffusés par les messages SIB11/12 [3GPP4]. Dans l'implémentation des constructeurs, les paramètres peuvent généralement être fixés par cellule : tous les mobiles d'une cellule utilisent donc les mêmes paramètres, quelles que soient les cellules voisines dont ils mesurent la qualité du canal pilote. Cependant les mobiles appartenant à une autre cellule peuvent recevoir des valeurs différentes de ces paramètres. Par conséquent, cela pose le problème particulier des mobiles en *soft hand-over*: pour certains constructeurs, les mobiles utilisent les paramètres fournis par la cellule où ils ont commencé leur appel. Cependant, si cette cellule quitte l'*Active Set*, le mobile utilisera alors les paramètres de la meilleure cellule de l'*Active Set*.
- Par conséquent pour ces constructeurs, dans le cadre d'une ingénierie automatique, il est inefficace de changer les fenêtres de macro-diversité d'une seule cellule, car cela n'impactera ni les mobiles entrant dans la cellule par *soft hand-over*, c'est à dire les cas d'ajout de la cellule considérée aux *Active Sets* des mobiles proches, ni les mobiles ayant cette cellule dans leur *Active Set*, mais cette cellule n'étant ni celle où ils ont commencé leur appel ni leur meilleure cellule.
- Le paramètre *offset* individuel par cellule (*Cell Individual Offset*) est un paramètre associé à chaque cellule de la liste des voisines. Il est ajouté aux valeurs mesurées de chaque cellule avant toute évaluation des différents évènements. Ce paramètre permet de favoriser ou désavantager l'entrée ou la sortie d'une cellule de l'*Active Set*, en modifiant indirectement les fenêtres d'ajout et de retrait pour cette cellule seulement. Si ce paramètre est non nul pour une cellule donnée, cela n'impactera pas le comportement des mobiles de la cellule ayant de nouveaux liens à créer. En revanche, cela aura un impact sur les mobiles entrant en macro-diversité avec la cellule considérée depuis n'importe quelle autre cellule, c'est à dire l'ajout de cette cellule dans tout *Active Set*. Ce paramètre peut permettre d'optimiser la macro-diversité dans une cellule particulière. Mais il est important de savoir comment et sur quelle échelle de temps, ce paramètre pourrait être mis à jour dans un réseau opérationnel, selon l'implémentation du constructeur.
- Les paramètres décrits ci-dessus permettent d'adapter les taux de macro-diversité sur une zone de plusieurs stations. En revanche, dans le cas d'une zone très limitée de surcharge (une seule cellule par exemple), il est difficile d'utiliser ces paramètres localement, car les effets risquent de se limiter à certains mobiles de la cellule.

Choix du paramètre de fenêtre d'ajout pour l'ingénierie automatique

L'objectif de l'ingénierie automatique est d'améliorer la qualité de service et les performances du réseau en définissant une politique de l'opérateur du réseau mobile permettant d'adapter efficacement node B par node B le paramétrage de la macro-diversité. Dans ce but, on retient dans un premier temps le paramètre de fenêtre d'ajout (*Add_win*) parmi les paramètres modifiables en sortie des contrôleurs automatiques. En effet, la valeur de ce paramètre a une incidence directe sur les indicateurs de qualité du réseau :

- Un paramètre de fenêtre d'ajout élevé implique un grand nombre d'utilisateurs en situation de *hand-over*, et par conséquent améliore la couverture *uplink*, augmente la qualité *downlink* dans une certaine mesure et aussi réduit le taux de blocage, au moins sur l'*uplink* (car les mobiles ont besoin de moins de puissance sur le lien UL). Mais la capacité *downlink* risque dans le même temps de diminuer, ce qui se traduit par l'augmentation du taux de coupure *downlink*. De plus la charge de signalisation peut devenir importante.
- Au contraire, un paramètre de fenêtre d'ajout faible implique moins de mobiles en situation de *hand-over*, et par conséquent une possible diminution de la qualité *uplink*,

ainsi qu'une augmentation de la puissance *uplink* transmise par les mobiles, causant un taux de blocage plus fort sur ce lien et finalement une perte de couverture.

Le paramètre que l'on choisit d'utiliser dans le cadre de l'ingénierie automatique de l'algorithme de macro-diversité est donc la fenêtre d'ajout *Add_win*. Par la suite, il sera intéressant de tester l'efficacité d'un contrôle automatique du paramètre de *CIO*.

Indicateurs de qualité pour la macro-diversité

Comme décrit sur la figure 1.1, le contrôleur automatique doit recevoir en entrées des indicateurs de qualité filtrés. Il est donc nécessaire d'établir pour chaque algorithme RRM la liste des indicateurs de qualité pertinents permettant un contrôle efficace de son paramétrage. Nous indiquons ci-dessous différents indicateurs de performance qui pourraient être utilisés dans le cadre du paramétrage automatique de l'algorithme de macro-diversité. Toutes les métriques énumérées ci-dessous sont des indicateurs calculés au niveau de la cellule :

- Charge *downlink* : pourcentage de puissance émise par le node B.
- Charge *uplink* : calculée à partir de l'interférence totale reçue à la station.

- Taux de coupure *uplink* (*UL dropping rate* ou *outage*) due à la couverture ou à la charge.
- Taux de coupure *downlink* (*DL dropping rate ou outage*) due à la couverture ou à la charge.
- Taux de coupure global par node B (tenant compte des coupures DL et UL).
- Taux de coupure CM couverture UL, taux de coupure CM couverture DL, taux de coupure CM charge UL, taux de coupure CM charge DL (CM signifiant contrôle du maintien) : les taux de coupure son classés selon les différentes causes identifiées.

- Taux de blocage *uplink* à l'admission de nouveaux appels (ne tient pas compte des blocages à la création de liens de macro-diversité).
- Taux de blocage *downlink* à l'admission de nouveaux appels (ne tient pas compte des blocages à la création de liens de macro-diversité).
- Taux de blocage global par node B (tenant compte des blocages DL et UL).
- Taux de blocage des liens de MD (macro-diversité) *downlink*.
- Taux de blocage total *downlink* tenant compte de tous les blocages DL.

- Les pourcentages de mobiles en macro-diversité à 1, 2 ou 3 liens.
- Le pourcentage global de mobiles en macro-diversité par station (pas de distinction du nombre de liens dans l'*Active Set*).

Les taux de mise à jour de l'*Active Set* des mobiles pourraient être des indicateurs utiles pour déceler un mauvais paramétrage du paramètre *Delta = Drop_win – Add_win* (par exemple, des changements fréquents d'*Active Set* indiqueraient un paramètre *Delta* trop faible).

Choix des indicateurs de qualité pour le paramétrage de la macro-diversité

Dans un premier temps, nous avons choisi l'indicateur de taux de blocage des liens de MD *downlink* pour évaluer l'efficacité d'une ingénierie automatique du paramètre *Add_win* de macro-diversité. En effet, nous avons testé une fonction coût qui s'est avérée efficace pour le contrôle automatique du paramètre *Add_win* et qui intégrait la valeur de taux de blocage des liens de MD *downlink* pour la station centrale ainsi que la moyenne et le maximum des valeurs de ce paramètre sur l'ensemble des stations voisines. Par ailleurs, pour évaluer les performances des contrôleurs automatiques, les indicateurs de taux de coupure et de blocage *downlink* pourront aussi être utilisés efficacement car ce sont des évaluations pertinentes de la qualité de service dans la cellule. Les indicateurs de mise à jour de l'*Active Set* pourront aussi

fournir des mesures de l'efficacité d'un contrôleur automatique de l'algorithme de macro-diversité. En particulier, une étude plus précise de ces indicateurs pourra s'avérer utile dans l'optique d'une gestion automatique du paramètre *Delta*. Enfin, certains indicateurs comme la charge *downlink* par cellule n'ont pas les mêmes variations pour la station centrale, celle en cours d'optimisation, et ses voisines. Cette différence de comportement pourra être utile pour l'élaboration de stratégies d'optimisation du paramétrage de l'algorithme de macro-diversité.

Exemple de front de paramétrage uniforme pour la macro-diversité

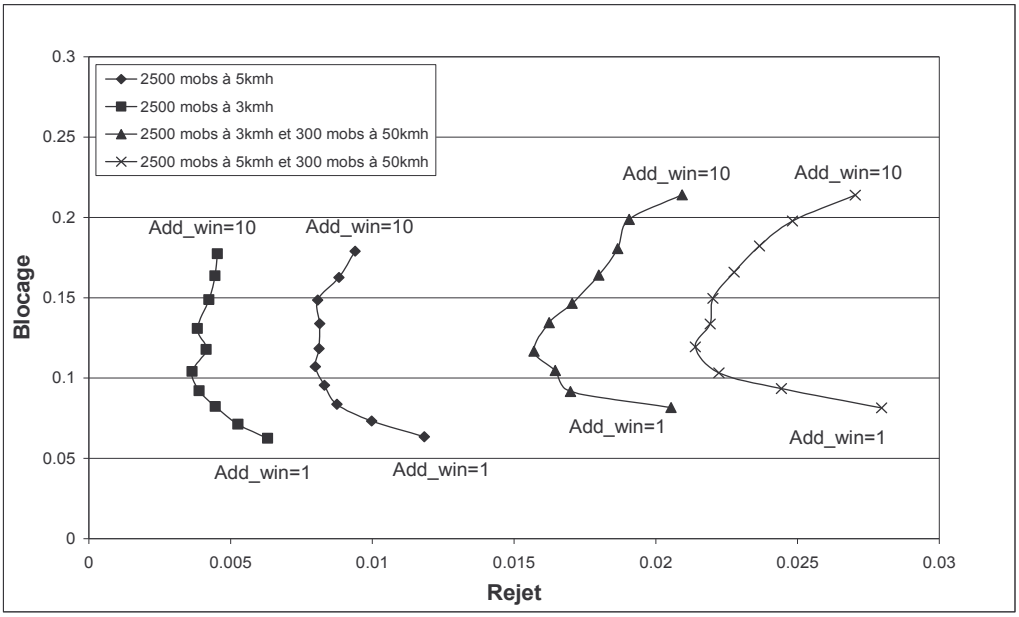


Figure 2.3 : Variations des taux moyens de blocage et de coupure en fonction du paramètre *Add_win* (statique et uniforme sur l'ensemble des stations) et du trafic

Sur la figure 2.3, nous avons représenté les courbes obtenues dans le plan taux de coupure - taux de blocage pour différents niveaux de trafic et différents paramétrages uniformes fixes de *Add_win* (c'est à dire sans ingénierie automatique). Ces courbes donnant les variations du taux de blocage en fonction du taux de coupure (les deux taux étant moyennés sur l'ensemble des stations du réseau), quand on fait varier le paramètre *Add_win* de 1 à 10 dB, sont appelées des fronts de paramétrage uniforme. On remarque que pour de faibles valeurs de ce paramètre, l'augmentation de *Add_win* réduit le taux de coupure, mais pour des valeurs plus importantes de *Add_win*, le taux de coupure augmente du fait d'une surcharge des stations causée par le nombre élevé des mobiles en macro-diversité. Dans les chapitres suivants, l'objectif de l'ingénierie automatique est d'obtenir des performances du réseau UMTS en deçà (en termes de taux de blocage et de coupure) de ces fronts de paramétrage uniforme, pour des niveaux de trafic équivalents.

2.1.3. Autres paramètres à optimiser dans les réseaux mobiles

Un contrôleur automatique pourra tenter d'améliorer la qualité de service et les performances du réseau en adaptant, node B par node B, les valeurs d'autres paramètres que ceux des algorithmes de macro-diversité ou de contrôle d'admission du réseau UMTS.

- Paramètres antennaires et puissances de canaux communs

Les paramètres d'azimuts, tilts, et puissances des canaux communs des nodes B auront a priori été optimisés dans la phase de design du réseau UMTS à l'aide d'un outil de type ACP. Cependant l'évolution de la demande en trafic (déplacement des zones de trafic, apparition de

nouveaux profils usagers…) risque de rendre moins efficace un paramétrage initialement quasi optimal. C'est pourquoi, il est nécessaire d'envisager la possibilité de corriger les valeurs de ces paramètres. En effet, certains constructeurs envisagent ainsi la possibilité de contrôler automatiquement les valeurs de tilts et puissances des canaux communs des nodes B au niveau de l'OMC. Deux types utilisations sont d'ailleurs considérés. La première consiste à modifier le paramétrage node B par node B, en fonction de variations anticipées du trafic: heure du jour, jour de la semaine, trafic saisonnier (vacances, jours fériés), manifestations exceptionnelles mais récurrentes (salons, compétitions sportives…). A l'aide de simulations et/ou de données de trafic, il aura été possible de calibrer plusieurs paramétrages typiques en fonction du trafic anticipé sur la zone. Le deuxième mode de fonctionnement est plus ambitieux mais plus risqué: en fonction des remontées de mesures, le contrôleur automatique adapterait le réglage en temps réel de ces paramètres – tilts et canaux communs.

- Autres paramètres radio

Avec l'augmentation du trafic de données sur le réseau UMTS, la gestion de la capacité radio par cellule va devenir un point critique. Un paramètre qui va donc devenir essentiel pour la capacité radio sur le lien descendant est la puissance maximale DL par canal DCH (*Dedicated channel*) en fonction du *bearer* demandé par l'utilisateur. Pour l'instant les valeurs de ce paramètre pour les différents *bearers* sont fixées constantes et homogènes sur l'ensemble du réseau. Dans un premier temps, nous proposons de les adapter node B par node B en fonctions des contraintes locales d'environnement et de trafic (peu ou beaucoup d'utilisateurs de service hauts débits sur la zone…). A plus long terme, nous envisageons une gestion dynamique de ce paramètre en fonction du trafic en temps réel sur la zone.

Un autre algorithme RRM important dans les réseaux UMTS est le contrôle de charge (ou *Congestion Control* [Sal03]) : cet algorithme assure que le système n'entre pas dans un état de surcharge (ou *overload*) et reste stable [Hol02]. Il sera donc essentiel de développer (comme pour le contrôle d'admission et la macro-diversité dans ce mémoire) des stratégies de contrôle automatique permettant d'adapter localement le paramétrage de cet algorithme aux données locales de trafic.

Enfin, l'algorithme de *Packet Scheduling* gère le trafic de données en mode paquet (allocation du canal de transport, surveillance des allocations de paquets et de la charge du système [Hol02]). Nous pourrons donc aussi intégrer les nouveaux algorithmes d'ordonnancement (ou *scheduling*) en développement [Ela03b] à nos méthodes de paramétrage automatique.

- Introduction au cas multi-système

Enfin, dans l'optique d'une gestion coopérative des différents réseaux mobiles, le 3GPP a normalisé un nouveau module de gestion conjointe de l'admission et de la re-sélection de réseau (*hand-over* inter-système) entre les réseaux GSM / GPRS et UMTS : le module *Common Radio Ressource Management* (C-RRM). Deux types d'architecture ont été proposés dans le rapport technique [3GPP1] de 2001 :

- L'approche de type serveur C-RRM : dans ce cas, les modules RRM (des systèmes GSM/GPRS et UMTS) sont conservés dans des nœuds séparés des modules C-RRM. Un serveur C-RRM, en fonction des informations fournies par les RNC et BSC, va déterminer une liste prioritaire de cellules candidates vers lesquelles l'appel doit être dirigé.
- L'approche C-RRM intégré : dans cette solution, les fonctionnalités C-RRM sont intégrées aux nœuds existants des réseaux UTRAN/GERAN.

Les algorithmes de sélection (à l'admission du mobile) et de re-sélection de réseau qui seront implémentés dans le module C-RRM (pour les deux architectures proposées), sont en cours de développement chez les constructeurs et opérateurs [Mea02] [Tol02a]. Dans une optique d'amélioration de la qualité de service et des performances de cette gestion coopérative des réseaux mobiles, on pourra donc adapter les méthodes de contrôle décrites dans ce mémoire, à la gestion automatique des paramètres (comme les seuils de re-sélection GSM / UMTS) qui seront introduits dans les futurs modules C-RRM, de gestion de l'admission et de la re-sélection de réseau entre les réseaux GSM / GPRS et UMTS.

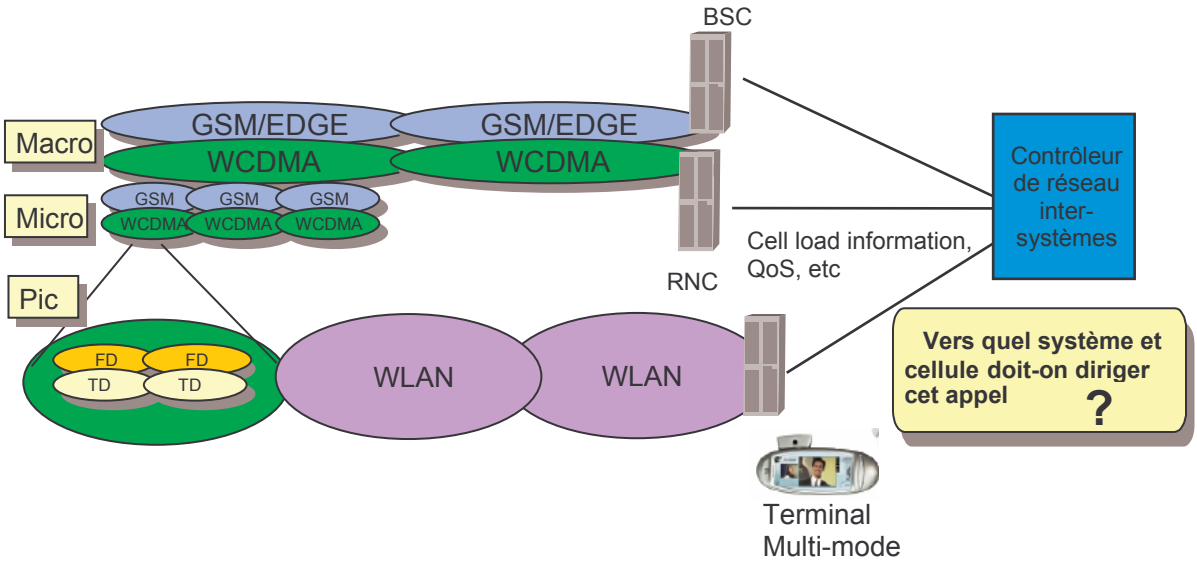


Figure 2.4 : Problème de la sélection de réseau pour la gestion coordonnée multi-systèmes

Par ailleurs, comme décrit sur la figure 2.4 ci-dessus, il est aussi envisagé d'intégrer la technologie WLAN dans une telle approche de gestion coordonnée des différentes technologies d'accès radio [3GPP2]. Il est donc important de développer des méthodes de contrôle automatique générique qui pourront être adaptées aux différents scénarii de gestion coordonnée multi-systèmes. Par exemple, un cadre théorique permettant de développer des stratégies de *Joint-RRM* (intégrant les systèmes UMTS, GSM/GPRS et WLAN), fondées sur des méthodologies de type *fuzzy-neural* a été introduit dans [Agu04].

2.2. Etude des performances d'un réseau UMTS

Dans la phase de dimensionnement du réseau, les opérateurs ont développé des méthodes analytiques d'évaluation de la couverture du service cible et d'estimation globale de la capacité cellulaire *downlink* et *uplink*. Ces formulations statiques vont nous permettre d'estimer grossièrement l'impact du paramétrage des principaux algorithmes RRM, contrôle de puissance, contrôle d'admission et de charge, macro-diversité, sur les performances globales du réseau. Dans la seconde partie de cette section, nous mentionnerons quelques approches analytiques plus sophistiquées permettant d'estimer plus finement les performances des réseaux. Ces méthodes pourraient nous permettre d'apprécier plus quantitativement l'influence du paramétrage RRM sur certaines performances du réseau.

2.2.1. Formulation statique du problème de capacité

Dans cette partie, on considère un réseau UMTS dans un environnement homogène, la distribution des sites (la grille des nodes B) étant relativement régulière. Les conditions de

propagation (canal de propagation multi-trajets) ne varient pas beaucoup sur la zone considérée et la distribution de trafic est supposée uniforme.

- ***Estimation de la capacité downlink***

Les canaux DPCH (*Dedicated physical channels*) sont contrôlés en puissance à la fréquence de 1500Hz. Nous supposons que le contrôle de puissance rapide est parfait et écrivons l'équation du lien élémentaire pour le mobile *j* [Hil00] :

$$\left(C/I\right)_j = \frac{P_j / PL_{BS,j}}{\alpha_j \left(P_{BS} - P_j\right)/PL_{BS,j} + \sum_{i \neq j} P_{BS_i}/PL_{i,j} + P_{Bruit}} \tag{2.5}$$

où $\left(C/I\right)_j$ est le rapport signal sur interférence requis, P_j la puissance requise au niveau du mobile *j* transmise par la station serveuse *BS*, α_j le facteur d'orthogonalité, P_{BS_i} la puissance totale transmise (en incluant les canaux communs) par la station *i*, P_{Bruit} la puissance du bruit et $PL_{i,j}$ le *path loss* entre la station i et le mobile j. Nous introduisons alors le facteur d'interférence *downlink* (cf. [Lai01]) qui correspond au rapport entre l'interférence inter-cellulaire (étude de l'espérance et de la variance de ce paramètre DL dans [Ela03a]) sur l'interférence intra-cellulaire :

$$F_j = \frac{PL_{BS,j}}{P_{BS}} \sum_{i \neq j} \frac{P_{BS_i}}{PL_{i,j}} \tag{2.6}$$

La puissance totale transmise par la station serveuse est égale à la somme des canaux de trafic et des canaux communs :

$$P_{BS} = \sum_j P_j + \sum_{CC} P_{CC} \tag{2.7}$$

En résolvant le système à deux d'équations, nous obtenons l'expression de la puissance totale transmise par la station serveuse :

$$\boxed{P_{BS} = \frac{\sum_{j=1}^{M} \beta_j PL_{BS,j} P_{Bruit} + \sum_{CC} P_{CC}}{1 - \sum_{j=1}^{M} \beta_j \left(\alpha_j + F_j\right)}} \tag{2.8}$$

Où *M* est le nombre de mobiles servis par la cellule et le paramètre β_j est égal à :

$$\beta_j = \frac{\left(C/I\right)_j}{\left(1 + \alpha_j \left(C/I\right)_j\right)} \tag{2.9}$$

A partir de l'expression (2.8) de la puissance totale P_{BS}, on définit le facteur de charge *downlink* η_{DL} [Lai01] par : $\eta_{DL} = \sum_{j=1}^{M} \beta_j \left(\alpha_j + F_j\right)$.

Donc, la puissance totale DL transmise par la station est égale à la somme de la puissance qui serait nécessaire en l'absence d'interférence plus le terme de *Noise Rise* du à l'interférence causée par l'utilisation d'une seule fréquence. Le *Noise Rise* en décibels est défini par :

$$NR_{DL} = -10\log\left(1 - \eta_{DL}\right) \tag{2.10}$$

En mono-service, nous introduisons la notion de capacité pole sur le lien *downlink*. Dans ce cas, en effet la puissance totale transmise par la station est donnée par la relation :

$$P_{BS} = \frac{\sum_{CC} P_{CC} + M \times \beta \times PL_{BS,moy} P_{Bruit}}{1 - M \times \beta \times \left(\alpha_{moy} + F_{moy}\right)} \tag{2.11}$$

Où $PL_{BS,moy}$, α_{moy} et F_{moy} sont les valeurs moyennes sur la cellule du path loss, du facteur d'orthogonalité, et du facteur d'interférence. On déduit alors l'expression de la capacité pole sur le lien DL par: $M_{Pole} = \frac{1}{\beta \times \left(\alpha_{moy} + F_{moy}\right)}$.

Prise en compte du facteur d'activité

Pour chaque service radio, nous introduisons un facteur d'activité sur les sens DL et UL; ν^{DL} et ν^{UL}. Ce facteur permet d'évaluer le taux d'utilisation moyen de la ressource radio pour les différents services accessibles (voix, visioconférence, données). Nous introduisons ce terme dans l'équation lien initiale et dans l'expression de la puissance totale transmise par la station. Nous obtenons au final la même expression que ci-dessus pour P_{BS}; le paramètre β_j étant maintenant égal à: $\beta_j = \frac{\left(C/I\right)_j}{\left(1 + \nu_{DL} \alpha_j \left(C/I\right)_j\right)}$.

Prise en compte de la macro-diversité

Sur le lien descendant, un mobile *j* peut être en macro-diversité avec plusieurs nodes B ou secteurs. Nous écrivons ci-dessous l'équation de lien avec recombinaison du signal sur le lien DL pour un mobile connecté avec trois nodes B, (généralisation de [Hil00]) :

$$\gamma_j = \left(C/I\right)_j = \left(C/I\right)_{j,BS} + \left(C/I\right)_{j,l} + \left(C/I\right)_{j,q} \tag{2.12}$$

Nous voulons voir l'impact de la prise en compte de la macro-diversité (à deux et trois liens, par exemple) sur l'expression de la puissance totale transmise par la station. Nous commençons par introduire le rapport Δ_j entre le rapport signal sur bruit du second meilleur lien sur le rapport signal sur bruit du meilleur lien, et p_{SHO} la proportion de mobiles ayant deux liens actifs, ainsi que les rapports, $\Delta_{j,1}$ et $\Delta_{j,2}$, et proportions p_{SSHO} équivalents pour les mobiles ayant trois liens actifs.

$$\Delta_j = \frac{\left(C/I\right)_{j,l}}{\left(C/I\right)_{j,BS}} \leq 1 \tag{2.13}$$

Et respectivement,

$$\begin{cases} \Delta_{j,1} = \dfrac{\left(C/I\right)_{j,l}}{\left(C/I\right)_{j,BS}} \leq 1 \\ \Delta_{j,2} = \dfrac{\left(C/I\right)_{j,q}}{\left(C/I\right)_{j,BS}} \leq 1 \end{cases} \tag{2.14}$$

Après quelques manipulations mathématiques élémentaires, nous déduisons les nouvelles expressions de la puissance totale transmise par la station et des paramètres β en mono-service (éq. (2.15)) :

$$P_{BS} = \frac{\sum_{CC} P_{CC} + M.P_{Bruit}\left[(1-p_{SHO}-p_{SSHO}).\beta_{NHO}.PL_{BS,moy}^{NHO} + 2.p_{SHO}.\beta_{SHO}PL_{BS,moy}^{SHO} + 3.p_{SSHO}.\beta_{SSHO}PL_{BS,moy}^{SSHO}\right]}{1-M\left[(1-p_{SHO}-p_{SSHO})\beta_{NHO}\left(\alpha_{moy}+F_{moy}^{NHO}\right)+2.p_{SHO}\beta_{SHO}\left(\alpha_{moy}+F_{moy}^{SHO}\right)+3.p_{SSHO}\beta_{SSHO}\left(\alpha_{moy}+F_{moy}^{SSHO}\right)\right]}$$

Avec

$$\beta_{NHO} = \frac{(C/I)_{BS}}{\left(1+\nu_{DL}\alpha_{moy}(C/I)_{BS}\right)} \tag{2.16}$$

$$\beta_{SHO} = \frac{(C/I)_{BS}}{\left(1+\Delta+\nu_{DL}\alpha_{moy}(C/I)_{BS}\right)} \tag{2.17}$$

$$\beta_{SSHO} = \frac{(C/I)_{BS}}{\left(1+\Delta_1+\Delta_2+\nu_{DL}\alpha_{moy}(C/I)_{BS}\right)} \tag{2.18}$$

Dans l'expression ci-dessus, M correspond au nombre de mobiles qui ont la station BS comme meilleure serveuse. Notons que la formule (2.15) généralise les résultats de l'article [Hil00] en exprimant dans le cas général la puissance totale transmise par la station centrale. Cela a permis de déduire la formule suivante de *Pole Point* généralisé sur le lien DL (éq. (2.19)) :

$$M_{pole} = \frac{1}{\left[(1-p_{SHO}-p_{SSHO})\beta_{NHO}\left(\alpha_{moy}+F_{moy}^{NHO}\right)+2.p_{SHO}\beta_{SHO}\left(\alpha_{moy}+F_{moy}^{SHO}\right)+3.p_{SSHO}\beta_{SSHO}\left(\alpha_{moy}+F_{moy}^{SSHO}\right)\right]}$$

Dans un document interne, nous avons étendu au cas *multi-services* l'expression (2.15) de la puissance totale transmise par la station, et déduit une nouvelle formule de *Pole Point* sur le lien DL (sous l'hypothèse d'une répartition connue du trafic entre les services).

Analyse de l'impact de différents paramètres du réseau sur la capacité DL

La puissance moyenne consommée au niveau de la station augmente (et donc la capacité DL diminue) quand :

- La puissance totale attribuée aux canaux communs augmente.
- Le *path loss* moyen $PL_{BS,moy}$, c'est-à-dire le rayon de la cellule, augmente.
- Les proportions p_{SHO} et p_{SSHO} de mobiles en macro-diversité dans la cellule augmente.
- Quand l'algorithme de macro-diversité est mal paramétré: des liens de MD inutiles sont activés, ce qui diminue globalement la capacité sur le lien DL.
- La puissance maximale par canal dédié (DCH) augmente (mais permet d'augmenter le rayon de couverture de la cellule).
- Le facteur d'interférence DL moyen sur la cellule augmente. Une telle situation peut correspondre à une mauvaise orientation d'une antenne voisine (azimut et tilt), une augmentation de la puissance transmise par les cellules voisines (sur les canaux communs et/ou dédiés), une augmentation de la charge des voisines (mauvais paramétrage du contrôle d'admissions, du contrôle de charge ou de la macro-diversité)…
- Le contrôle de puissance rapide fonctionne mal: soit la puissance transmise sur le lien est trop faible et l'information transmise est inutile (taux d'erreur trop important au décodage) soit elle est trop forte. Dans les deux cas, on génère inutilement de l'interférence sur le canal *downlink*.

A l'inverse, des algorithmes de contrôle d'admission et macro-diversité correctement paramétrés (le paramétrage ayant été adapté localement à la demande en trafic et aux contraintes d'environnement), permettront d'optimiser la capacité sur le lien descendant, qui sera fort probablement le lien limitant dans les réseaux UMTS

- ***Estimation de la capacité uplink***

Dans [Spi00], il a été établi l'expression suivante pour la charge sur le lien UL tenant compte du gain de sectorisation G_{Sect}, et du facteur d'activité du service ν_j^{UL} :

$$X_{UL} = \sum_{j=1}^{M} \frac{1}{1 + \frac{1}{(C/I)_j^{UL} \nu_j^{UL}}} \left(1 + F_j \times G_{Sect}\right), \tag{2.20}$$

Où $(C/I)_j$ est le rapport signal sur bruit requis sur le lien UL, F_j le facteur d'interférence UL et M le nombre de mobiles en communication avec la station BS. Cette formule tient compte de la saturation des mobiles et du phénomène de *hand-over* par l'intermédiaire de marges.

En *mono-service*, nous pouvons déduire une expression de la capacité pole sur le lien UL :

$$M_{Pole} = \frac{1}{1 + F_{moy}} \left(1 + \frac{1}{(C/I)^{UL} \nu^{UL}} \right), \tag{2.21}$$

Où F_{moy} est le facteur d'interférence UL moyen sur la cellule.

En *multi-services*, nous réécrivons l'expression de la charge UL en regroupant les mobiles ayant le même service :

$$X_{UL} = \sum_{Services} \sum_{j=1}^{M_S} \frac{1}{1 + \frac{1}{(C/I)_S^{UL} \nu_S^{UL}}} \left(1 + F_j\right) = \sum_{Services} \frac{M_S}{1 + \frac{1}{(C/I)_S^{UL} \nu_S^{UL}}} \left(1 + F_{moy,S}\right) \tag{2.22}$$

Où M_s est le nombre de mobiles pour le service s, et $F_{moy,s}$ le facteur d'interférence UL moyen sur la cellule pour le service s. Pour déterminer la capacité UL moyenne, on se fixe une charge limite sur le sens montant, et on peut déduire le nombre maximal admissible de mobiles par cellules (en définissant par exemple les pourcentages de mobiles par service en multi-services).

Analyse de l'impact de différents paramètres du réseau sur la capacité UL

Dans cette modélisation, l'influence du paramétrage du réseau est concentrée dans la variable du facteur d'interférence moyen F_{moy}. Sur le lien UL, la capacité cellulaire diminue quand le facteur d'interférence augmente :

- Quand l'algorithme de macro-diversité est mal paramétré: en effet, dans ce cas, certains liens UL ne sont pas choisis, alors qu'ils permettraient une diminution de la puissance transmise par le mobile et donc de réduire le niveau d'interférence UL.
- Quand la charge UL des cellules voisines augmente (mauvais paramétrage du contrôle d'admissions ou du contrôle de charge).
- Quand une des cellules voisines a une mauvaise orientation d'une de ses antennes ou un mauvais paramétrage de ses canaux communs. Dans ce cas, une partie des mobiles qu'elle

devrait couvrir, sont pris en charge par la station centrale : ils doivent augmenter leur puissance de transmission en augmentant ainsi l'interférence totale sur l'UL.

A l'inverse, un paramétrage efficace et adapté localement à la demande en trafic des nodes B permettra de limiter ces effets et d'améliorer la capacité UL. En particulier un bon paramétrage de l'algorithme de macro-diversité permettra de diminuer la puissance transmise par les mobiles (avec une amélioration de la couverture UL) et donc l'interférence totale UL.

2.2.2. Etude de la couverture

Tout opérateur souhaite optimiser le taux de couverture de son réseau mobile. Pour cela, dans la phase de dimensionnement, il va déterminer le nombre minimal de sites (avec un outil de type ACP) permettant de couvrir la zone cible en optimisant notamment le paramétrage antennaire de ces sites. Par la suite, en fonction de l'évolution du trafic (distribution et niveau), il va pouvoir adapter son paramétrage antennaire, mais surtout optimiser localement le paramétrage des algorithmes RRM, afin d'atteindre le meilleur compromis réalisable entre les objectifs de couverture et de capacité [Vee99]. Nous rappelons qu'un mobile est couvert, s'il reçoit correctement les canaux communs de la cellule ainsi que le canal DCH du service cible, et s'il peut émettre le niveau de puissance nécessaire pour que le rapport signal sur interférence reçu par la station de base soit suffisamment fort.

- **Couverture des canaux communs**

Les autres canaux communs sont fixés relativement au canal pilote CPICH. On considère que l'ensemble des canaux communs est reçu correctement, quand la qualité *Ec/Io* reçue du canal CPICH est supérieure à un seuil déterminé expérimentalement. Pour améliorer la couverture des canaux communs, on peut soit augmenter la puissance du CPICH (au détriment de la capacité de la cellule), soit diminuer le niveau d'interférence DL en optimisant par exemple le paramétrage des algorithmes de macro-diversité et de contrôle d'admission.

- **Couverture des canaux de trafic: bilans de liaison radio**

Les bilans de liaison radio permettent d'estimer le rayon maximal par service d'une cellule typique du réseau. Dans les bilans de liaison radio UL et DL, on prend en compte [Spi00] :

- Le *path loss* de propagation, les gains d'antennes, les pertes feeders, les gains de diversité, les marges de *fast fading* (ou marges de saturation), et les marges de *shadowing*.
- Les gains de *soft hand-over*.
- Les marges d'interférence.

En UL, une optimisation du paramétrage RRM des nodes B pourra avoir un impact sur les gains de *soft hand-over* et les niveaux d'interférences dans la cellule (on pourra alors diminuer en proportion les marges d'interférence). En DL, cela impactera principalement la puissance maximale par canal dédié (hausse ou baisse selon les cas) et le pourcentage de mobiles en macro-diversité (et donc la marge supplémentaire due à la MD). Enfin sur les liens UL et DL, une diminution du *Noise Rise* (c'est-à-dire une diminution du niveau d'interférence) permettra d'améliorer la couverture du service mais aussi la capacité de chaque lien.

La sortie des bilans de liaison UL et DL correspond donc à un *path loss* maximum admissible par service, qui dépend du type d'environnement, du type de pénétration et d'effets liés au W-CDMA (macro-diversité, *shadowing*, *fast fading*…). A partir de ce *path loss* maximum admissible, on détermine à l'aide d'un modèle de propagation [Lai00]: le rayon cellulaire, la

distance inter-site typique et le pourcentage de couverture pour le service cible (valable pour les sites tri-sectoriels).

$$d_{intersite} = 1.5 \times d_{CellRange} \tag{2.23}$$

$$S_{CoverageArea} = \frac{9}{8}\sqrt{3} \times d_{CellRange}^2 = 1.95 \times d_{CellRange}^2 . \tag{2.24}$$

En conclusion, si l'on peut estimer le gain réalisé par optimisation de certains algorithmes RRM sur certains paramètres des bilans de liaison, on peut en déduire une évaluation grossière de l'amélioration du taux de couverture dans le réseau.

2.2.3. Evaluation statistique des performances des réseaux

Des méthodes analytiques plus complexes étudient l'estimation des performances des réseaux UMTS, en se fondant en général sur des approches statistiques.

Ainsi, l'article [Ela03a] propose d'étudier les variations du SIR UL et d'en déduire ses impacts sur la capacité du réseau UMTS. D'autres méthodes ont introduit une notion de probabilité de dysfonctionnement [Cha01a] pour analyser la capacité du réseau UMTS. Ainsi sur le lien descendant, des approches analytiques comme [Cho01] [Zha01a] définissent des notions de probabilité de dysfonctionnement dans le cas de trafic multi débits. Sur le lien montant, les méthodes décrites dans [Kim00] et [Zha01b] introduisent la notion de dysfonctionnement en la fondant respectivement sur le SIR atteint au niveau de la station et sur le niveau total d'interférence dans le réseau.

Nous avons généralisé ces approches en proposant une méthode d'analyse de la probabilité de dysfonctionnement intégrant les liens montants et descendants [Dub03] [Brev04]. Dans le cadre d'une telle approche, il est possible d'estimer plus quantitativement l'influence de certains paramètres RRM sur les performances du réseau UMTS.

2.3. Conclusions

Dans ce chapitre, les algorithmes de contrôle d'admission et de macro-diversité ont été analysés dans une optique de paramétrage automatique des réseaux UMTS. Cette étude a permis de trier et sélectionner les indicateurs de performance significatifs de l'état des cellules du réseau et les paramètres des ces algorithmes les plus influents sur les performances du système. Dans les chapitres 6 et 7, nous testerons l'efficacité des méthodes de contrôle automatique développées dans ce mémoire appliquées au contrôle d'admission et à la macro-diversité. Nous avons mentionné les autres algorithmes et paramètres auxquels nous pourrions appliquer l'ingénierie automatique : le contrôle de charge, le *packet scheduling*, le contrôle de puissance… Nous envisageons aussi d'étendre ces méthodes de contrôle à des contextes multi-système comme : le module C-RRM pour la gestion commune du GSM et de l'UMTS, ou l'intégration du WLAN dans des approches de type *Joint-RRM*.

Dans la seconde partie du chapitre, nous avons introduit des méthodes d'estimation de la capacité et de la couverture dans les réseaux UMTS. De telles modélisations permettent d'estimer grossièrement l'impact du paramétrage RRM sur les performances du réseau. Enfin, nous avons mentionné quelques approches plus sophistiquées d'estimation de la capacité, fondées sur des méthodes statistiques. Une approche introduisant par exemple une notion de probabilité de dysfonctionnement dans les réseaux mobiles pourrait fournir une analyse plus quantitative de l'influence du paramétrage RRM sur les performances du réseau.

3. Contrôle des paramètres RRM par logique floue

L'objectif de ce chapitre est d'introduire un système générique permettant de contrôler les paramètres RRM de l'ensemble des stations du réseau UMTS, voire d'autres réseaux mobiles mono et multi-systèmes. Les réseaux mobiles sont des systèmes dynamiques (de part la mobilité et la nature aléatoire du trafic) et complexes, notamment à cause des phénomènes d'interférences. Dans ce chapitre, nous voulons introduire un cadre théorique au problème difficile d'ingénierie automatique des réseaux mobiles. Dans cette optique, les Systèmes d'Inférence Floue (SIF) permettent une modélisation appropriée de notre problème d'ingénierie automatique. Les entrées et les sorties du système sont respectivement les indicateurs du réseau et les corrections à appliquer aux paramètres RRM. Ce sont des valeurs réelles continues. La modélisation par un SIF permet de ramener le problème du contrôle à un ensemble de règles de décisions élémentaires qui peuvent se déduire de l'expérience ou de l'analyse des experts radio. Un SIF se compose principalement d'une base de règles de décisions élémentaires de la forme (voir paragraphe 3.1) :

$$Si\ (X \text{ est } L_X \text{ et } Y \text{ est } L_Y)$$

$$Alors\ Z \text{ est } O_Z.$$

Où X et Y sont les variables d'entrées (ou état) du système, L_X et L_Y des labels[*] linguistiques caractérisant ces entrées, Z la variable de sortie et O_Z le label caractérisant cette sortie.

A partir du cadre général du SIF ainsi défini, dans la suite du chapitre nous allons introduire une structure adaptée à nos contrôleurs automatiques des paramètres RRM. Il apparaît en effet qu'un contrôleur de type FKBC (*Fuzzy Knowledge Based Controller*) répond de manière satisfaisante aux contraintes de notre problème. Les composants d'un processus FKBC comprennent les éléments suivants[**] : le *Module de fuzzification* transformant les variables d'entrées continues en des labels linguistiques, la *Base de connaissance* qui se décompose en la *base de données* (qui contient l'ensemble des informations relatives à l'environnement du contrôleur ainsi qu'aux variables et ensembles flous) et la *base de règles* (qui contient l'ensemble des règles de paramétrage élémentaires), le *Moteur d'inférence* permettant de déterminer la valeur floue globale de sortie à partir des contributions de chaque règle élémentaires, et le *Module de defuzzification* transformant la valeur floue de sortie en une valeur déterministe (continue ou discrète) à appliquer au système.

La suite du chapitre aura pour objectifs de définir concrètement l'ensemble des composants du processus FKBC pour le cas du paramétrage des algorithmes RRM et d'imaginer des solutions permettant d'introduire ces contrôleurs automatiques dans les réseaux mobiles UMTS : type d'architectures possibles et prise en compte des différentes contraintes temporelles (tailles des fenêtres de filtrage, granularité des mesures)… En particulier, les moteurs d'inférence pour le contrôle des paramètres du contrôle d'admission et de la macro-diversité seront représentés mathématiquement par un système de matrices de décisions, qui recevra en entrées les valeurs floues des indicateurs de performances de la station centrale et de ses voisines. L'expression du moteur d'inférence par un système de matrices et le choix de fonctions d'appartenance élémentaires (formes triangulaires) pour les ensembles flous permettront d'ailleurs une représentation paramétrique simple des éléments composants (voir en annexes ; la paramétrisation des matrices et extrémités d'intervalles) le contrôleur de logique floue. Cette modélisation des contrôleurs facilitera en particulier le processus d'optimisation par essaim de particules décrit au chapitre 4.

---

[*] Un label est un terme qualitatif rendant compte de l'état d'un indicateur : par exemple *faible*, *moyen* et *fort* sont des labels qui permettront de qualifier le taux de blocage dans le réseau.

[**] Chapitre 3 de [Dia96].

3.1. Modélisation par un système d'inférence floue

3.1.1. Caractéristiques d'un système d'inférence floue

La figure 3.1 ci-après présente les différents éléments d'un système d'inférence floue (ou *Fuzzy Inference System*) :

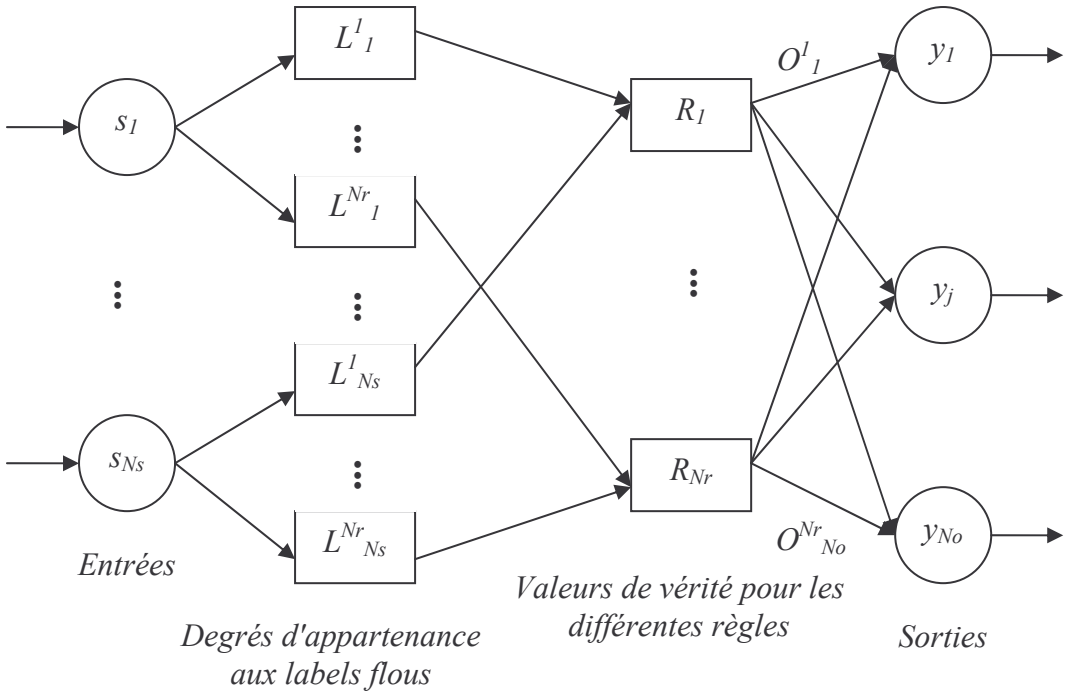


Figure 3.1 : Diagramme d'un système d'inférence floue.

Un système d'inférence floue (SIF) est défini à partir de N_R règles de décisions R_i de la forme suivante :

$$R_i: Si\ (s_1 \text{ est } L^i_1 \text{ et } \ldots\ s_{Ns} \text{ est } L^i_{Ns})$$
$$Alors\ (y_1 \text{ est } O^i_1 \text{ et } \ldots\ y_{No} \text{ est } O^i_{No})$$

Avec

R_i	la ième règle de la base de règles,
$\underline{s} = (s_1, s_2, \ldots, s_{Ns})$	le vecteur d'entrée (ou état),
$\underline{S} = S_1 \times S_2 \times \ldots \times S_{Ns}$	l'ensemble de définition des variables d'entrées,
L^i_j	terme linguistique (label flou) correspondant à la variable d'entrée s_j dans la règle R_i,
$\mu_{L^i_j}$	fonction d'appartenance au label flou L^i_j,
$\underline{y} = (y_1, y_2, \ldots, y_{No})$	le vecteur des *No* variables de sortie,
O^i_j	terme linguistique (label flou) correspondant à la variable de sortie y_j dans la règle R_i.

3.1.2. Cas des réseaux mobiles UMTS

Dans l'optique du contrôle des paramètres RRM des réseaux mobiles UMTS, nous avons défini et introduit le système d'inférence floue suivant.

La base des règles de décisions

Les règles de décisions ont tout d'abord été déduites de l'expérience humaine acquise en ingénierie des réseaux mobiles. Elles ont été énoncées sous forme de propositions linguistiques simples, parfaitement compréhensibles par tout un chacun, du type :

" Si le taux de blocage des appels est fort et le taux de coupure des appels est faible Alors augmenter le seuil de contrôle d'admission" (i)

Il s'agit à présent de formuler ces règles linguistiques simples de type *" Si ... Alors ..."* sous formes de relations mathématiques qui puissent être utilisées au cours d'un processus de paramétrage automatique du réseau mobile. Une modélisation mathématique possible de ces règles mathématiques est fournie sous la forme de matrices de décisions:

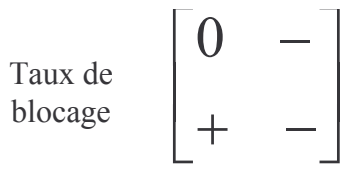


Figure 3.2 : Exemple de matrice de décisions

La figure 3.2 décrit une matrice de décisions élémentaire déduite de l'expérience acquise en ingénierie des réseaux mobiles. La partie gauche de la matrice correspond à des taux de coupure faibles, et la partie droite à des taux de coupure forts. La partie haute de la matrice correspond à des taux de blocage faibles, et la partie basse à des taux de blocage forts. Pour déterminer cette matrice, la proposition linguistique (i) a été complétée par deux autres propositions empiriques :

" Si le taux de blocage des appels est faible et le taux de coupure des appels est fort Alors baisser le seuil de contrôle d'admission" (ii)

" Si le taux de blocage des appels est fort et le taux de coupure des appels est fort Alors baisser le seuil de contrôle d'admission" (iii)

Enfin on complète la dernière case de la matrice de décision de manière intuitive :

" Si le taux de blocage des appels est faible et le taux de coupure des appels est faible Alors ne pas modifier le seuil de contrôle d'admission" (iv)

Dans le cadre du processus de contrôle automatique des paramètres RRM décrit au paragraphe 3.2.1, nous définirons de la même manière un ensemble de matrices de décisions plus complexes permettant le paramétrage des algorithmes d'admission des mobiles et de macro-diversité des réseaux UMTS :

- Nombres de colonnes et de lignes de la matrice déterminés en fonction du nombre de labels flous définis pour l'indicateur considéré.
- Remplissage direct des cases de la matrice pour lesquelles une règle empirique est connue (notamment les sommets de la matrice et la diagonale).
- Remplissage en utilisant des symétries et/ou des intuitions logiques fondées sur l'expérience et l'analyse des experts radio pour les cases restantes de la matrice.

Les matrices ainsi déterminées font la synthèse mathématique d'une première série de règles de décisions pour notre SIF définie de manière empirique. Dans le paragraphe 3.2.1, nous

introduirons cet ensemble de règles dans la base de connaissance du processus de contrôle de type FKBC (*Fuzzy Knowledge Based Controller*).

Définitions des entrées (ou état) du système d'inférence floue

A cette étape, il s'agit de déterminer les entrées qui définissent de manière pertinente l'état du système à contrôler. Ainsi, on déduira les mesures à collecter dans le réseau permettant de calculer l'ensemble des indicateurs de performances (ou *Key Performance Indicators* ; KPI), qui seront fournis en entrées des contrôleurs de logique floue. Les mesures élémentaires ainsi que les KPI dépendront évidement des algorithmes de gestion des ressources radio (ou RRM) que l'opérateur souhaite paramétrer de manière automatique. Par exemple pour le contrôle d'admission, un vecteur de qualité de service en entrée du système d'inférence floue peut être défini comme le vecteur des couples (taux blocage, taux coupure) pour la station centrale (station 0) et ses voisines (stations proches géographiquement et/ou ayant un niveau d'interférence non négligeable pour la station 0) :

$$\underline{s} = (b_0, d_0) \times (b_1, d_1) \times \ldots \times (b_{Ns}, d_{Ns}), \tag{3.1}$$

où *Ns* est le nombre de stations voisines de la station centrale. Dans ce cas, l'ensemble de définition correspondant est l'hypercube $[0,1]^2 \times [0,1]^2 \times \ldots \times [0,1]^2$. Le vecteur d'entrée défini ci-dessus a ainsi été utilisé dans le cadre de simulations du paramétrage automatique du seuil de charge cible downlink dans un réseau mobile UMTS.

Définitions des labels flous des entrées

La proposition linguistique (i) est une règle fondée sur l'expérience des experts en matière d'ingénierie des réseaux mobiles. Son énoncé qualitatif convient à une compréhension humaine mais pas à un processus de paramétrage automatique, car, par exemple, selon les circonstances le terme *"fort"* peut correspondre à différentes valeurs de taux de blocage. Dans un système d'inférence floue, pour permettre une telle modélisation " imprécise" des métriques, on définit les labels flous pour les entrées du système ainsi que les fonctions d'appartenance à ces labels qui prendront leurs valeurs dans l'ensemble *[0,1]*. Lors de la définition des ensembles flous, les fonctions d'appartenance vérifient la contrainte forte de partition, réalisant ainsi une partition floue forte du domaine d'entrée :

$$\forall i \in \{1,\ldots,N_S\}, \forall s_i \in S_i, \quad \sum_{j=1}^{N_L(i)} \mu_{L_j^i}(s_i) = 1, \tag{3.2}$$

$N_L(i)$ est le nombre d'ensembles flous utilisés pour partitionner l'ensemble de définition de la variable d'entrée s_i. Par exemple, on définit l'ensemble des valeurs floues du taux de blocage par cellule b_j par l'ensemble de labels qualitatifs suivants: *{faible, moyen, fort, et très fort}*. De la même façon, on définit l'ensemble de labels flous du taux de coupure par cellule d_j par : *{faible, moyen, fort, et très fort}*.

La première étape du processus de contrôle flou consiste en la *fuzzification*[*] des indicateurs de performances. Une valeur d'un indicateur KPI est ainsi transformée à l'aide des fonctions d'appartenance en un vecteur de degrés d'appartenance aux différents labels flous.

[*] Transformation des valeurs déterministes continues en valeurs floues.

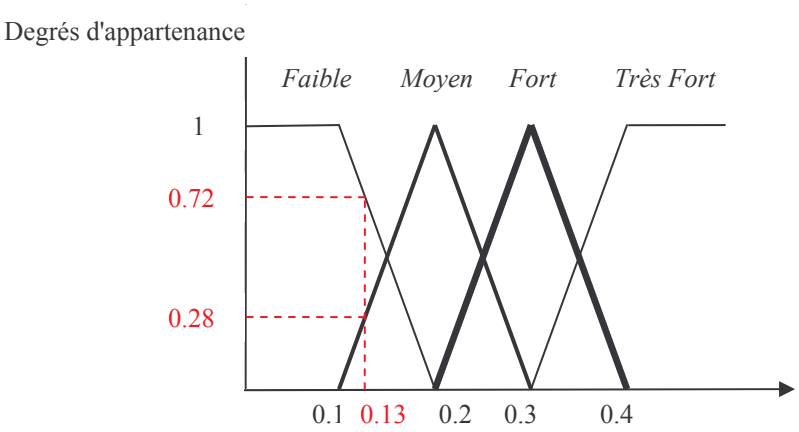


Figure 3.3 : *Fuzzification* d'une valeur continue de taux de coupure.

Dans l'exemple de la figure 3.3, la valeur mesurée de 0.13 du taux de coupure est transformée en une valeur floue (0.72, 0.28) des labels *Faible* et *Moyen*.

Dans ce mémoire, comme décrit sur la figure 3.3 et au paragraphe 3.2.1, les fonctions d'appartenance sont construites en utilisant des formes triangulaires. En effet, cela permet une représentation paramétrique simple des partitions floues qui facilitera par la suite l'optimisation du processus de contrôle (voir chapitre 4). Ci-dessus, les partitions floues sont définies par les bases des triangles isocèles avec des décalages permettant de vérifier la relation de normalisation (3.2) : {0, 0.1, 0.2, 0.3, 0.4}.

Définitions des labels flous de sorties

Tout comme pour les labels flous d'entrées, on définit les valeurs floues de sorties par un ensemble de labels qualitatifs, qui correspondent dans notre cas à la modification d'un paramètre d'un algorithme RRM: *{forte diminution, faible diminution, pas de modification, faible augmentation, et forte augmentation}*. Ces labels qualitatifs O^i_j correspondent chacun à une valeur numérique fixée à l'initialisation, la variable de sortie y_j étant une synthèse numérique de ces valeurs qui dépend de la règle d'agrégation des valeurs floues choisie (voir paragraphe 3.3.2.2).

Par exemple, les variables de sortie peuvent correspondre à une modification du seuil de charge d'admission (paramètre de contrôle d'admission) ou à une modification du paramètre de fenêtre d'admission (paramètre de macro-diversité) : pour le paramètre *Add_win*, les valeurs numériques correspondant aux labels flous de sortie peuvent être {-0.1 dB, -0.05 dB, 0, +0.05 dB, +0.1 dB}.

3.2. Définition d'une architecture de contrôle du paramétrage automatique

Dans un premier temps, nous avons classiquement opté pour une stratégie de contrôle des paramètres RRM par un processus de type FKBC (*Fuzzy Knowledge Based Controller*: processus de contrôle basé sur des connaissances floues), car une telle méthode présente de nombreux avantages, notamment dans l'optique d'une utilisation opérationnelle :

- Il s'agit d'un moyen simple, pratique et synthétique pour énoncer des règles de contrôle (traduites ensuite mathématiquement par un système de matrices de décisions comme décrit au paragraphe 3.3.2).

- Il est inutile de définir un modèle analytique de contrôle des paramètres RRM.
- Dans ce modèle, l'implémentation d'un mécanisme d'optimisation pour le contrôleur est pratique :
    - Description paramétrique souple des partitions floues et des matrices de décisions
    - Possibilité de définir des fonctions coûts globales permettant de comparer l'efficacité des contrôleurs de logique floue
    - Adaptation simple d'une stratégie d'optimisation par essaim de particules (OEP)

Dans la suite du chapitre, nous présenterons la première architecture de contrôle (de type FKBC) des algorithmes RRM par logique floue que nous avons développée. Cette architecture nous a permis d'implémenter et tester des concepts élémentaires de paramétrage automatique – *on-line* et *off-line* – pour les algorithmes d'admission et macro-diversité. Par la suite, une nouvelle méthode de contrôle et d'optimisation, l'apprentissage par renforcement, a été étudiée et adaptée au paramétrage automatique des réseaux mobiles. Ce processus, décrit au chapitre 4, se fonde aussi sur le système d'inférence floue introduit auparavant, mais permet un mode d'apprentissage plus riche (prise en compte parallèle des données de chaque cellule, au lieu de l'utilisation de fonctions coûts globales dans l'OEP) et utilise des notions plus facilement transposables aux réseaux réels.

3.2.1. Description du contrôle automatique par logique floue

A partir de la modélisation par un système d'inférence floue introduite au paragraphe 3.1, on veut définir une méthode de réglage automatique des paramètres RRM. On propose donc le schéma global de fonctionnement suivant qui permettra le contrôle du paramétrage des réseaux mobiles :

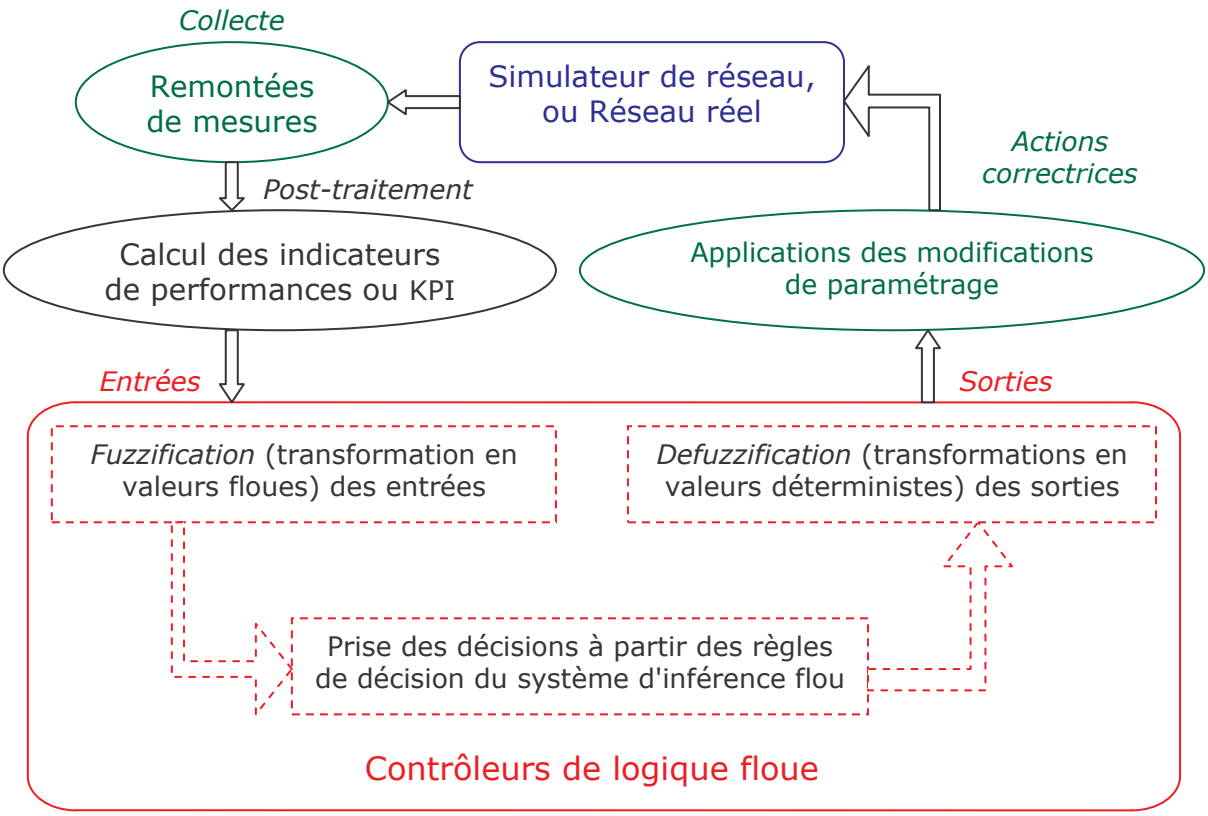


Figure 3.4: Schéma de fonctionnement du contrôle automatique par logique floue.

Le système surveille les performances du réseau en collectant périodiquement les métriques nécessaires et les post-traitant pour obtenir des indicateurs de performances significatifs fournis en entrées des contrôleurs de logique floue. Comme l'indique le schéma de fonctionnement sur la figure 3.4, tout processus de contrôle flou à l'intérieur d'un CLF se décompose en trois phases principales. La première étape, de *fuzzification* des données, consiste à transformer chaque valeur d'entrée déterministe en un vecteur de degrés d'appartenance aux différents labels flous. La deuxième étape, l'inférence de décisions, consiste à prendre des décisions en appliquant un ensemble de règles de type *" Si ... Alors ..."* définies dans le SIF. La dernière étape, de *defuzzification* des corrections, consiste à transformer les valeurs floues de sortie en grandeurs déterministes. Enfin en sortie des CLF, les actions correctrices – modifications du paramétrage des algorithmes RRM pour les stations identifiées – sont appliquées au réseau mobile.

En résumé, les contrôleurs de logique floue analysent régulièrement les indicateurs de performances du réseau mobile en se fondant sur l'expérience des ingénieurs radio et ajustent de manière automatique les paramètres RRM pertinents à l'aide de boucles de contrôles.

Les paramètres RRM suivants pourraient être ainsi contrôlés automatiquement :

- Pour l'algorithme de contrôle d'admission [Din04] [Ye05] : le seuil d'admission, les largeurs de bande allouées par appel, le seuil du niveau de signal reçu dans le cas des réseaux HCS (paramètre utilisé pour privilégier une couche par rapport à une autre [Lo03]), le débit maximum autorisé pour certains mobiles (cas avec trafic de données élastique) ainsi que le niveau maximal de dégradation…
- Pour les algorithmes de macro-diversité [Hom04] : la taille de la fenêtre d'ajout des liens de MD (*Add_win)*, la taille de la fenêtre de retrait des liens de MD (*Drop_win)*, la taille de la fenêtre de remplacement des liens de MD (*Rep_win)*, la taille de l'*Active Set*, la liste des cellules voisines, le paramètre CIO…

Dans l'optique d'un paramétrage automatique de certains de ces paramètres RRM, les indicateurs de performances (KPI) significatifs pourraient être respectivement:

- Pour l'algorithme de contrôle d'admission [Sen04] : la charge du réseau par cellule, le taux de coupure des appels, le taux de blocage des appels, la distribution du trafic, les informations de mobilité (probabilité de blocage des liens de MD) et le nombre d'appels dans les files d'attente ou *buffers* (en particulier pour les environnements avec des microcellules [Lo03])…
- Pour les algorithmes de macro-diversité [Cha01b] : le taux de *hand-over*, le taux d'échec de *hand-over*, le taux de coupure des appels, la charge du réseau, les informations de mobilité, la puissance des signaux pilotes, la QoS perçue, le taux d'erreur binaire, la couverture du réseau …

3.2.2. Types d'architectures réalistes envisageables

La maîtrise des coûts opérationnels (coûts d'OPEX) des réseaux cellulaires est devenue un thème de grand intérêt, du fait de la taille et de la complexité croissantes de ces réseaux. Il est donc essentiel d'étudier et définir une architecture possible qui permettra la surveillance et la gestion automatiques des réseaux mobiles, et réduira ainsi les coûts d'OPEX. La figure 3.5 suivante décrit une architecture hiérarchique classique et générique de surveillance et management pour les réseaux d'accès radio (systèmes GSM et UMTS par exemple) :

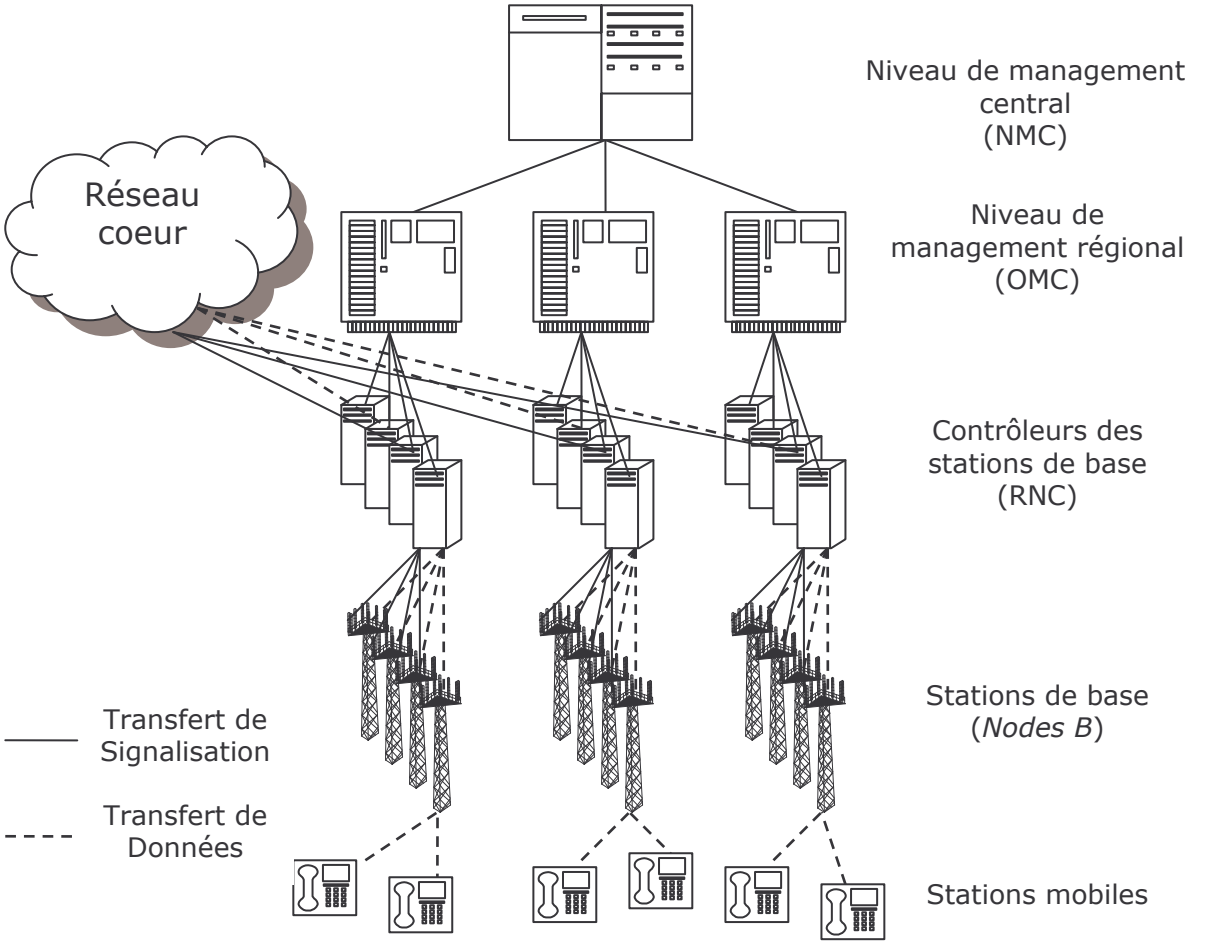


Figure 3.5 : Système hiérarchisé de management dans un réseau d'accès radio (avec la terminologie UMTS indiquée).

Les opérateurs de réseaux mobiles utilisent pour l'instant des systèmes centralisés de surveillance et management de réseau. Ces systèmes permettent par exemple d'identifier et de résoudre des dysfonctionnements (*troubleshooting* et *fault management*), de fournir des services avancés et personnalisés aux clients, de garantir la sécurité du réseau. A partir des noeuds centraux de gestion du réseau, comme l'OMC*, il est aussi possible de reconfigurer les paramètres du réseau d'accès radio définis par station.

A ce jour, le réglage des paramètres du réseau mobile est effectué "à la main" par un ingénieur radio. Cet expert se fonde sur des données d'analyses de performances et sur sa propre expérience, pour modifier et optimiser le paramétrage du réseau. Cela est effectué en envoyant des commandes de (re)configuration aux éléments du réseau d'accès.

3.2.2.1 Définition des étapes nécessaires à une implémentation dans les réseaux réels

L'automatisation des tâches de surveillance et de gestion des réseaux mobiles implique de pouvoir collecter et post-traiter des informations du réseau, et en fonction de ces résultats, d'appliquer des commandes correctrices de (re)configuration au réseau, à l'intérieur d'une architecture globale de boucles de rétroactions [Mag01] [Gus02].

* *Operations & Management Center* : centre de surveillance et gestion du réseau.

Collecte des mesures

La collecte de mesures dans le réseau permet de surveiller les performances des éléments observés, *node B* ou RNC par exemple, avec des remontées de données aussi proches que possibles du temps réel (c'est-à-dire des remontées continues et rapides) pour un fonctionnement en mode *on-line*. En revanche, dans le cadre d'un fonctionnement *off-line*, il n'est pas essentiel que ces remontées de données respectent les contraintes temporelles nécessaires aux applications temps réel.

Par ailleurs les tâches de collecte de mesures peuvent être réparties entre plusieurs niveaux du réseau [Mag01] : les agents de niveau inférieur effectuent des mesures élémentaires telles que le BER, taux d'erreur binaire, le rapport signal sur interférence, les puissances pilotes reçues… Les agents de niveau supérieur, OMC par exemple, peuvent à partir de ces mesures élémentaires calculer des quantités plus abstraites telles que les indicateurs de performances (KPI) du réseau : capacité, qualité de service, taux de coupure et de blocage des appels…

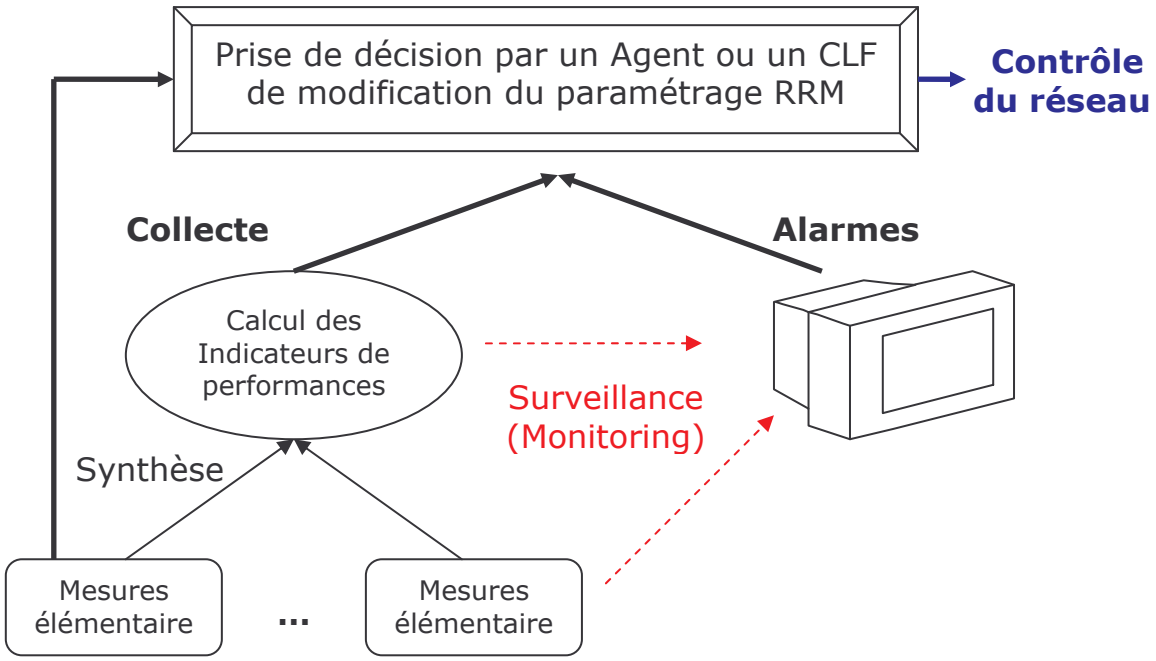


Figure 3.6 : Schéma de remontées de mesures et surveillance dans un réseau mobile.

Post-traitement des mesures

Cet étape comprend le filtrage des mesures et le calcul des indicateurs de performances :

- Il est nécessaire de filtrer les mesures élémentaires, à cause des fluctuations irrégulières des données instantanées inhérentes à la transmission sur le canal radio (déplacement des mobiles, effets du *shadowing* et du *fast fading*, débuts et fins de communications aléatoires dans le réseau…). Un filtrage de type fenêtre glissante permet de lisser les variables d'interférences, puissances ou taux d'erreur. Ainsi on obtient des valeurs moyennées significatives de l'état "instantané" du réseau. La taille des filtres dépendra d'ailleurs des différents usages de l'ingénierie automatique : ingénierie dynamique directement sur le réseau réel (mode *on-line*) ou reconfiguration statique du paramétrage du réseau (mode *off-line*).

- A partir des valeurs filtrées des mesures élémentaires, on calcule les KPI du réseau. Ces quantités constituent des indicateurs plus pertinents de l'état d'un cellule ou d'une zone, et plus adaptées à la prise de décisions de reconfiguration du paramétrage du réseau par un agent ou un système automatique de contrôle. Pour chaque algorithme RRM pris en compte par le processus, une phase d'analyse (exemples au paragraphe 3.3.2.1) devra nécessairement déterminer les indicateurs les plus significatifs de l'état de d'une cellule pour modifier localement le paramétrage de l'algorithme.

Remarque : il faudra adapter les limites des ensembles flous définis pour la *fuzzification* des indicateurs de performances en fonction des tailles des filtres de moyenne glissante utilisés et de la granularité des mesures collectées dans le réseau.

Applications des actions de contrôle

Les actions de contrôle du système peuvent s'appliquer directement sur les paramètres du réseau ou commander des modifications à un niveau inférieur (dans une structure hiérarchique). Une action de contrôle dépend de valeurs de mesures, de KPI, de la stratégie de management de l'opérateur, et éventuellement d'autres commandes de contrôle d'un niveau supérieur. Idéalement, le contrôle devrait être effectué aussi bas que possible dans la structure du réseau, pour permettre une bonne réactivité du système, et éviter des remontées de données jusqu'au nœud de management central (délais de remontée des métriques et de descente des commandes, dimensionnement approprié des canaux de signalisation, capacité de traitement adaptée au niveau du noeud…) dont on pourrait faire l'économie. Cependant, la possibilité de disposer de modules de contrôle dans les niveaux inférieurs des réseaux mobiles dépendra des intentions des constructeurs et de l'avancement de la normalisation au 3GPP sur ces questions. Théoriquement, on pourrait imaginer des structures de contrôles hiérarchisées sur plusieurs niveaux, comme décrit sur la figure 3.7 ([Sal04] [Gaï05] [Sal05]). Par ailleurs, deux stratégies d'applications des actions de contrôle sont envisagées : modifications simultanées de toutes les stations du réseau ou modification d'une seule station par groupe d'interférence (voir chapitre 5).

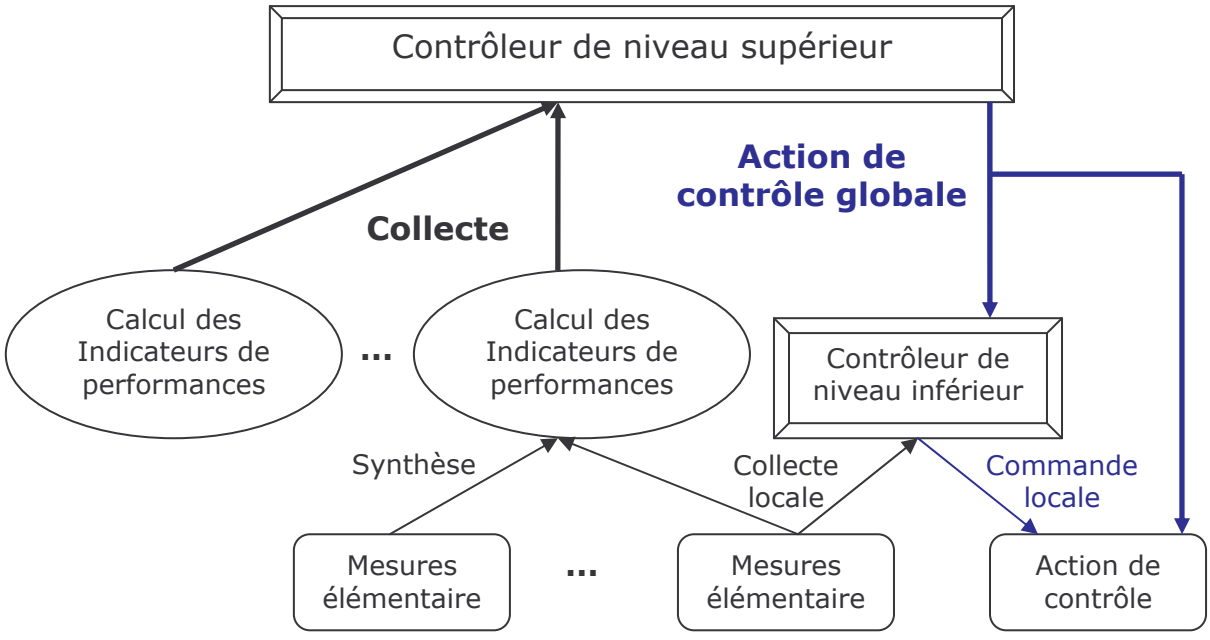


Figure 3.7 : Schéma d'applications des actions de contrôle dans un réseau mobile.

A court terme, des constructeurs (Nokia par exemple) envisagent des applications de *self-tuning* de certains paramètres RRM du node B à partir de l'OMC. Il s'agirait donc d'une structure centralisée de paramétrage automatique avec remontée des compteurs au niveau de l'OMC, traitement de données et modification des paramètres RRM à ce niveau (avec les délais d'applications des commandes de contrôle intrinsèques à une telle solution).

3.2.2.2 Modèle centralisé de paramétrage automatique

Classiquement dans les réseaux mobiles, les applications de surveillance et gestion du réseau (figure 3.5) sont effectuées au niveau d'un nœud central de management[*]. Dans une optique de configuration automatique du réseau, un modèle centralisé nécessiterait donc que des volumes importants de données soient remontés régulièrement des nœuds élémentaires du réseau d'accès radio vers le nœud central. Un fonctionnement *on-line* semble difficile à envisager car cela implique les contraintes supplémentaires suivantes pour limiter les délais liés aux remontées et traitement de l'information, et à la transmission des actions de contrôle :

- Détermination du niveau du réseau jusqu'où chaque type de mesure doit être remonté pour calculer les KPI correspondants (figure 3.6),
- Si possible, distribution dans le réseau des modules de calcul des KPI,
- Capacité de traitement performante au niveau des modules de calcul des KPI,
- Design de l'architecture optimale des canaux de signalisation et des canaux de transmission des commandes de contrôle,
- Dimensionnement adéquat des canaux de signalisation et des files d'attente,
- Capacité de traitement de grandes quantités de données (pour des réseaux mobiles étendus et complexes) très performante pour des applications de gestion du réseau en temps réel ou semi temps réel au niveau du nœud de management central.

Par conséquent dans une architecture centralisée, une implémentation de type *off-line* du paramétrage automatique apparaît plus accessible. En effet, il y a moins de contraintes liées à ce mode de fonctionnement:

- Pas de contrainte temporelle de remontées des mesures et d'applications des commandes de contrôle (contrairement au cas *on-line*),
- Détermination du niveau du réseau jusqu'où chaque type de mesure doit être remonté pour calculer les KPI correspondants,
- Mais, possibilité d'utiliser des outils de capture de mesures: on évite ainsi la remontée de grandes quantités de données sur les canaux de signalisation du réseau d'accès radio,
- Possibilité de traiter hors-ligne les données remontées du réseau sur des machines dédiées. Donc la capacité de traitement nécessaire au niveau du nœud de management central est réduite.

Ainsi avec l'architecture centralisée actuelle de surveillance et management des réseaux UMTS, nous pourrions dés à présent envisager une solution de paramétrage automatique en mode *off-line* au niveau de l'OMC. Une telle utilisation devrait prendre en compte les contraintes réelles de délais de remontées de mesures et de transmissions aux éléments du réseau des commandes de contrôle.

[*] L'OMC en UMTS, par exemple.

3.2.2.3 Modèle des agents distribués du paramétrage automatique

Du fait de la taille des réseaux mobiles, du grand nombre de mesures nécessaires dans une optique de paramétrage automatique, et de la nature distribuée des systèmes cellulaires, une architecture permettant la distribution des fonctions de surveillance et de contrôle dans le réseau semble plus adaptée à un mode de fonctionnement *on-line*.

Dans une telle architecture, les données de mesures devraient en principe être collectées et post-traitées près de la source avant d'être transférées au niveau hiérarchique supérieur. Plus globalement, il serait souhaitable dans les systèmes distribués d'éclater les applications de surveillance et de contrôle en sous-applications élémentaires qui seraient réparties entre les nœuds de gestion et les nœuds de trafic du réseau. Tout élément du réseau qui comporte ainsi une fraction distribuée des applications de gestion serait désigné par le terme d'*agent*. L'introduction dans les réseaux mobiles d'agents de surveillance et contrôle distribués présenterait les avantages suivants:

- Coordination des différentes commandes locales de contrôle au niveau de l'agent,
- Répartition des agents sur plusieurs niveaux du réseau (figure 3.7) : les agents de niveau inférieur seront adaptés au contrôle des problèmes nécessitant une réponse rapide et locale, les agents des niveaux supérieurs permettront de développer des stratégie globales de contrôle.
- Transmission de données (métriques et KPI) et de commandes de contrôle entre différents agents de différents niveaux du réseau mobile (voire de différents systèmes à plus long terme).

Un type d'architecture distribué nécessiterait donc de définir des interfaces ouvertes et standardisée permettant aux *agents* d'interagir à différents niveaux du système (exemple de l'interface IRP, *Integration Reference Points*, dans [Mag01]). Une telle utilisation d'un modèle par agents distribués intégrant des fonctions de mesures et de contrôle permettrait d'obtenir un système de surveillance et de management robuste (contrairement au cas d'une structure centralisée très sensible aux pannes éventuelles du nœud central de management) et de structure auto-similaire aux différentes échelles du réseau.

3.3. Description des contrôleurs de logique floue modélisé par un processus FKBC

La figure 3.4 décrit le schéma global de fonctionnement du contrôle automatique du paramétrage RRM des réseaux mobiles que nous proposons. Cette figure met en évidence les trois principales fonctionnalités d'un CLF. Mais la description complète et précise de la structure d'un contrôleur de logique floue, modélisé par un processus FKBC, comporte l'ensemble des éléments suivants (Chapitre 3 de [Dia96]) :

- *Module de fuzzification* : Ce module normalise les variables déterministes $\underline{s}$ d'état du processus (division par des facteurs d'échelle) avant de les transformer en des valeurs floues. Cela permet de les rendre compatibles avec la représentation des variables d'état par des valeurs floues dans les prémisses des règles d'inférence du système (partie *"Si ..."* des règles). Dans le module de *fuzzification*, deux classes de stratégies de *fuzzification* sont possibles pour le contrôleur, ce choix étant déterminé par le type de *moteur d'inférence* : l'inférence fondée sur la *composition* et celle fondée sur les *règles individuelles* (Chapitre 2 de [Dia96]).

- *Base de connaissance* : La base de connaissance d'un contrôleur de logique floue est constituée d'une base de données et d'une base de règles. La *base de données* fournit l'ensemble des informations nécessaires à un fonctionnement correct de la base de règles, du module de *fuzzification* et du module de *defuzzification*. Ces informations incluent les partitions floues (et les fonctions d'appartenance) des variables d'état et de sortie du processus ainsi que les facteurs d'échelle appliqués aux métriques en entrées. La *base de règles* permet de représenter de manière structurée la politique de contrôle définie par un ingénieur radio expérimenté sous la forme d'un ensemble de règles de production telles que :

  Si (état du processus) *Alors* (sortie de contrôle)

  La première partie de l'énoncé (partie *"Si ..."* des règles) est appelée l'antécédent (ou prémisse) de la règle et donne une description de l'état du processus sous forme d'une combinaison logique de propositions floues élémentaires. Par exemple; "*Si* b_i est *Fort* et d_i est *Moyen*".
  La deuxième partie de l'énoncé (partie *"Alors ..."* des règles) est appelée la conséquence de la règle et donne une description de la sortie de contrôle sous forme d'une combinaison logique de propositions floues élémentaires. Ainsi ces propositions déterminent les valeurs linguistiques (ou labels) que prendront les variables de sortie pourvu que les états courants du processus correspondent (au moins à un certain degré) à l'état du processus énoncé dans l'antécédent de la règle.

- *Moteur d'inférence* (ou de prise de décisions) : Comme indiqué ci-dessus dans le *module de fuzzification*, il existe deux types de moteurs d'inférence : l'inférence fondée sur la *composition* et celle fondée sur les *règles individuelles*. Avec le premier type d'inférence, on commence par agréger toutes les règles individuelles en une seule relation floue intégrant la signification de chaque règle. Alors, le moteur d'inférence réalise la composition entre le vecteur d'entrée *fuzzifié* et la relation floue globale agrégeant la signification de toutes les règles. Avec le second type d'inférence, la fonction principale du moteur d'inférence est de calculer la valeur globale de la variable de sortie à partir des contributions individuelles de chaque règle. Chaque contribution individuelle représente les valeurs des variables de sortie élémentaires calculées pour chacune des règles.

- *Module de defuzzification* : Ce module transforme tout d'abord les variables floues en sortie du *moteur d'inférence* en des valeurs déterministes. Différents types d'opérateurs de *defuzzification* peuvent être choisis: *defuzzification* par centre de gravité, par min-max, par produit - somme… On applique ensuite les facteurs d'échelle inverses pour déterminer finalement les valeurs de modifications des paramètres RRM à appliquer sur le réseau mobile.

La figure 3.8 suivante résume la structure complète du contrôleur de logique floue dans une modélisation par un processus de type FKBC (*Fuzzy Knowledge Based Controller*). Nous avons indiqué les différents flux d'information :

- Le flux de traitement des indicateurs de performances en entrées du contrôleur aboutissant aux valeurs de corrections en sortie du réseau. Les entrées et sorties (en noir) sont déterministes, alors que les étapes de traitements intermédiaires (en gris) utilisent des valeurs floues.

- Les informations fournies par la base de données aux différents éléments du contrôleur lors de la phase de définition du CLF.

- Les informations linguistiques fournies par la base de règles linguistiques, définies de façon expérimentale, à la base de règles floues intégrée dans le moteur d'inférence.

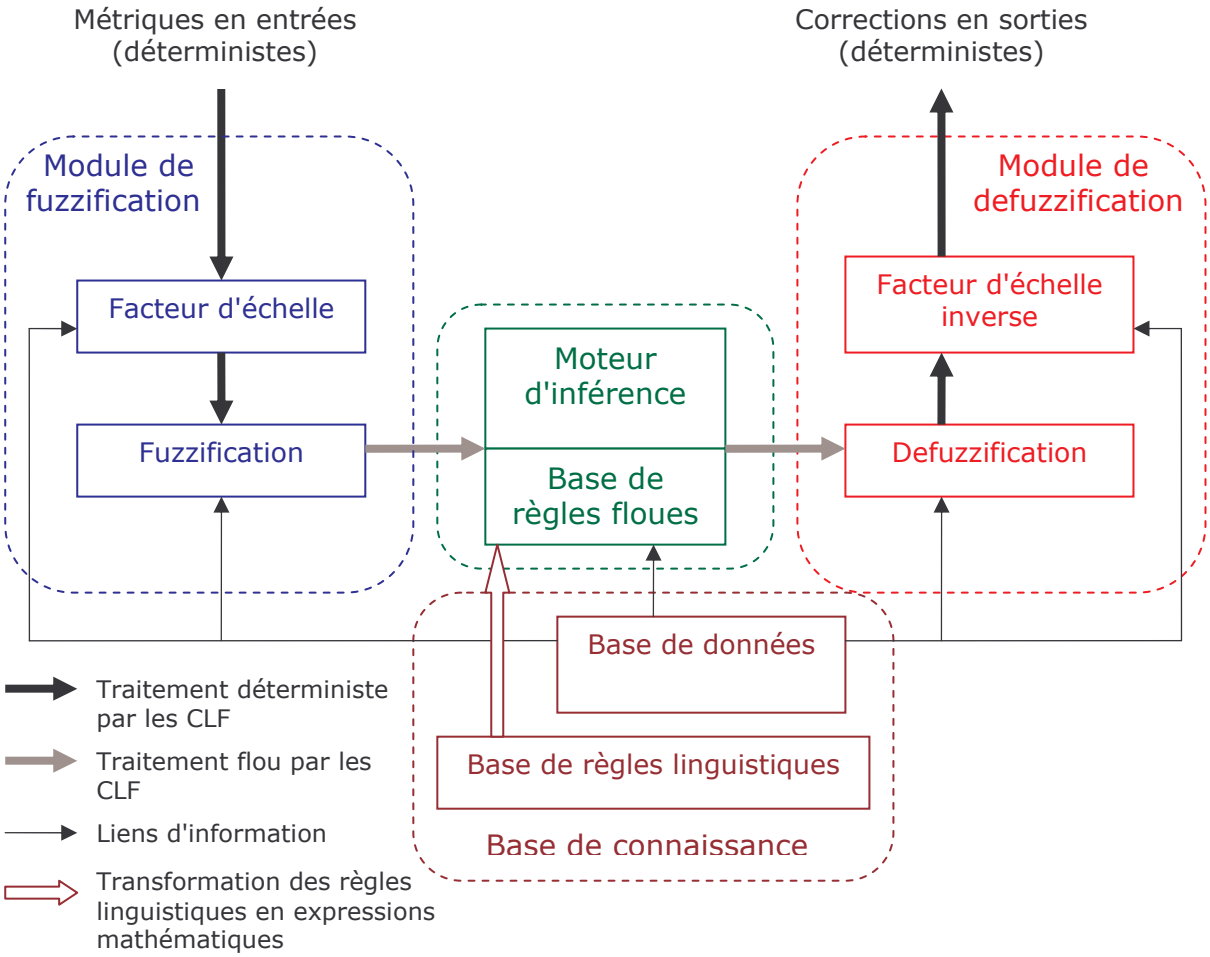


Figure 3.8 : Structure complète du contrôleur de logique floue, modélisé par un processus FKBC.

3.3.1. Procédure de *fuzzification*

Dans ce paragraphe, nous expliquons comment les métriques déterministes en entrées des contrôleurs sont transformées en valeurs floues (phase de *fuzzification*) sur lesquelles s'appliqueront les règles d'inférence définies au paragraphe 3.3.2.

3.3.1.1 Module de fuzzification des indicateurs de performances

Nous rappelons que deux types de moteur d'inférence sont possibles dans un contrôleur FKBC: le premier dans lequel l'ensemble des règles élémentaires est d'abord fusionné en une seule règle globale, et le second dans lequel on applique d'abord chaque règle élémentaire avant de calculer la variable de sortie globale à l'aide d'opérateurs flous. Dans les deux cas, il est nécessaire de transformer les indicateurs réels continus de l'état du réseau en des valeurs floues utilisables par le moteur d'inférence.

Les indicateurs de performance filtrés et normalisés (avec des facteurs d'échelle) sont donc mappés sur les ensembles des partitions floues qui ont été préalablement définis. Ainsi

comme expliqué au paragraphe 2.1.2, on transforme des valeurs réelles continues des indicateurs de blocage, coupure, blocage des liens MD ou le paramètre *Add_win*, en des valeurs floues qui vont être traitées dans le processus de contrôle.

Exemple : On considère que la figure 3.9 représente les ensembles *Moyen* et *Fort* d'une partition floue pour le taux de blocage de la station centrale (le type de partition choisi dans l'ensemble du mémoire est tel qu'une valeur réelle ne peut appartenir au maximum qu'à deux ensembles flous différents). Dans ce cas, une valeur de taux de blocage de 0.635 pour la centrale sera transformée en la valeur floue : *(Moyen, Fort)* avec des degrés d'appartenance de *(0.65, 0.35)*.

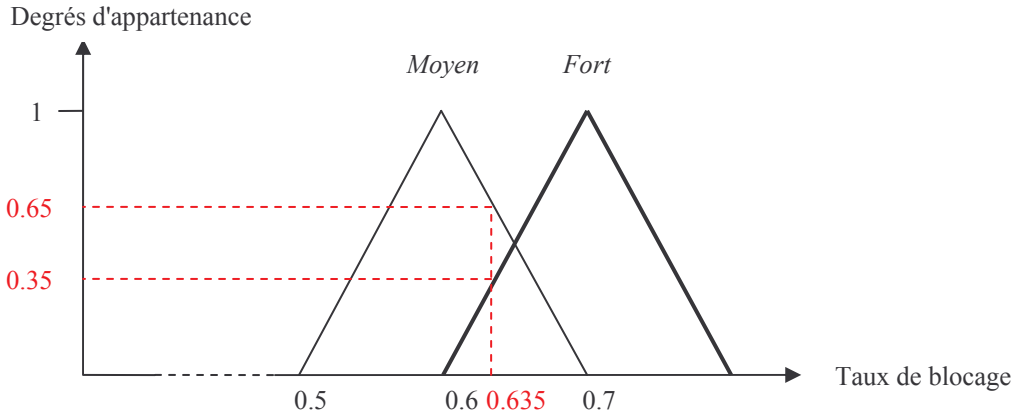


Figure 3.9 : *Fuzzification* d'une valeur continue de taux de blocage

3.3.1.2 Description de la base de données (*Data base*)

Pour construire correctement *notre base de données*, nous devons choisir les fonctions d'appartenance (des partitions floues) correspondant aux différents labels introduits, les limites maximum et minimum des ensembles ainsi que les facteurs d'échelle (de normalisation) appliqués aux métriques en entrées.

Définition des partitions floues

Pour la phase de *fuzzication* des indicateurs de performances, il est nécessaire d'introduire des partitions floues correspondant aux différents labels de chaque indicateur. Essentiellement, une valeur floue correspond à un vecteur d'état d'indicateurs qualitatifs : *faible, moyen, fort* et *très fort* taux de blocage par exemple. Les fonctions d'appartenance permettent de définir pour chaque valeur réelle dans l'intervalle considéré un degré d'appartenance à chaque valeur qualitative. Pour les ensembles flous définis dans ce mémoire, les fonctions d'appartenance vérifient la contrainte forte de partition. Ainsi le type de partition choisi implique que pour chaque valeur d'entrée s_i, il y a au plus deux ensembles flous réellement actifs (c'est-à-dire avec la fonction membre $\mu > 0$) et au moins un ensemble flou actif (l'ensemble de définition S_i est entièrement couvert par les ensembles flous). Les partitions floues, ainsi définies, des indicateurs de performances forment donc une hypergrille sur l'espace d'entrée. Un point de l'hypergrille correspond à la prémisse d'une règle d'inférence s'appliquant à ces indicateurs.

Dans notre méthode de contrôle, nous utiliserons deux types d'ensembles flous faciles à paramétrer : les ensembles flous uniformément et non uniformément répartis (voir figures 3.10 et 3.11).

Exemples de partitions d'ensembles flous à définir à l'initialisation du processus

Les ensembles flous pour les indicateurs de performances qui seront définis à l'initialisation du processus de contrôle automatique des algorithmes d'admission et de macro-diversité pourront être par exemple :

- Les ensembles flous des taux de blocage et de coupure pour la station centrale ainsi que pour ses voisines.
- Les ensembles flous des indicateurs de taux de blocage des liens de MD pour la station centrale, le maximum des taux de blocage des liens de MD sur l'ensemble des voisines, et la moyenne des taux de blocage des liens de MD sur l'ensemble des stations de la zone (y compris la centrale).
- Les ensembles flous pour la correction du paramètre *Add_win*.

- Ensembles flous uniformes

Entre les bornes minimum et maximum des ensembles flous, l'intervalle est divisé en parties égales et les fonctions d'appartenance sont définies par des triangles isocèles identiques pour tous les labels. Sur la figure 3.10, en dessous de X_{min} l'indicateur x est considéré à 100% comme *très faible* et au-dessus de X_{max} l'indicateur x est considéré à 100% comme *très fort*.

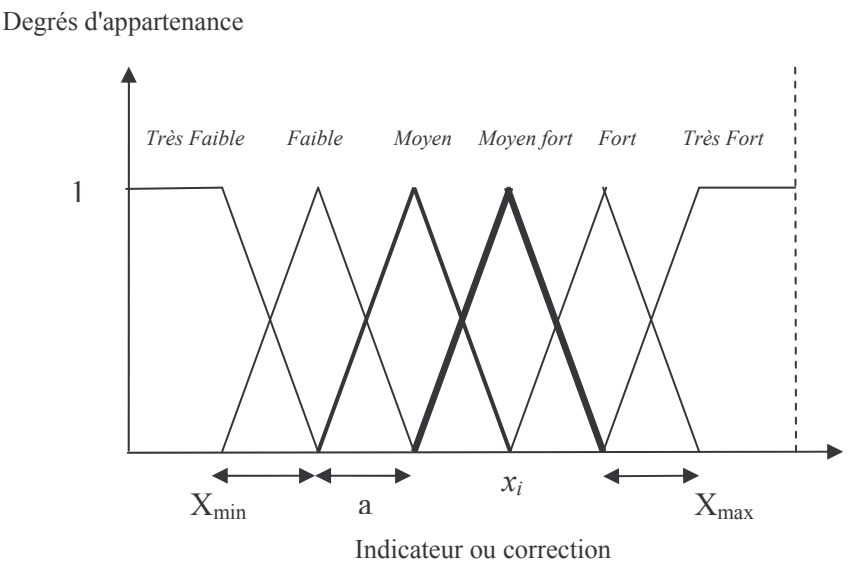


Figure 3.10 : Répartition uniforme des ensembles flous.

On définit la partition floue uniforme par le triplet N, x_{min}, et x_{max}, où N est le nombre d'ensembles, x_{min} et x_{max} sont les valeurs minimale et maximale définies sur la figure 3.10. Pour des valeurs inférieures ou supérieures à x_{min} et x_{max}, une valeur unité est attribuée respectivement à l'ensemble flou correspondant. Les formules suivantes expriment les abscisses x_i des extrémités des différents ensembles flous :

$$x_i = x_{\min} + (i-1)\,a\,; \quad i = 1,...,N \tag{3.3}$$

$$a = \frac{x_{\max} - x_{\min}}{N-1} \tag{3.4}$$

- Ensembles flous de type *linear gradual scaling*

Un cas particulier de partition non uniforme des ensembles flous est fourni par le type *linear gradual scaling* présenté sur la figure 3.11 suivante :

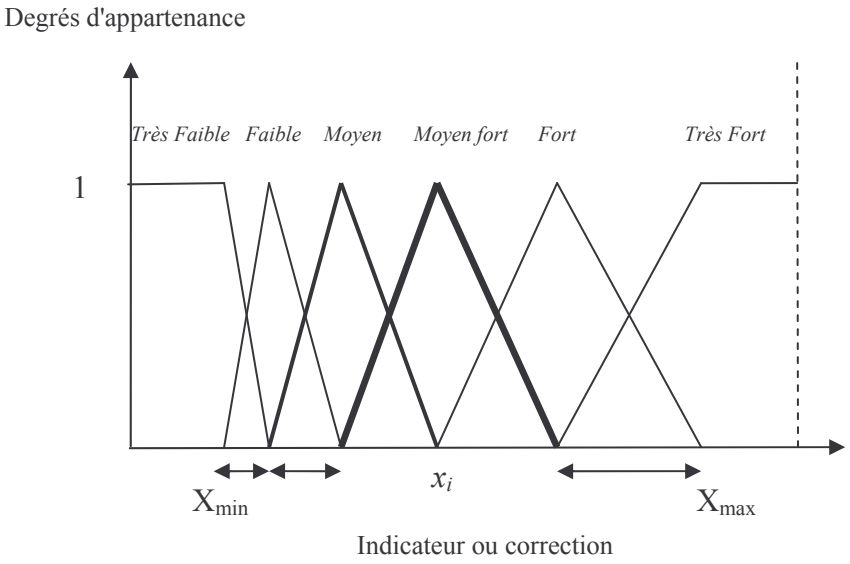


Figure 3.11 : Répartition de type *linear gradual scaling* des ensembles flous (dans ce cas $x_{min} = x_1$ et $x_{max} = x_n$).

On définit la partition floue de type *linear gradual scaling* par les variables N, x_{min}, x_{max}, et β, où N est le nombre d'ensembles, x_{min} et x_{max} sont les valeurs minimale et maximale, et β correspond au rapport entre le segment le plus grand et le segment le plus petit. Pour des valeurs inférieures ou supérieures à x_{min} et x_{max} respectivement, une valeur unité est attribuée à l'ensemble flou correspondant. Les formules suivantes expriment les abscisses x_i des extrémités des différents ensembles flous :

$$x_1 = x_{\min}$$
$$x_i = x_{\min} + (i-1)\,a + \frac{(i-2)\times(i-1)}{2}\,d; \quad i = 2,...,N \tag{3.5}$$

Où les variables a et d sont définies à partir des variables d'entrées, N, x_{min}, x_{max}, β, à l'aide des équations suivantes:

$$a = \frac{2}{(N-1)} \times \frac{(x_{\max} - x_{\min})}{(1+\beta)} \tag{3.6}$$

$$d = \frac{a \times (\beta - 1)}{(N-2)} \tag{3.7}$$

- Ensembles flous quelconques

La répartition uniforme est assez contraignante (peu de souplesse dans les modifications possibles). Cependant il est possible de reporter quasiment toute la variabilité du processus de contrôle sur les règles de décisions et dans ce cas les partitions uniforme et de type *linear gradual scaling* sont suffisantes. Mais sinon, il faudrait utiliser un type d'ensemble flou plus général, par exemple comme sur la figure 3.12.

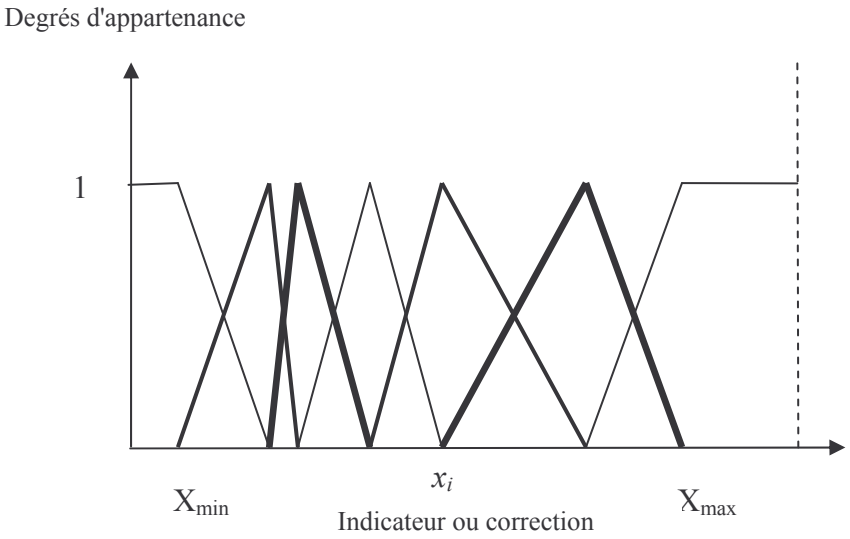


Figure 3.12 : Famille irrégulière de sept valeurs floues. Les valeurs de rang 1, 3, 5, 7 sont définies par 7 paramètres. Les autres s'en déduisent.

Définition des facteurs d'échelle

Les facteurs d'échelle permettent de ramener les processus de contrôle à des processus de contrôle génériques sans unité. Ainsi, on pourra normaliser les valeurs réelles mesurées de chaque indicateur de performance, et définir les ensembles flous correspondants en partitionnant l'ensemble [0,1].

3.3.2. Module de prise de décisions

3.3.2.1 Description de la base des règles (*Rule base*)

Dans le cadre d'un algorithme de contrôle fondé sur la logique floue, la définition des règles de paramétrage se fait dans un langage naturel à l'aide d'un ensemble de règles du type :

Si (état du processus) *alors* (sortie de contrôle).

Pour construire correctement notre *base de règles*, nous devons choisir respectivement :

- Les variables définissant les états du processus et les sorties de contrôle.
- Les énoncés des prémisses (partie *"Si…"*) et des conséquences (partie *"alors…"*) des règles de décisions.
- Les ensembles de variations des labels (ou variables linguistiques) décrivant l'état du processus et les variables de sortie.

Cet ensemble de choix et définitions permet d'induire l'ensemble des règles d'inférence du SIF.

Exemple de règles de décisions pour le paramétrage de l'algorithme de contrôle d'admission dans un réseau UMTS

Déduit des règles d'optimisation expérimentales déterminées les ingénieurs de design radio ou fruit d'une analyse de l'impact des paramètres RRM sur les indicateurs de performances, on énonce les règles linguistiques de paramétrage sous la forme :

Si (le taux de blocage est faible et le taux de coupure est très fort)
Alors (Baisser (fortement) le seuil d'admission de charge *downlink*)

Cette règle linguistique s'applique à la correction du paramètre seuil d'admission de charge cible *downlink*. Elle se fonde sur l'expérience du compromis coupure – blocage, nécessaire pour toute politique visant à contrôler l'algorithme d'admission. En effet, en baissant ce seuil, l'algorithme de contrôle d'admission va avoir tendance à refuser plus de mobiles demandant l'admission dans le réseau (à condition que la charge mesurée dans la cellule soit supérieure à la charge cible). En conséquence, le taux de blocage va a priori augmenter dans le réseau. La deuxième conséquence est que, du fait de la diminution du nombre de mobiles admis dans le réseau, le nombre moyen de mobiles présents dans le réseau va diminuer ce qui, a priori, va permettre une baisse du taux de coupure des mobiles. Il est important de remarquer que :

- Plus la diminution du seuil est importante et plus, a priori, les baisses et hausses respectives des taux de coupure et blocage sont grandes.
- Ce contrôle est réversible : en augmentant à nouveau le seuil on inversera les baisses et hausses respectives des taux de coupure et blocage.
- Ce contrôle est dans une certaine mesure symétrique : quand le taux de blocage est fort et le taux de coupure est faible, une hausse du seuil de charge cible *downlink* permet a priori de diminuer le blocage et augmenter le taux de coupure.

On a ainsi déterminé expérimentalement ou analytiquement plusieurs règles de paramétrage du seuil de charge cible *downlink* [Pic04] [Dub05], basées sur la situation de la station centrale, et qui sont résumées dans la table suivante :

Mesure de correction de la charge cible DL		Taux de coupure			
		Faible	Moyen	Fort	Très fort
Taux de blocage	Faible	Rien	Baisser (légèrement)	Baisser (Moyennement)	Baisser (Fortement)
	Moyen	Augmenter (légèrement)	?*	?	Baisser (?)
	Fort	Augmenter (Moyennement)	?	?	Baisser (?)
	Très fort	Augmenter (Fortement)	?	?	Baisser (?)

Table 3.1 : Correction de la charge cible DL en fonction des taux de blocage et de coupure pour la station centrale

Les points d'interrogations dans la table 3.1 ci-dessus signifient que l'on se trouve dans un cas intermédiaire de l'état du réseau et donc notre expérience ne permet pas de définir une correction à appliquer systématiquement dans ce cas-là. On pourra donc soit tester "à la main" plusieurs paramétrages possibles du CLF soit utiliser une stratégie d'optimisation décrite au chapitre 4. Deux situations sont possibles pour ces cas indéterminés:

- Soit l'état intermédiaire du réseau ne permet pas déterminer le type et l'amplitude de la correction du seuil.
- Soit on connaît le type de correction à faire (augmenter ou diminuer le seuil) mais on ne sait pas quelle est l'amplitude de la meilleure correction (faible, moyenne ou forte).

* Impossible de déterminer une correction à appliquer systématiquement dans cet état intermédiaire du réseau.

Impact de l'état des stations voisines sur le paramétrage du contrôle d'admission

Dans l'optique d'une amélioration du paramétrage des algorithmes RRM, il est nécessaire de tenir compte des indicateurs de performances de la station centrale, mais aussi de ceux des stations voisines. En effet, dans le cas des réseaux mobiles UMTS, comme il y a une fréquence porteuse unique (une fréquence pour le sens descendant et une pour le sens montant en FDD) dans le réseau et comme les signaux transmis dans les cellules voisines ne sont pas protégés par l'orthogonalité des codes de canalisation (*channelisation codes*, [Hol02]), les stations voisines génèrent une interférence brouilleuse sur la station centrale (et réciproquement). Donc l'état des stations voisines (le niveau de charge notamment) a une influence directe sur l'optimisation possible de la gestion de l'admission des mobiles par la cellule centrale (et réciproquement).

Etat des stations voisines	Peu chargées	Moyennement chargées	Fortement chargées
Contraintes pour la station centrale	Pas de contrainte exercée du fait de l'état des voisines	Nécessité de limiter la charge et l'interférence globales de la zone	Nécessité de diminuer la charge et l'interférence globales de la zone
Conséquences sur la gestion de l'admission des mobiles par la centrale	Pas d'influence sur la prise de décision par la station centrale	Incite la centrale à baisser légèrement la charge cible DL	Incite la centrale à baisser fortement la charge cible DL

Table 3.2 : Impact de l'état des stations voisines sur la gestion de l'admission des mobiles par la station centrale

Pour les stations voisines, on considère qu'une cellule est chargée dès lors que le taux de blocage ou le taux de coupure est élevé : dans les deux cas cette situation incitera la station centrale à diminuer son niveau de charge.

Prise en compte de l'écart entre la charge effective et le seuil de charge cible de l'itération précédente dans la cellule

Sur la figure 2.2 du chapitre 2, on observe les variations du seuil de charge cible lors d'un paramétrage automatique qui ne prend pas en compte l'écart entre la charge courante et le seuil de charge cible courant. Une telle stratégie de paramétrage automatique peut être sous optimale voire contreproductive.

Par exemple sur la figure 3.13, alors que la charge cible est déjà nettement supérieure à la charge courante de la cellule, le contrôleur prend la décision de continuer à augmenter la charge cible en se fondant uniquement sur les indicateurs de performances de la cellule (sans tenir compte de l'écart entre la charge et le seuil de charge cible courants). Le risque est donc de continuer à augmenter le seuil de charge cible DL, alors que cette décision est inefficace du point de vue des indicateurs de performance et diminue la réactivité du contrôleur : en effet le délai sera d'autant plus long si celui-ci doit ensuite diminuer la valeur de ce seuil. Ainsi sur la figure 3.13, le contrôleur continue d'augmenter le seuil de charge cible DL après le temps T_0, alors que le réseau continue à admettre des nouveaux mobiles dans la cellule, du fait que la charge DL est inférieure à la charge cible. Or à partir du temps T_1, comme on a admis beaucoup trop de mobiles dans la cellule, la station entre dans un état de fort taux de coupure et fort blocage. Il devient alors urgent de diminuer la charge de la station et donc de diminuer

le seuil de charge cible DL. Ce n'est pas avant le temps T_2 que la charge cible DL va redevenir inférieure à la charge de la station : on peut dire que nous avons un délai de réactivité du système égale à T_2 - T_1.

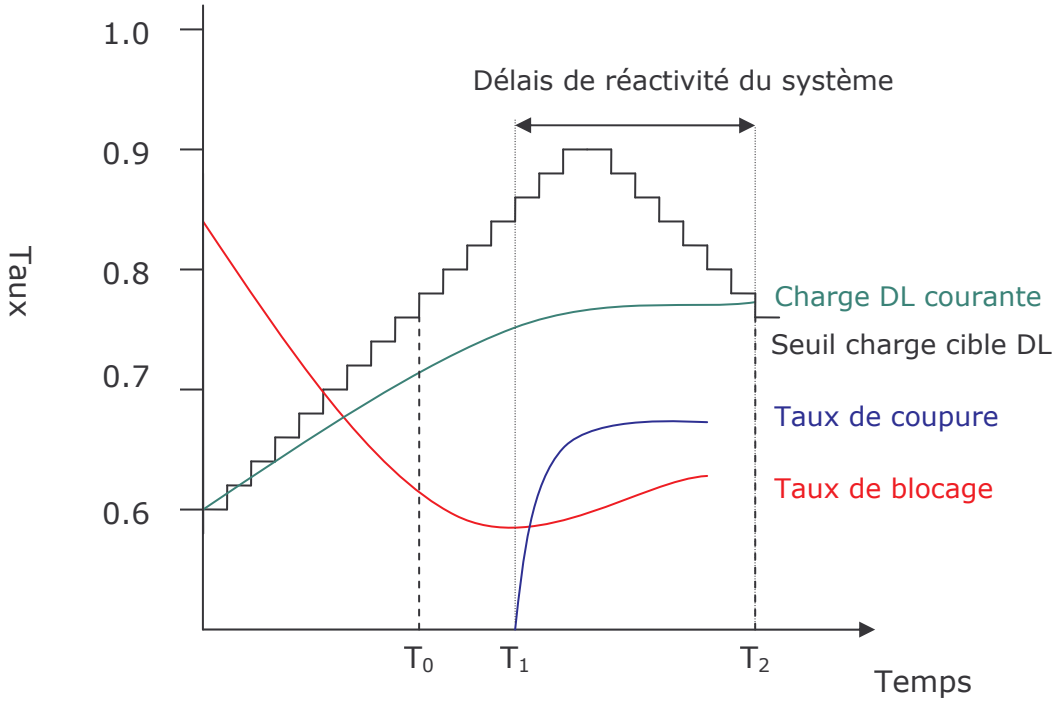


Figure 3.13 : Sans prise en compte de la différence entre la charge et le seuil de charge cible courants

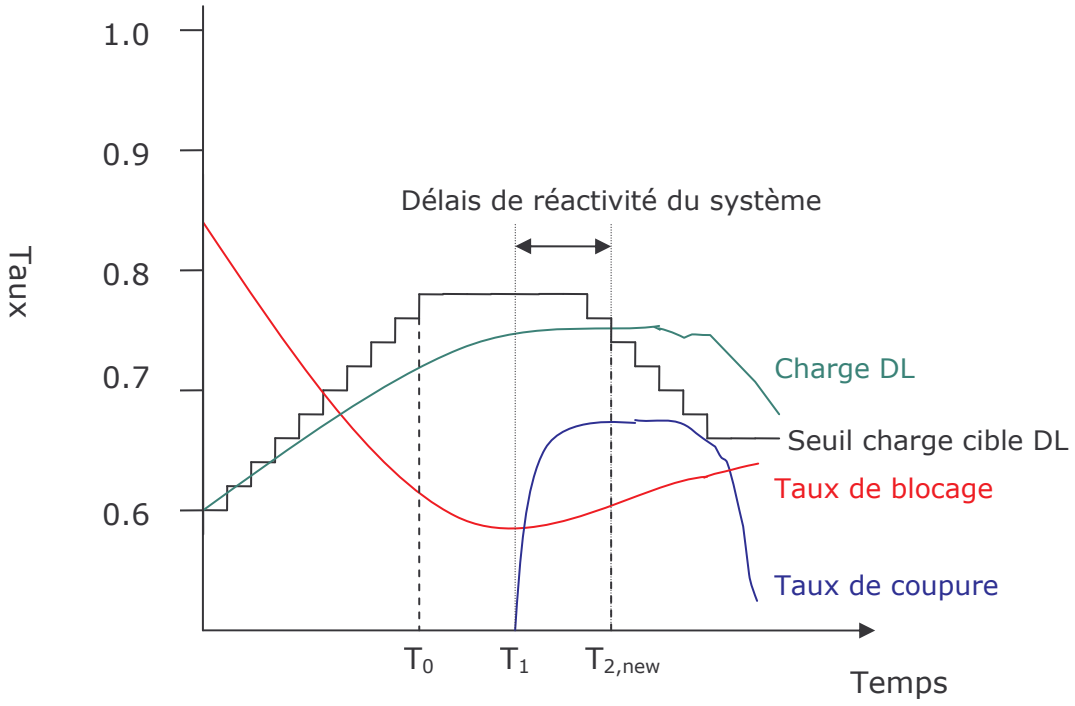


Figure 3.14 : Avec prise en compte de la différence entre la charge et le seuil de charge cible courants

En revanche sur la figure 3.14, pour déterminer la correction à l'itération courante Δc^n le contrôleur prend en compte la différence entre la charge de la cellule et le seuil de charge

cible courants, $Load^n - Load_{cible}^{n-1}$, ainsi que la correction instantanée $\Delta c_{Inst.}^{n}$ * estimée pour l'état courant. A partir du temps T_0, le contrôleur n'augmente plus la charge cible DL, étant donné que la charge cible est déjà nettement supérieure à la charge DL, et que les corrections précédentes avaient déjà augmenté ce paramètre. Comme précédemment, vu que la charge DL est inférieure à la charge cible, le réseau continue à admettre des nouveaux mobiles dans la cellule. De nouveau, à partir du temps T_1, à cause du nombre de mobiles actifs dans la cellule, la station entre dans un état de fort taux de coupure et fort blocage. Le contrôleur diminue alors le seuil de charge cible DL. Mais à présent, c'est au temps $T_{2,new}$ que la charge cible DL redevient inférieure à la charge de la station : le délai de réactivité du contrôleur égal à $T_{2,new}$ - T_1 a donc été nettement diminué par rapport à la situation décrite sur la figure 3.13.

En conclusion, il semble essentiel d'introduire dans le système de contrôle par logique floue du seuil de charge cible DL un moyen de prendre en compte l'écart entre la charge et la valeur de la charge cible courants dans la cellule.

Exemple de règles de décisions pour le paramétrage de l'algorithme de MD

Comme pour l'algorithme d'admission ci-dessus, nous allons à présent résumer certaines règles de contrôle de l'algorithme de macro-diversité qui ont pu être déduites empiriquement ou analytiquement par les ingénieurs de gestion du réseau (cf. rapport de stage de V. Diascorn [Dia04] et [Dub05]). Ainsi dans le cas du contrôle du paramètre de fenêtre d'ajout *Add_win*, l'analyse de l'influence de ce paramètre sur les indicateurs de performances nous a permis de déterminer deux types de règles linguistiques qui donneront deux matrices de décisions différentes : les règles linguistiques locales et les règles linguistiques globales.

Approche locale de la gestion du paramètre *Add_win* de macro-diversité

- Tout d'abord, on considère le cas d'une station centrale chargée, c'est-à-dire avec un fort taux de blocage des liens créés par macro-diversité, dans une région peu chargée.

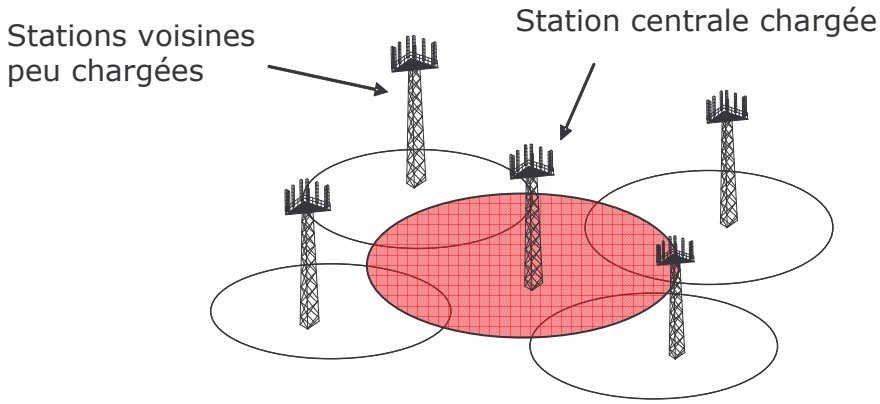


Figure 3.15 : Cas d'une station centrale chargée avec des voisines peu chargées

En augmentant la taille de la fenêtre d'ajout de la station centrale, cela permet aux mobiles de la cellule de créer plus facilement des liens de macro-diversité. Ainsi le taux de blocage des liens créés par MD de la station centrale diminue. Inversement, si on diminue la taille de la fenêtre d'ajout des stations voisines, cela implique que la station centrale sera moins

* Correction "instantanée" du paramètre de seuil de charge cible DL dépendant seulement de l'état sur la zone à l'itération courante. Cette valeur est une correction sans mémoire (ne tenant pas compte de la correction Δc^{n-1} réalisée pour la station centrale à l'itération précédente).

facilement ajoutée à l'*Active Set* des mobiles servis par les cellules voisines. En conséquence, le taux de blocage des liens créés par MD des stations voisines va augmenter un peu, et celui de la centrale va a priori diminuer. On déduit de ces constatations deux règles linguistiques pour la gestion de l'algorithme de macro-diversité :

Si (le taux de blocage des liens MD de la centrale est fort
et le maximum du taux de blocage des liens MD des voisines est faible)

Alors (Augmenter la taille de la fenêtre d'ajout *Add_win* de la centrale).

Si (le taux de blocage des liens MD de la centrale est faible
et le maximum du taux de blocage des liens MD des voisines est fort)

Alors (Diminuer la taille de la fenêtre d'ajout *Add_win* de la centrale).

- On considère à présent le cas d'une station centrale chargée avec une majorité de stations voisines déchargées et quelques voisines chargées.

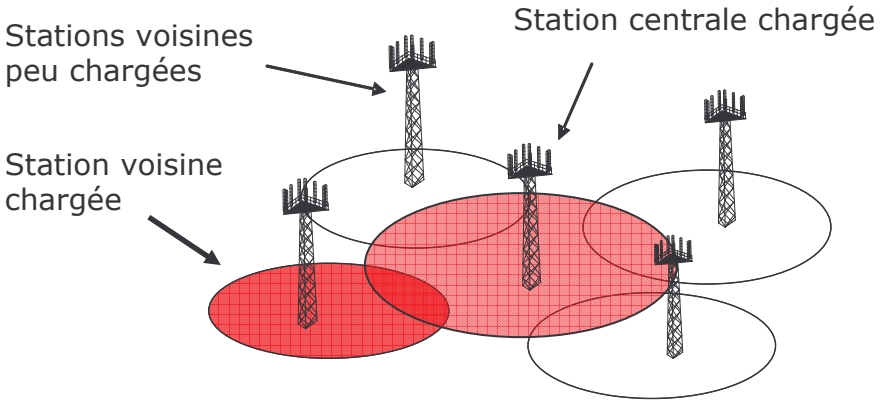


Figure 3.16 : Cas d'une station centrale chargée avec des voisines peu chargées et des voisines chargées

Si on augmentait la taille de la fenêtre d'ajout de la station centrale, cela permettait aux mobiles de la cellule de créer plus facilement des liens de macro-diversité. Mais le risque serait alors de dégrader le taux de blocage des liens MD des stations voisines chargées: dans ce cas on ne modifie donc pas le paramètre *Add_win* de la station centrale. Par ailleurs, si on diminue la taille de la fenêtre d'ajout des stations voisines peu chargées, cela permet que la station centrale soit moins facilement ajoutée à l'*Active Set* des mobiles servis par ces cellules voisines. Par conséquent, le taux de blocage des liens créés par MD des stations voisines peu chargées va augmenter un peu, et celui de la centrale va a priori diminuer. On déduit de la première constatation une nouvelle règle linguistique pour la gestion de l'algorithme de macro-diversité :

Si (le taux de blocage des liens MD de la centrale est fort
et le maximum du taux de blocage des liens MD des voisines est fort)

Alors (Ne pas modifier la taille de la fenêtre d'ajout *Add_win* de la centrale).

En revanche, la deuxième constatation revient à une règle linguistique énoncée au paragraphe précédent, c'est-à-dire :

Si (le taux de blocage des liens MD de la centrale est faible
et le maximum du taux de blocage de liens MD des voisines est fort)

Alors (Diminuer la taille de la fenêtre d'ajout *Add_win* de la centrale).

- Enfin, on considère le cas d'une station centrale peu chargée dans une région peu chargée. Dans une approche locale de la gestion du paramètre *Add_win*, on choisit de ne pas modifier la valeur de ce paramètre. On en déduit donc la règle linguistique suivante :

Si (le taux de blocage des liens MD de la centrale est faible
et le maximum du taux de blocage des liens MD des voisines est faible)

Alors (Ne pas modifier la taille de la fenêtre d'ajout *Add_win* de la centrale).

On a ainsi déterminé expérimentalement ou analytiquement plusieurs règles fondées sur une approche locale de paramétrage de la taille de la fenêtre d'ajout *Add_win*, basées sur les performances de la station centrale et de la plus chargée des stations voisines, et qui sont résumées dans la table suivante :

Gestion du paramètre *Add_win* de la centrale	Etat de la station centrale en terme de taux de blocage des liens MD		
Etat de la pire station voisine, en terme de taux de blocage des liens MD	Faible	Moyen	Fort
Faible	Ne pas modifier *Add_win*	?*	Augmenter *Add_win*
Moyen	?	?	?
Fort	Diminuer *Add_win*	?	Ne pas modifier *Add_win*

Table 3.3 : Approche locale de gestion du paramètre *Add_win*.

Les points d'interrogations dans la table 3.3 ci-dessus signifient que l'on se trouve dans un cas intermédiaire de l'état local du réseau et donc notre expérience ne permet pas de définir une correction à appliquer systématiquement dans ce cas-là. On pourra donc soit tester "à la main" plusieurs paramétrages possibles du CLF soit utiliser une stratégie d'optimisation parmi celles décrites au chapitre 4.

Approche globale de la gestion du paramètre *Add_win* de macro-diversité

- Tout d'abord, on considère le cas d'une station centrale chargée dans une région globalement chargée, c'est-à-dire avec un fort taux moyen de blocage des liens créés par macro-diversité.

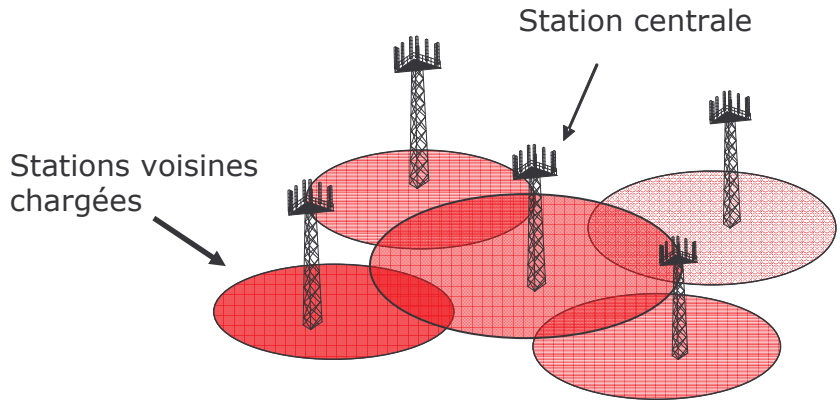


Figure 3.17 : Cas d'une station centrale chargée dans une région globalement chargée

* Impossible de déterminer une correction à appliquer systématiquement dans cet état intermédiaire du réseau.

Dans une approche globale de gestion du paramètre *Add_win*, il est essentiel pour une telle situation de diminuer la charge globale sur la zone. Dans ce but, nous proposons de réduire le paramètre *Add_win* pour la station centrale et l'ensemble des stations chargées de la région. Ainsi, la charge globale de la zone va a priori décroître, ce qui libérera une partie de la puissance DL des stations chargées. De cette manière, on pourra globalement faciliter la création de nouveaux liens de MD et donc diminuer le taux de blocage des liens de MD. On déduit de ces constatations une première règle linguistique de gestion de l'algorithme de macro-diversité dans une approche globale, qui permet de diminuer la charge de la station centrale (pour chaque voisine, sa charge sera ensuite diminuée par l'application de cette même règle, une fois qu'elle sera considérée à son tour comme la station centrale chargée d'une zone chargée) :

Si (La moyenne du taux de blocage des liens MD sur l'ensemble des voisines est forte

et la charge *downlink* de la station centrale est forte)

Alors (Diminuer la taille de la fenêtre d'ajout *Add_win* de la centrale).

Dans la règle ci-dessus, on évalue la charge de la station centrale à l'aide de l'indicateur de charge *downlink* (rapport entre la puissance DL émise par le secteur et la puissance DL maximale).

- On considère à présent le cas d'une station centrale peu chargée dans une région globalement peu chargée. A priori, dans une telle situation, il n'est pas nécessaire de modifier la taille de la fenêtre d'ajout *Add_win* de la centrale. Cependant, une stratégie d'anticipation des risques peut consister à profiter de cet état favorable du réseau pour relever la valeur du paramètre *Add_win* jusqu'à une valeur cible moyenne et se réserver ainsi une certaine marge de manœuvre dans les futures situations de charge dans la zone. En effet sans cela, étant donné que l'on a tendance à diminuer la valeur de l'*Add_win* d'une station dès qu'une de ses voisines est chargée, le risque est de saturer en permanence (et même quand cela n'est pas nécessaire) à la valeur minimale de ce paramètre. Ainsi on déduit une nouvelle règle linguistique de gestion de l'algorithme de macro-diversité dans une approche globale :

Si (La moyenne du taux de blocage des liens MD sur l'ensemble des voisines est faible
et la charge *downlink* de la station centrale est faible)

Alors (Augmenter la taille de la fenêtre d'ajout *Add_win* de la centrale).

Les algorithmes d'optimisation des CLF vont permettre d'explorer et sélectionner différents types de gestion globale de l'algorithme de macro-diversité pour les situations intermédiaires des performances du réseau. Dans ce but, au paragraphe 3.3.2.2, nous allons décrire une méthode pour transformer les règles linguistiques énoncées ci-dessus de gestion des algorithmes RRM, en règles d'inférences mathématiques exprimées sous la forme d'un système paramétrable de matrices de décisions. Ce système de matrices composera le moteur d'inférence de notre modèle de contrôleur FKBC.

Prise en compte des stations voisines dans le simulateur de réseau mobile

Vue l'importance des stations voisines dans les règles linguistiques énoncées ci-dessus, nous expliquons dans ce paragraphe comment les stations voisines d'une centrale ont été définies et prises en compte dans un simulateur de réseaux mobiles développé en interne par France Télécom Recherche & Développement.

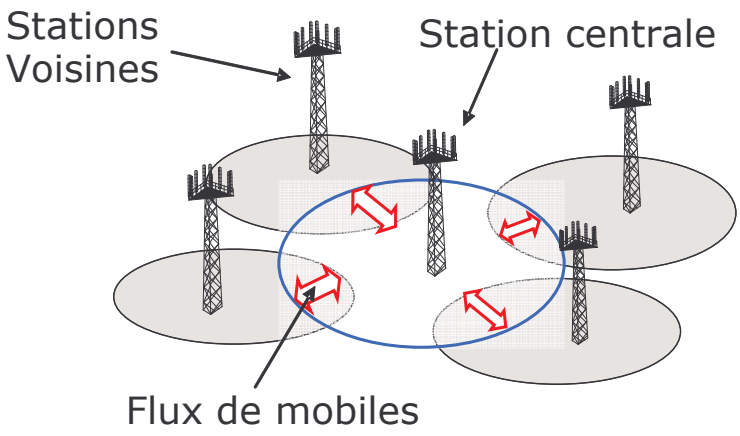


Figure 3.18 : Définition des voisines de la station centrale

Pour sélectionner les stations voisines d'une station centrale dans notre simulateur de réseaux mobiles, on peut :

1. Soit, comme indiqué dans la norme et testé sur les réseaux expérimentaux, établir pour chaque station une liste de cellules voisines déclarées.
2. Soit sélectionner les voisines en fonction des stations ayant des flux de mobiles avec la station centrale.

Dans le cas 2, on peut ainsi estimer des coefficients de pondération de la contribution de chaque voisine à la station centrale, en se fondant sur les quantités de flux de trafic échangées entre la voisine et la station centrale.

Dans le module d'ingénierie automatique (contrôle des algorithmes RRM), il est alors possible de tenir compte de l'influence des stations voisines sur le paramétrage de la centrale, soit à travers une moyenne simple des indicateurs de coupure et blocage (cas 1), soit à travers une moyenne pondérée utilisant les coefficients de pondération déterminés à partir des flux de trafic (cas 2).

3.3.2.2 Description du moteur d'inférence

Pour construire correctement notre *moteur d'inférence*, nous devons choisir respectivement :

- Le type de moteur d'inférence.
- L'expression mathématique d'une règle d'inférence seule.
- L'expression mathématique d'un ensemble de règles d'inférence.

Ci-après, nous allons voir comment un ensemble de règles de contrôle de la forme *"Si ... Alors ..."* peut être transformé en un système de matrices de décisions. Les matrices de décisions sont en effet une manière simple, pratique et synthétique pour énoncer des règles de contrôle. De plus, dans l'optique d'une optimisation du processus de contrôle, les matrices de décisions peuvent être décrites sous des formes paramétriques simples.

Choix du type de moteur d'inférence

Deux types de moteur d'inférence sont possibles dans un contrôle de type FKBC :

- L'inférence fondée sur *la composition (composition based inference)* : les règles élémentaires sont fusionnées en une seule règle globale (qui pourrait être traduite

mathématiquement par une hypermatrice de décisions) et une relation de composition permet d'appliquer cette règle globale au vecteur d'état *fuzzifié* du système.

- L'inférence fondée sur *les règles individuelles (individual-rule based inference)* : on commence par appliquer chaque règle élémentaires aux différents indicateurs *fuzzifiés* du système, et on combine les différents corrections floues obtenues à l'aide d'un opérateur d'agrégation d'ensembles flous (Max - min, somme - produit…)

Nous préférons le second type d'inférence, car elle permet de conserver la structure des règles élémentaires. En effet, après l'optimisation d'un contrôleur de logique floue (chapitre 4), cette structure d'inférence pourrait nous permettre de déduire d'un CLF optimisé, de nouvelles règles élémentaires efficaces de management des réseaux mobiles UMTS qui n'avaient pas encore été identifiées par les experts radio.

Opérateurs d'agrégation

Avec un moteur d'inférence fondée sur *les règles individuelles*, nous devons déterminer comment agréger plusieurs valeurs floues en une nouvelle valeur floue. Ci-dessous, nous appellerons R_i les règles de décisions élémentaires, D un opérateur d'agrégation générique, a et b respectivement les première et seconde variables floues à agréger et c la valeur floue agrégée.

Nous supposons que les règles de décisions sont telles que :

$$\begin{aligned} &A_1\ R_1\ B_1 \rightarrow C_1 \\ &A_1\ R_2\ B_2 \rightarrow C_2 \\ &A_2\ R_3\ B_1 \rightarrow C_3 \\ &A_2\ R_4\ B_2 \rightarrow C_2 \end{aligned} \tag{3.8}$$

Nous devons donc déterminer les degrés d'appartenance de la variable floue c au labels C_1, C_2 et C_3 à partir des degrés d'appartenance élémentaires des variables a et b, et de l'opérateur d'agrégation. La correspondance logique des règles de décisions indiquées ci-dessus donne :

$$\begin{aligned} &C_1 \Leftrightarrow \left[A_1\ ET\ B_1\right] \\ &C_2 \Leftrightarrow \left[A_1\ ET\ B_2\right] OU \left[A_2\ ET\ B_2\right] \\ &C_3 \Leftrightarrow \left[A_2\ ET\ B_1\right] \end{aligned} \tag{3.9}$$

Différents opérateurs sont utilisables dans notre contexte :

- L'opérateur Maximum – Minimum : cet opérateur d'agrégation remplace l'opérateur logique ET par la fonction minimum et l'opérateur logique OU par la fonction maximum :

$$c_2 = Max\left(Min\left(a_1, b_2\right), Min\left(a_2, b_2\right)\right) \tag{3.10}$$

- L'opérateur Somme – Produit : cet opérateur d'agrégation remplace l'opérateur logique ET par la fonction produit et l'opérateur logique OU par la fonction somme :

$$c_2 = a_1 \times b_2 + a_2 \times b_2 \tag{3.11}$$

- L'opérateur Maximum – Produit : cet opérateur d'agrégation remplace l'opérateur logique ET par la fonction produit et l'opérateur logique OU par la fonction maximum :

$$c_2 = Max\left(a_1 \times b_2, a_2 \times b_2\right) \tag{3.12}$$

Dans les CLF, nous utiliserons plutôt l'opérateur d'agrégation somme – produit, car il s'agit d'un opérateur de type centre de gravité.

Expression mathématique d'une règle d'inférence

Il s'agit à présent d'exprimer mathématiquement les règles d'inférences linguistiques introduites au paragraphe 3.3.2.1 pour le paramétrage des algorithmes de contrôle d'admission et de macro-diversité: pour cela, nous allons introduire et définir les matrices de décisions pour le contrôle des algorithmes RRM déduites de la base des règles linguistiques.

Pour la table 3.1, il semble naturel de transformer les règles énoncées en une matrice de décisions dont les lignes et colonnes correspondent respectivement aux labels des taux de blocage et de coupure. Dans ce modèle, chaque élément de la matrice correspond à un label de correction induit en fonction du couple de labels blocage – coupure en entrées.

Labels en entrées	Coupure faible	Coupure moyen	Coupure fort	Coupure très fort
Blocage faible	Rien	Baisser (légèrement)	Baisser (Moyennement)	Baisser (Fortement)
Blocage moyen	Augmenter (légèrement)	?*	?	Baisser (?)
Blocage fort	Augmenter (Moyennement)	?	?	Baisser (?)
Blocage très fort	Augmenter (Fortement)	?	?	Baisser (?)

Table 3.4 : Corrections du seuil de charge cible DL en fonction du couple de labels blocage – coupure pour la station.

En examinant l'ensemble des couples de labels flous coupure – blocage, on construit ci-dessous la matrice de décisions permettant de déterminer la correction $\Delta c_{Centrale}$ qui tient compte seulement de l'état de la station centrale, dans le cadre du contrôle de l'algorithme d'admission. Par rapport à la table 3.4, on introduit ci-dessous les ensembles flous de correction *EC1*,…, *EC7* correspondant respectivement aux corrections linguistiques: *Baisser (Fortement),..., Augmenter (Fortement)*. La matrice de décisions a ensuite été complétée en fixant les termes de la diagonale à zéro (Ensemble de correction *EC4*) et en déterminant les éléments restants par interpolation linéaire. La matrice finale (table 3.5) est donc anti-symétrique :

ET	ED1	ED2	ED3	ED4
EB1	***EC4*** **0**	***EC3*** **-**	***EC2*** **--**	***EC1*** **---**
EB2	***EC5*** **+**	EC4 0	EC3 -	***EC2*** **--**
EB3	***EC6*** **++**	EC5 +	EC4 0	***EC3*** **-**
EB4	***EC7*** **+++**	EC6 ++	EC5 +	EC4 0

Table 3.5 : Matrice de décision pour la station centrale (en gras les éléments directement déduits des règles).

* Impossible de déterminer une correction à appliquer systématiquement dans cet état intermédiaire du réseau.

Dans la table 3.5, les ensembles flous de blocage sont notés EB, ceux de coupure ED (*dropping*) et les ensembles flous de correction EC. On formalise ainsi sous la forme de matrices de décisions les règles linguistiques énoncées au paragraphe 3.3.2.1.

Matrices de décisions pour le paramétrage du contrôle d'admission

De cette façon, on introduit à partir des règles linguistiques, énoncées au 3.3.2.1, trois matrices de décisions pour le paramétrage automatique du contrôle d'admission :

- La matrice qui détermine, à partir des labels de blocage (lignes) et coupure (colonnes) pour la station centrale, la correction $\Delta c_{Centrale}$. Cette correction ne tient pas compte de l'état des voisines (cas d'une station isolée).
- La matrice qui détermine, à partir des labels de blocage (lignes) et coupure (colonnes) agrégés (moyenne ou moyenne pondérée) sur l'ensemble des voisines, la correction $\Delta c_{Voisines}$. Cette valeur $\Delta c_{Voisines}$ estime l'impact de l'état des stations voisines sur la correction à appliquer au paramétrage de la station centrale.
- En fusionnant (voir les détails au paragraphe suivant) ces corrections $\Delta c_{Centrale}$ et $\Delta c_{Voisines}$, on obtient la valeur floue de correction instantanée Δc_{Inst} du paramètre de seuil de charge cible DL tenant seulement compte de l'état sur la zone à l'itération courante. Cette valeur $\Delta c_{Inst.}$ est une correction instantanée (ou sans mémoire), c'est-à-dire qu'elle ne tient pas compte de la correction réalisée pour la station centrale à l'itération précédente.
- La matrice qui détermine, en fonction de la valeur floue de correction instantanée Δc_{Inst} et de l'écart entre la charge effective et le seuil de charge cible de l'itération précédente, la valeur floue de correction finale Δc_{Finale}.

Matrices de décisions pour le paramétrage de la macro-diversité

On applique la même méthode pour définir les deux matrices de décisions déduites des règles linguistiques de paramétrage de l'algorithme de macro-diversité : la matrice correspondant à une approche locale et celle correspondant à une approche globale.

- La matrice qui détermine, à partir du maximum des taux de blocage des liens MD des voisines (lignes) et du taux de blocage des liens MD de la centrale (colonnes), la correction de la taille de fenêtre d'ajout de la cellule Δc_{Locale} dans une approche locale.
- La matrice qui détermine, à partir de la moyenne du taux de blocage des liens MD sur l'ensemble des voisines (lignes) et de la charge *downlink* de la station centrale, la correction de la taille de fenêtre d'ajout de la cellule $\Delta c_{Globale}$ dans une approche globale.
- En fusionnant (voir les détails au paragraphe suivant) ces corrections Δc_{Locale} et $\Delta c_{Globale}$ on obtient la valeur floue de correction Δc_{Totale} de la taille de la fenêtre d'ajout *Add_win*.

Par exemple, la table 3.6 ci-dessous donne un exemple de matrice de correction de la taille de fenêtre d'ajout de la cellule Δc_{Locale} dans une approche locale (la matrice de décisions a été complétée par symétrie en fixant un seul ensemble de correction par diagonale) :

ET	EB MD1	EB MD2	EB MD3	EB MD4	EB MD5
EB MD1 Max. Vois.	**EC Local 3 0**	EC Local 4 +	EC Local 4 +	EC Local 5 ++	***EC Local 5*** **++**
EB MD2 Max. Vois.	EC Local 3 0	EC Local 3 0	EC Local 4 +	EC Local 4 +	EC Local 5 ++
EB MD3 Max. Vois.	EC Local 2 -	EC Local 3 0	EC Local 3 0	EC Local 4 +	EC Local 4 +
EB MD4 Max. Vois.	EC Local 2 -	EC Local 2 -	EC Local 3 0	EC Local 3 0	EC Local 4 +
EB MD5 Max. Vois.	***EC Local 1*** **--**	EC Local 2 -	EC Local 2 -	EC Local 3 0	**EC Local 3 0**

Table 3.6 : Matrice de décisions pour l'approche locale de gestion de la macro-diversité (en gras les éléments directement déduits des règles)

Dans le paragraphe suivant, on propose une formulation paramétrique des matrices de décisions introduites ci-dessus. Une fois les matrices de décisions écrites sous cette forme paramétrique simple, on pourra améliorer l'efficacité du système de matrices de contrôle des algorithmes RRM à l'aide d'un des processus d'optimisation (décrits au chapitre 4) : optimisation par essaim particulaire ou apprentissage par renforcement.

Représentation paramétrique des matrices

Dans l'optique d'une optimisation du processus de contrôle, on propose plusieurs représentations paramétriques simples des matrices de décisions:

- Représentation paramétrique par interpolation linéaire des extrémités des matrices

Dans cette approche, on paramètre les quatre sommets de la matrice de décision, et les points intermédiaires sont interpolés linéairement. Les sommets et éléments de la matrice sont définis par des entiers relatifs correspondant à des numéros d'ensembles flous. Cela introduit donc des contraintes de granularité pour les paramètres représentatifs de la matrice. Ainsi par exemple, la matrice de décision pour la station centrale introduite au paragraphe précédent (pour le paramétrage du contrôle d'admission) peut être représentée par:

0	*-1*	*-2*	***-3***
+1	*0*	*-1*	*-2*
+2	*+1*	*0*	*-1*
+3	*+2*	*+1*	***0***

Figure 3.19 : Matrice de décisions pour la station centrale définie à l'aide de ses quatre sommets

La matrice de décision pour le contrôle de décision est donc paramétrée par *{0, -3, +3, 0}*. L'espace de recherche pour l'optimisation de cette matrice paramétrée par ses quatre sommets sera donc de la forme : *{-3, -2, -1, 0, +1, +2, +3}*. Un tel paramétrage des matrices de décisions permet donc de réduire considérablement la taille de l'espace de recherche lors du processus d'optimisation (cf. chapitre 4).

- Autres représentations paramétriques des matrices de décisions.

Deux autres représentations paramétriques des matrices de décisions sont proposées en annexes du mémoire : détermination des éléments de la matrice par combinaison linéaire de deux matrices élémentaires ou par un calcul de distance à un point[*].

[*] Voir la section 10.2 des annexes.

- Introduction de contraintes dans la représentation paramétrique des matrices.

Nous avons expliqué ci-dessus qu'il pouvait exister une contrainte de *granularité* (nombres entiers par exemple) pour les éléments de la matrice de décision. Par ailleurs, il peut être aussi nécessaire de contraindre ces éléments à varier lors de l'optimisation dans un intervalle limité (correspondant par exemple au nombre d'ensembles flous pris en compte). Ces deux contraintes seront prises en compte lors de l'optimisation des contrôleurs.

Par ailleurs, à partir de l'expérience des ingénieurs radio et de l'analyse des algorithmes RRM formalisées sous forme de règles linguistiques, il est possible de déduire des contraintes de relations d'ordres entre les éléments et sommets des matrices de décisions :

- Pour l'algorithme de contrôle d'admission, dans une situation de très fort blocage et de faible taux de coupure (sommet bas et gauche) les tables 3.4 et 3.5 recommandent d'augmenter très fortement le seuil de charge cible de la centrale. A priori aucun autre élément de la matrice de décisions ne doit permettre une augmentation plus forte du paramètre. Inversement, dans une situation de très fort blocage et de faible taux de coupure (sommet haut et droit), les tables 3.4 et 3.5 recommandent de diminuer très fortement le seuil de charge cible de la centrale. A priori aucun autre élément de la matrice de décisions ne doit permettre une augmentation plus forte du paramètre. On en déduit donc des contraintes de relations d'ordre entre les paramètres de la matrice $\left[m_{ij}^{Centrale}\right]$ de taille $n_{Bloc.} \times n_{Coup.}$ ($n_{Bloc.}$ étant le nombre de lignes et $n_{Coup.}$ le nombre de colonnes) : $\forall i \in \left\{1,...,n_{Bloc.}\right\},\ \forall j \in \left\{1,...,n_{Coup.}\right\},\ m_{1n_{Coup.}}^{Centrale} \le m_{ij}^{Centrale} \le m_{n_{Bloc.}1}^{Centrale}$.

- Pour l'algorithme de macro-diversité, dans une situation de faible taux de blocage des liens de MD sur la centrale et de fort taux maximum de blocage de liens de MD sur l'ensemble des voisines (sommet bas et gauche), la table 3.6 recommande de diminuer très fortement la taille de la fenêtre d'ajout A*dd_win* de la centrale. A priori aucun autre élément de la matrice de décisions ne doit permettre une diminution plus forte du paramètre. Inversement, dans une situation de fort taux de blocage des liens de MD sur la centrale et de faible taux maximum de blocage de liens de MD sur l'ensemble des voisines (sommet haut et droit), la table 3.6 recommande d'augmenter très fortement la taille de la fenêtre d'ajout A*dd_win* de la centrale. A priori aucun autre élément de la matrice de décisions ne doit permettre une augmentation plus forte du paramètre. On en déduit donc des contraintes de relations d'ordre entre les paramètres de la matrice $\left[m_{ij}^{Locale}\right]$ de taille $n_{Bloc.} \times n_{Bloc.}$ ($n_{Bloc.}^{MD}$ étant le nombre de lignes et colonnes de la matrice) : $\forall i, j \in \left\{1,...,n_{Bloc.}^{MD}\right\}$, $m_{n_{Bloc.}^{MD}1}^{Locale} \le m_{ij}^{Locale} \le m_{1n_{Bloc.}^{MD}}^{Locale}$.

Pour chaque matrice du système d'inférence, on déduit donc un ensemble de contraintes de granularité, d'intervalles et de relations d'ordres que les paramètres de la matrice devront respecter au cours du processus d'optimisation. Il est aussi possible de relâcher certaines des contraintes de relation d'ordre pour vérifier qu'il n'y ait pas de solution optimale hors du cadre de compréhension introduit ci-dessus.

L'expression mathématique d'un ensemble de règles d'inférence

A partir des matrices de décisions déduites des règles linguistiques de la base de données, la structure des moteurs d'inférence pour le contrôle d'admission et la macro-diversité doit être déterminée. Pour atteindre cet objectif, il nous reste à spécifier les règles permettant de fusionner les corrections floues élémentaires calculées à partir des différents indicateurs de performances: par exemple comment va-t-on réaliser le compromis entre les corrections

$\Delta c_{Centrale}$ et $\Delta c_{Voisines}$ qui tiennent compte respectivement de l'état de la centrale et de l'état des voisines pour le contrôle du seuil de charge cible DL ?

La solution la plus simple est d'introduire une matrice de compromis qui reçoive en entrées le couple de corrections élémentaires, par exemple le couple $(\Delta c_{Centrale}, \Delta c_{Voisines})$, et détermine la en sortie la correction floue globale Δc_{Inst} à appliquer à la station centrale. Le nombre de lignes et colonnes de cette matrice sont respectivement égaux aux nombres d'ensembles flous pour la première correction élémentaire et pour la deuxième correction élémentaire.

Exemple pour la gestion du paramètre de charge cible DL

Nous introduisons donc une matrice permettant de réaliser le compromis entre la correction déduite de l'état de la station centrale et l'état des stations voisines. Nous supposons que trois états moyens sont définis pour les voisines qui correspondent (voir table 3.2) à trois labels d'influence des voisines $\Delta c_{Voisines}$: {*Baisse forte, baisse moyenne, pas de modification*}.

Pour la correction $\Delta c_{Centrale}$, trois labels sont définis dans notre exemple {*Baisser, pas de modification, Augmenter*}. La table 3.7 suivante donne un exemple de compromis que l'on peut qualifier d'équitable entre la centrale et ses voisines :

Compromis Centrale / voisines	**$EC_{Voisines}0$ Baisse forte**	**$EC_{Voisines}1$ Baisse moyenne**	**$EC_{Voisines}2$ Pas de modif.**
EC0 Baisser	EC0 Baisser	EC0 Baisser	EC0 Baisser
EC1 Pas de modif.	EC0 Baisser	EC0 Baisser	EC1 Pas de modif.
EC2 Augmenter	EC0 Baisser	EC1 Pas de modif.	EC2 Augmenter

Table 3.7 : Matrice de compromis entre la centrale et ses voisines pour la correction du seuil de charge cible DL.

De la même façon dans le cas de la macro-diversité, nous introduisons une matrice permettant de fusionner les corrections floues élémentaires Δc_{Locale} et $\Delta c_{Globale}$ et obtenir la valeur floue de correction Δc_{Totale} de la taille de la fenêtre d'ajout *Add_win*. Ces matrices de compromis sont paramétrées de la même façon que les autres matrices du moteur d'inférence (voir figure 3.19). On pourra donc les intégrer au processus d'optimisation par essaim de particules.

3.3.2.3 Description de deux moteurs d'inférence pour le contrôle de paramètres RRM

L'objectif global du contrôleur est d'adapter un des paramètres RRM sélectionnés*, en fonction des indicateurs de QoS identifiés (pour la station centrale et ses voisines) en vue d'améliorer les performances du réseau. En entrées du contrôleur CLF, sont disponibles les indicateurs de QoS filtrés pour la station et ses voisines ainsi que les modifications faites à l'itération précédente du paramétrage RRM. En sorties du contrôleur, on détermine de nouvelles corrections du paramètre RRM identifié. Ces ajustements de paramétrage sont fondés sur une approche de contrôle par logique floue. Ci-dessous, nous décrivons les deux

* Le seuil de charge cible *downlink* pour le contrôle d'admission ou la taille de la fenêtre d'ajout *Add_win* pour la macro-diversité.

algorithmes de paramétrage automatique respectivement mis en oeuvre pour le contrôle d'admission et la macro-diversité.

Pour le contrôle d'admission

Nous récapitulons ci-dessous l'ensemble des étapes mises en œuvre lors du fonctionnement du contrôleur de logique floue pour le paramétrage de l'algorithme de contrôle d'admission :

- Les entrées du contrôleur

Pour la station centrale, les taux de blocage et de coupure filtrés par moyenne glissante.
Pour les stations voisines, les taux de blocage et de coupure filtrés par moyenne glissante.
La valeur à l'itération précédente *(N-1)* du paramètre de seuil de charge cible *downlink*.

- *Fuzzification* des indicateurs de performances

Les taux de blocage et de coupure de la station centrale sont transformés en valeurs floues à l'aide des fonctions d'appartenance aux ensembles flous de blocage et de coupure (étape de *fuzzification*). Les taux de blocage et de coupure agrégés sur l'ensemble des voisines (par moyenne simple ou pondérée) sont transformés en valeurs floues à l'aide des fonctions d'appartenance aux ensembles flous de blocage et de coupure définies pour les voisines.

- Matrices de décisions de niveau 1

A partir des valeurs floues de blocage et de coupure de la station centrale, on détermine une valeur floue $\Delta c_{Centrale}$ de correction en utilisant la matrice de décisions pour la station centrale (comme si la station centrale était isolée). A partir des valeurs floues de blocage et de coupure agrégées sur l'ensemble des voisines, on détermine une valeur floue $\Delta c_{Voisines}$, correction s'appliquant à la centrale, décrivant l'influence de l'état des voisines sur le paramétrage de la station centrale.

- Matrice de décisions de niveau 2

La matrice de compromis centrale / voisines permet de faire la synthèse des valeurs floues de corrections $\Delta c_{Centrale}$ et $\Delta c_{Voisines}$ qui ont été respectivement déterminées en fonction de l'état de la station centrale et de l'état de ses voisines. Cette valeur $\Delta c_{Inst.}$ est une correction instantanée (ou sans mémoire), qui ne tient pas compte de la correction réalisée pour la station centrale à l'itération précédente.

- Matrice de décisions de niveau 3

La valeur floue de correction instantanée $\Delta c_{Inst.}$ est comparée à la valeur de correction de l'étape *N-1*. En fonction de l'efficacité de la correction à l'étape *N-1*, on décide de modérer ou de confirmer la valeur floue de correction $\Delta c_{Inst.}$. On obtient ainsi la valeur floue de correction finale Δc_{Finale}.

- *Defuzzification* de la valeur correction

A l'aide de l'opérateur de *defuzzification*, des degrés d'appartenance aux ensembles flous des variables intermédiaires et des valeurs numériques des labels flous de sortie, on transforme la valeur floue de correction Δc_{Finale} en une grandeur réelle de modification à appliquer au seuil de charge cible downlink.

L'ensemble des étapes décrites ci-dessus est résumé dans le schéma de fonctionnement (figure 3.20) suivant du contrôleur de logique floue.

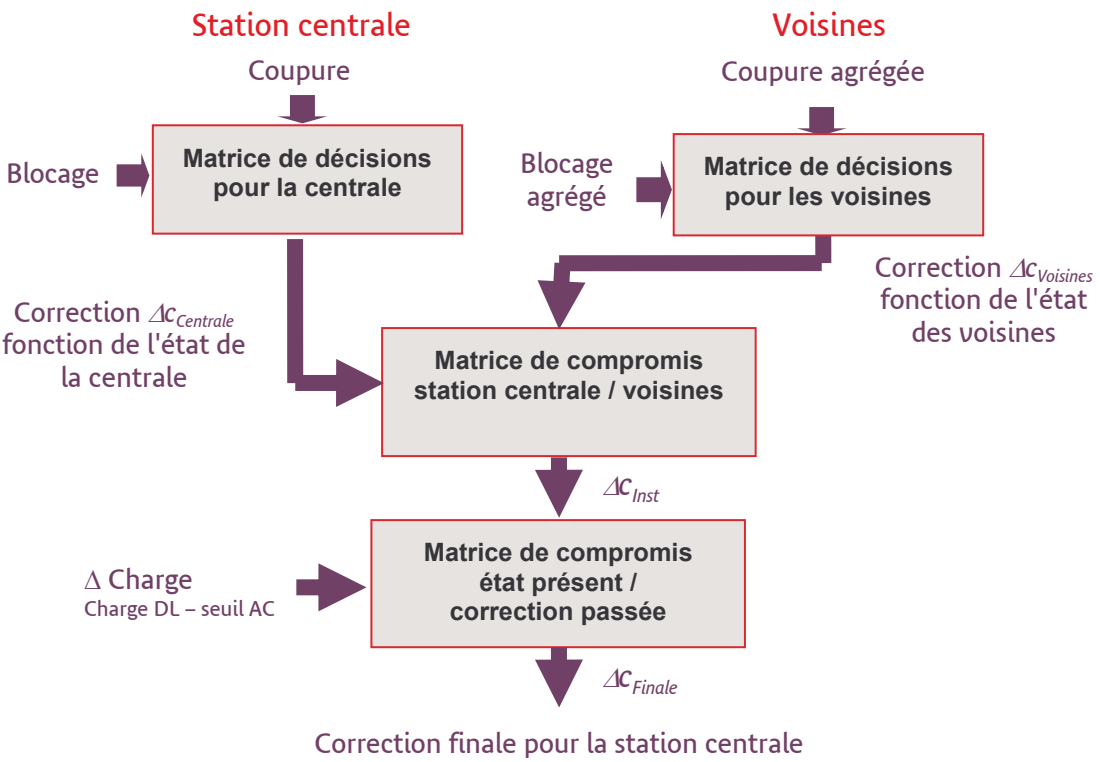


Figure 3.20 : Schéma de fonctionnement du contrôleur de logique floue pour le paramétrage du contrôle d'admission

Pour la macro-diversité

Nous récapitulons ci-dessous l'ensemble des étapes mises en œuvre lors du fonctionnement du contrôleur de logique floue pour le paramétrage de l'algorithme de macro-diversité :

- Les entrées du contrôleur

Pour la station centrale, le taux de blocage des liens créés par MD filtré par moyenne glissante.
Pour les stations voisines, le taux de blocage des liens de MD filtré par moyenne glissante.
La valeur à l'itération précédente *(N-1)* du paramètre de taille de fenêtre d'ajout *Add_win*.

- *Fuzzification* des indicateurs de performances

Le taux de blocage des liens de MD et la charge *downlink* de la station centrale sont transformés en valeurs floues à l'aide des fonctions d'appartenance aux ensembles flous de blocage des liens de MD et de charge DL. Les taux de blocage des liens de MD maximum et moyen sur l'ensemble des voisines sont transformés en valeur floue à l'aide des fonctions d'appartenance aux ensembles flous de blocage des liens de MD définies pour les valeurs maximum et moyenne sur l'ensemble des voisines. Le paramètre *Add_win(N-1)* (valeur de l'itération précédentes) est transformé en valeur floue à l'aide des fonctions d'appartenance aux ensembles flous définis pour le paramètre *Add_win*.

- Matrices de décisions de niveau 1

A partir des valeurs floues de taux de blocage des liens de MD pour la station centrale et du maximum des taux de blocage des liens de MD pour les voisines, on détermine une valeur floue Δc_{Locale} de correction en utilisant la matrice de décisions dans l'approche locale. A partir des valeurs floues de la moyenne des taux de blocage des liens de MD pour les voisines et de la charge *downlink* de la station centrale, on détermine une valeur floue $\Delta c_{Globale}$ de correction en utilisant la matrice de décisions dans l'approche globale.

- Matrice de décisions de niveau 2

La matrice de compromis local / global permet de faire la synthèse des valeurs floues de corrections Δc_{Locale} et $\Delta c_{Globale}$. On détermine ainsi une valeur floue de compromis local / global $\Delta c_{Compromis}$ du paramètre *Add_win*.

- Matrice de décision de niveau 3

La valeur floue de compromis $\Delta c_{Compromis}$ est comparée à la valeur de correction de *Add_win* à l'étape *N-1*. En fonction de l'efficacité de la correction à l'étape *N-1*, on décide de modérer ou de confirmer la valeur floue de correction $\Delta c_{Compromis}$. On obtient ainsi la valeur floue de correction finale Δc_{Finale}.

- *Defuzzification* de la valeur correction

A l'aide de l'opérateur de defuzzification, des degrés d'appartenance aux ensembles flous des variables intermédiaires et des valeurs numériques des labels flous de sortie, on transforme la valeur floue de correction Δc_{Finale} en une grandeur réelle de modification à appliquer à la taille de la fenêtre d'ajout *Add_win*.

L'ensemble des étapes décrites ci-dessus est résumé dans le schéma de fonctionnement (figure 3.21) suivant du contrôleur de logique floue.

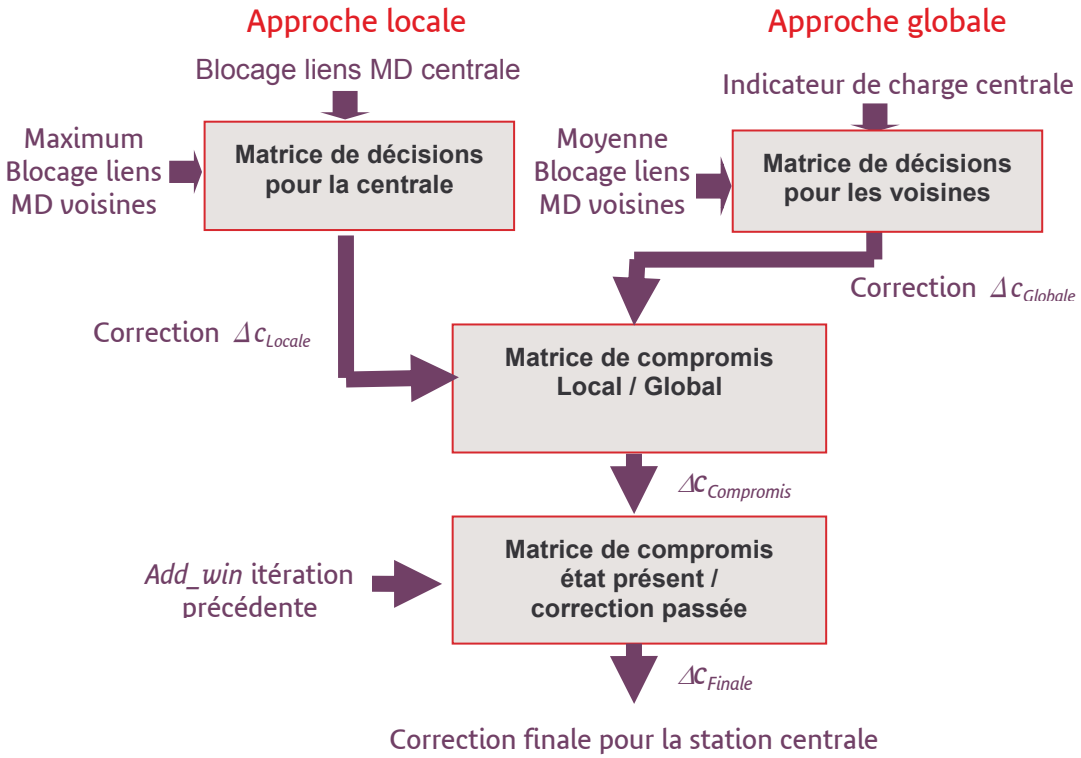


Figure 3.21 : Schéma de fonctionnement du contrôleur de logique floue pour le paramétrage de la macro-diversité

Fonctionnement conjoint des deux contrôleurs

Deux types de fonctionnement conjoint des contrôleurs de logique floue de l'algorithme de contrôle d'admission et de l'algorithme de macro-diversité :

- Fonctionnement alternatif

Tout d'abord, on effectue une adaptation automatique du paramétrage du contrôle d'admission en mode *off-line* (voir chapitre 5). On analyse les performances obtenues en termes

d'indicateurs et de fonctions coûts (voir chapitre 4). Puis on effectue une adaptation automatique du paramétrage de la macro-diversité en mode *off-line*, et on analyse les performances obtenues en termes d'indicateurs et de fonctions coûts. On compare alors les performances et les fonctions coûts obtenues après applications successives des deux contrôleurs en mode *off-line*.

- Fonctionnement simultané

En mode *on-line*, on effectue l'adaptation automatique simultanée du paramétrage du contrôle d'admission et du paramétrage de la macro-diversité. On comparera alors les performances et les fonctions coûts obtenues après cette adaptation conjointe avec les performances et fonctions coûts obtenues par l'adaptation automatique de l'un ou l'autre des algorithmes RRM considérés.

3.3.3. Procédure de defuzzification

Le module de *defuzzification* permet de transformer la valeur floue de correction finale en une valeur réelle de modification du paramètre RRM à modifier. Pour construire notre *module de defuzzification* du processus FKBC, nous devons choisir les opérateurs de *defuzzification* ainsi que les valeurs numériques correspondant aux labels de sortie.

3.3.3.1 Valeurs de correction en sorties

Tout d'abord, il s'agit de définir les valeurs numériques correspondant aux différents labels de sortie. L'expérience acquise sur les réseaux expérimentaux et les simulateurs de réseau mobiles UMTS nous a permis de définir des valeurs typiques a priori de correction des paramètres RRM considérés. Ces valeurs typiques de corrections ont ensuite été testées avec des simulations logicielles de paramétrage automatique de réseaux mobiles UMTS.

Paramètres RRM	Correction faible	Correction forte
Seuil de charge cible DL	< 0.05	≥ 0.05
Taille de la fenêtre d'ajout de macro-diversité (dB)	< 0.02	≥ 0.02

Table 3.8 : Labellisation des corrections des paramètres RRM

L'optimisation de cette typologie des valeurs de correction des paramètres RRM peut aussi être introduite dans les processus d'optimisation du SIF. En entrées du processus d'optimisation, on définira les ensembles de variations de chaque label de correction de chaque paramètre RRM. L'optimisation par essaim de particules (chapitre 4) permet de tester différents jeux de valeurs pour ces labels (exploration) et d'en déduire des combinaisons efficaces (meilleures particules). Dans le cas de l'apprentissage par renforcement, on va discrétiser les ensembles de variations des sorties du SIF, et l'algorithme de *Q-learning* va permettre d'apprendre les combinaisons de valeurs maximisant les fonctions de qualité du processus (chapitre 4). Ces deux méthodes proposent ainsi deux procédés d'exploration qui pourront compléter l'expérience des ingénieurs en charge du paramétrage radio.

3.3.3.2 Types d'opérateurs de defuzzification

Les opérateurs de *defuzzification* permettent de déterminer la valeur déterministe de sortie Δc_{Sortie} de notre processus à partir de la valeur floue finale. La valeur floue de correction finale

Δc_{Finale} a été obtenue par agrégation de plusieurs valeurs floues intermédiaires: les labels flous de sortie sont notés $O_1, ..., O_N$, les valeurs numériques de ces labels $o_1, ..., o_N$ et les degrés d'appartenance correspondants $\alpha_1, ..., \alpha_N$. Deux opérateurs de *defuzzification* sont utilisables dans notre processus de contrôle :

- Opérateur du maximum de *defuzzification* : la sortie est égale à la valeur numérique du label correspondant au degré d'appartenance le plus fort (ou la moyenne des valeurs numériques si plusieurs labels ont le même degré d'appartenance) :

$$\Delta c_{Sortie} = o_{i_{Max}} \text{ où } \underset{1 \le i \le N}{i_{Max}} = \arg\max(\alpha_i) \tag{3.13}$$

- Opérateur du centre de gravité de *defuzzification* : la sortie est égale à la somme pondérée (par les degrés d'appartenance) des valeurs numériques des labels correspondants :

$$\Delta c_{Sortie} = \sum_{1 \le i \le N} o_i \times \alpha_i \tag{3.14}$$

Dans les CLF, nous utiliserons plutôt l'opérateur du centre de gravité de *defuzzification*, car cet opérateur tient compte des valeurs numériques de l'ensemble des labels activés au cours du processus.

3.4. Conclusions

Nous avons commencé par décrire le problème général du contrôle des paramètres RRM d'un réseau UMTS à l'aide d'un système d'inférence floue (SIF). Cette modélisation pourra s'appliquer aussi au contrôle d'autres réseaux mobiles (réseaux de type GSM/GPRS/EDGE et CDMA) voire être utilisée dans une approche multi-système. Nous rappelons qu'un SIF est composé d'une base de règles de décisions qui peuvent être déduites, dans le cas des réseaux UMTS, de l'expérience des ingénieurs radio et de l'analyse de l'impact du paramétrage des algorithmes RRM.

Dans le cadre de ce SIF, nous avons modélisé le contrôleur automatique par une structure de type FKBC. Les composants d'un processus FKBC comprennent les éléments suivants. Le *Module de fuzzification* transforme les variables d'entrées continues en des labels linguistiques cohérents avec les prémisses des règles d'inférence. La *Base de connaissance* se décompose en la *base de données* et la *base de règles*. Le *Moteur d'inférence* permet de déterminer la valeur floue globale de sortie à partir des contributions individuelles de chaque règle élémentaires. Le *Module de defuzzification* transforme la valeur floue de corrections en une valeur déterministe (continue ou discrète). Cette modélisation des contrôleurs nous a permis de déduire un schéma général de fonctionnement représenté sur la figure 3.22 (pertinent a priori pour toute technologie de réseau d'accès radio mobile), déduit des figures 3.20 et 3.21 et que l'on pourra tester pour le contrôle d'autres paramètres RRM :

- Deux types de données en entrées du contrôleur : indicateurs de performances locaux et globaux.
- Mémorisation par le contrôleur de l'état et/ou des corrections de l'itération précédente.
- Application par le contrôleur des règles élémentaires de contrôle locales et globales.
- Compromis entre les données locales, globales et passées en appliquant les opérateurs d'agrégation des valeurs floues.
- Synthèse des corrections dans la valeur floue de correction finale Δc_{Finale}.

- Au lieu de corriger le paramétrage de la station centrale en fonction de l'état des stations voisines (hypothèse sous-jacente aux figures 3.20 et 3.21), le contrôleur automatique pourrait être adapté pour modifier directement le paramétrage des voisines, si une telle méthode s'avérait plus efficace.

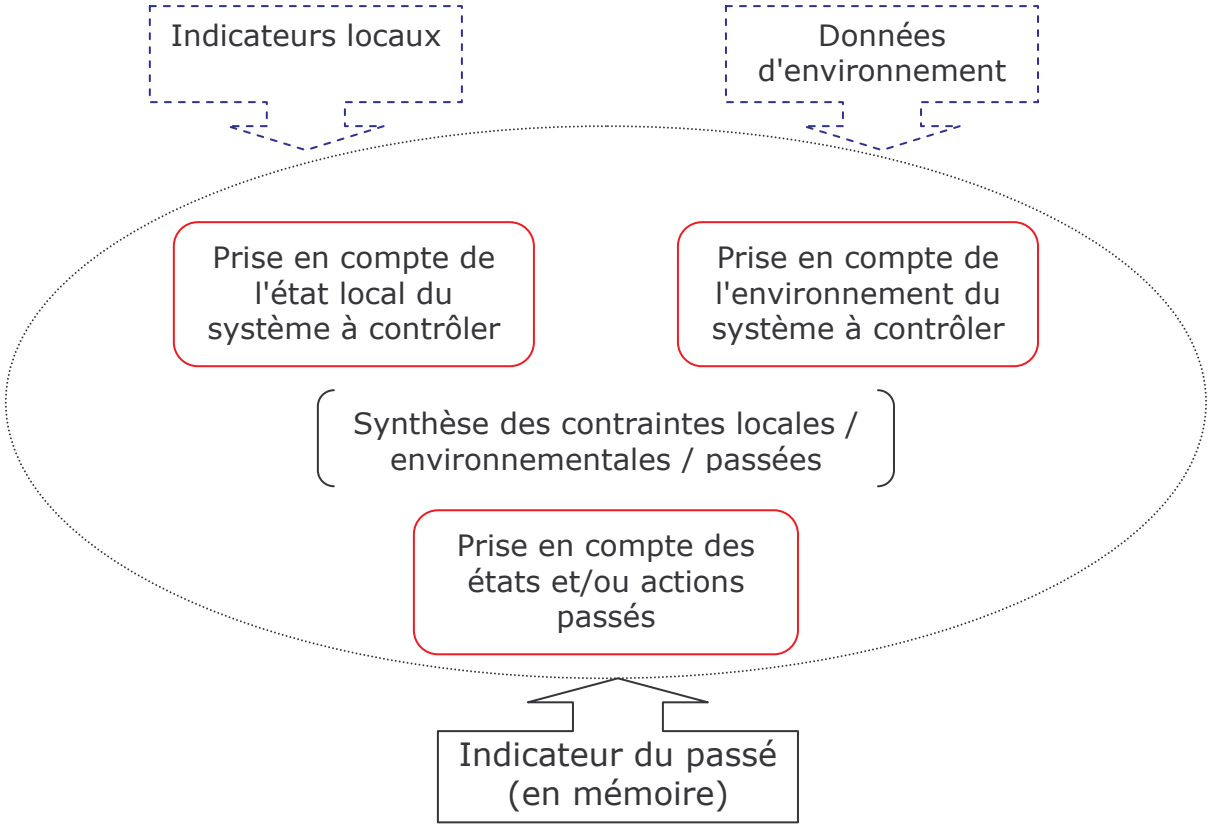


Figure 3.22 : Schéma générique de fonctionnement d'un contrôleur FKBC

La représentation paramétrique du moteur d'inférence et de la base de données du contrôleur FKBC sera très pratique dans le cadre de l'optimisation par essaim de particules, décrite au chapitre 4. En revanche, dans le cas de l'optimisation par apprentissage par renforcement, on préférera revenir à une description du problème de paramétrage automatique par un système d'inférence floue pour avoir une structure de contrôle plus générale. En effet, l'apprentissage par renforcement permet un mode d'apprentissage plus riche (prise en compte parallèle des données de chaque cellule, au lieu de l'utilisation de fonctions coûts globales dans l'OEP) et utilise des notions plus facilement transposables aux réseaux réels.

Enfin, nous avons décrit deux aspects essentiels de l'introduction des contrôleurs automatiques dans les réseaux mobiles UMTS. Le premier point concerne les différentes échelles de temps relatives à un tel processus de contrôle : granularité des indicateurs de performances, taille des fenêtres de filtrage, intervalles entre les modifications des paramètres RRM, temps de réponse du réseau à une modification du paramétrage… Le second point fondamental dans l'optique d'une mise en oeuvre opérationnelle est la définition d'une architecture permettant le contrôle automatique des paramètres RRM. Ainsi, nous avons introduit deux types d'architectures de paramétrage automatique des réseaux mobiles: un modèle centralisé et un modèle par agents distribués.

4. Méthodes d'optimisation des contrôleurs de logique floue

Les objectifs de l'optimisation des CLF sont d'améliorer de manière automatique les performances de ces contrôleurs et de les adapter localement à un environnement réel et une demande en trafic caractéristique de la zone considérée. Cette optimisation vise donc à obtenir un paramétrage adapté et plus efficace des contrôleurs de logique floue.

Deux modes de fonctionnement de l'optimisation peuvent être envisagés :

- L'optimisation dynamique des contrôleurs CLF comme décrit sur la figure 4.1 : un tel fonctionnement semble cependant difficilement utilisable dans les réseaux mobiles à court terme, à cause du risque de faire rentrer le réseau dans un état chaotique et la nécessité conséquente de valider toute modification du contrôleur.
- Une adaptation *off-line* des contrôleurs en utilisant des métriques et KPI radio collectés, des outils de simulations de gestion dynamique des réseaux mobiles et des tests de validation de ces contrôleurs optimisés en simulation sur des réseaux expérimentaux.

Les connaissances fondées sur l'analyse et l'expertise des ingénieurs radio seront intégrées au cours du processus d'optimisation sous forme de connaissance a priori (notamment dans le processus d'apprentissage par renforcement). Réciproquement, les contrôleurs optimisés pourront fournir en fin de processus des règles de contrôle nouvelles qui n'avaient pu être découvertes expérimentalement auparavant.

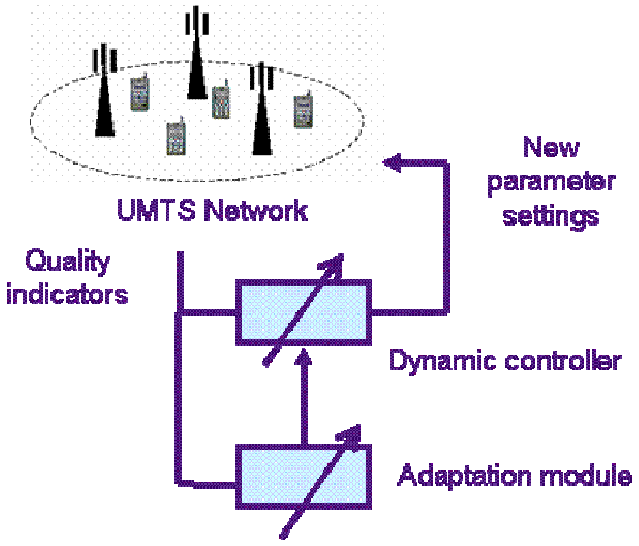


Figure 4.1 : Adaptation dynamique des paramètres des contrôleurs de logique floue

Deux stratégies d'optimisation fondées sur des approches différentes sont proposées dans ce chapitre :

- Une stratégie utilisant l'optimisation par essaim de particules qui appartient à la famille des méthodes d'optimisation itératives stochastiques. Dans ce cas, la notion d'optimisation plus efficace se traduira mathématiquement au cours du processus par l'amélioration de fonctions coûts qui seront définies au paragraphe 4.1.3.2 dans le but de comparer les performances des contrôleurs de logique floue.

- Une stratégie d'optimisation par apprentissage par renforcement (*Reinforcement learning*) qui introduit des récompenses et punitions (valeurs de renforcement) en fonction des performances des contrôleurs en terme de qualité de service mesurée dans le réseau. Nous adapterons l'algorithme de *Q-learning* au système d'inférence floue introduit au chapitre 3. Cet algorithme se décompose en deux phases principales : une phase d'exploration et une phase d'exploitation. Durant la phase d'exploration, l'algorithme va prendre des décisions (de corrections des paramètres RRM) partiellement aléatoires qui vont ensuite, à l'aide d'une boucle de rétroaction et des valeurs de renforcements, lui permettre de distinguer les corrections efficaces des corrections inutiles en fonction de l'état de la cellule et ses voisines. Une fois que l'on considère que l'algorithme a suffisamment visité les différents états du système, on passe à la phase d'exploitation. Dés lors, à chaque itération, l'algorithme va sélectionner systématiquement la "meilleure" correction à appliquer au paramétrage du node B. Avec l'algorithme d'OEP, les indicateurs des différentes cellules sont agrégés dans les fonctions coûts en n'exploitant pas une partie importante de l'information. Au contraire, l'algorithme de *Q-learning* va tenir compte, à chaque itération, de l'état, du renforcement reçu et de la correction appliquée au niveau de chaque cellule de la zone considérée. Le système d'inférence floue va donc apprendre simultanément sur l'ensemble des nodes B.

Dans la première partie du chapitre, nous développons une méthode d'optimisation des contrôleurs FKBC à partir d'une version simplifiée de l'OEP. Dans la seconde partie nous adaptons l'algorithme de *Q-learning* au problème d'optimisation par apprentissage du SIF défini pour le contrôle automatique d'un paramètre RRM dans les réseaux UMTS.

4.1. Optimisation par essaim de particules

L'optimisation par essaim de particules appartient à la famille de méthodes d'optimisation itératives stochastiques. Un avantage essentiel est que ces méthodes peuvent s'appliquer aussi bien aux problèmes à variables discrètes qu'aux problèmes à variables continues. Dans notre cas, nous avons voulu appliquer la méthode d'OEP au problème de l'optimisation des contrôleurs de logique floue qui ont été introduits au paragraphe 3.3 et sont entièrement décrits par :

- Un système de matrices de décisions paramétrées chacune par leurs quatre extrémités pour *le moteur d'inférence* (variables discrètes).
- Les limites des partitions floues introduites pour chaque indicateur de performance pour *la base de données* (variables continues).

Dans le laboratoire NET de France Télécom Recherche & Développement, une méthode d'optimisation des contrôleurs par essaim de particules a été proposée en collaboration par M. Clerc, Z. Altman et H. Dubreil (réf. [Dub05]).

4.1.1. Introduction à l'optimisation par essaim de particules

L'OEP est une technique d'optimisation fondée sur un ensemble de taille fixe d'agents, appelés particules, se déplaçant dans l'espace de recherche à D dimensions. La position de chaque particule correspond à un paramétrage du contrôleur de logique floue. La vitesse de chaque particule correspond aux modifications de ce paramétrage entre deux simulations de réseau mobile. On initialise aléatoirement la position et la vitesse de chaque particule. L'ensemble des particules se déplace dans l'espace de recherche, et on évalue régulièrement les performances de chaque particule à l'aide de fonctions coûts. L'évolution de chaque

particule permet une exploration "intelligente" de l'espace de recherche. L'ensemble des particules forme l'essaim, de taille *S* fixée au départ.

Principes d'évolution de l'essaim de particules

A la première itération, les positions et vitesses de chaque particule P_i sont initialisées aléatoirement (sauf la position de la particule P_0 qui correspond au paramétrage par défaut du contrôleur). Puis à chaque itération du processus, trois éléments essentiels sont pris en compte pour la détermination de la nouvelle position de la particule P_i :

- La vitesse propre de la particule P_i (calculée à l'itération précédente).
- La meilleure position dans le passé de la particule P_i, cette influence étant interprétée comme une attraction exercée par la meilleure position mémorisée de la particule.
- Pour la particule P_i, on tire aléatoirement sur l'ensemble des particules une liste de particules informatrices. On va tenir compte de la meilleure performance de la meilleure particule informatrice pour le calcul de la nouvelle position de P_i. Cette influence peut être interprétée comme une attraction exercée par la meilleure des particules informatrices.

On combine linéairement ces trois effets à l'aide de coefficients de confiance aléatoires, comme décrit sur la figure 4.2 suivante :

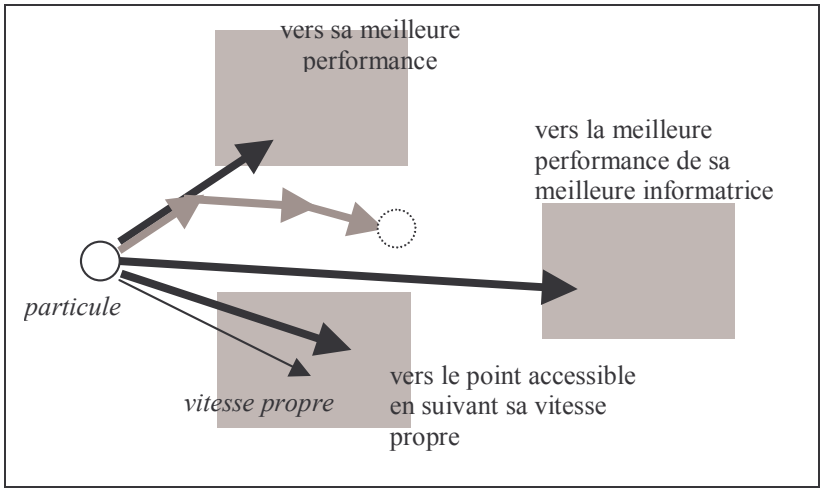


Figure 4.2 : Détermination de la nouvelle position d'une particule dans un processus OEP (Les trois flèches grisées représentent la combinaison vectorielle des trois effets pris en compte)

Dans toutes les versions d'OEP « classiques », on combine ainsi linéairement les vecteurs des trois effets pris en compte via des coefficients c_1 et c_2 de confiance aléatoires. Dans notre cas, nous considérons des distributions de probabilité simplifiées pour les coefficients de confiance : c_1 est constant, alors que c_2 est une variable aléatoire de type D-rectangulaire. Nous avons donc les relations suivantes pour la mise à jour de la vitesse $\underline{v}$ et de la position $\underline{x}$ d'une particule :

$$\begin{cases} v_d \leftarrow c_1 \times v_d + c_{2,d}\left(p_d - x_d\right) + c_{2,d}\left(g_d - x_d\right) \\ x_d \leftarrow x_d + v_d \end{cases}$$

Où d est une dimension dans $\{1,\ldots, D\}$, $\underline{p}$ un point tiré aléatoirement dans le voisinage de la meilleure position de la particule dans le passé et $\underline{g}$ un point tiré aléatoirement dans le voisinage de la meilleure position dans le passé de la meilleure particule informatrice.

Avantages de l'OEP

L'optimisation par essaim de particules est une méthode d'optimisation itérative stochastique qui s'applique aussi bien aux problèmes à variables continues (limites des ensembles flous) qu'aux problèmes à variables discrètes (valeurs entières des sommets des matrices de décisions dans notre cas), contrairement à d'autres méthodes d'optimisation. De plus, cette méthode permet en général de converger rapidement (réf. [Cle02]) vers une solution approchée de bonne qualité. Il s'agit là d'un point important car l'intervalle de temps entre deux modifications de paramétrage dépend de la durée de simulation du réseau mobile considéré (fonction du nombre de stations actives). Et plus l'évaluation des fonctions coûts dure longtemps, plus il est nécessaire d'avoir une méthode d'optimisation qui converge rapidement.

Il existe deux familles d'algorithmes OEP : les versions paramétriques et adaptatives. La version paramétrique, utilisée la suite de ce document, nécessite de configurer certains paramètres comme la taille de l'essaim, les coefficients de confiance… C'est une méthode d'optimisation très largement répandue dont le fonctionnement est relativement simple et qui peut être implémentée très facilement. La version adaptative évite à l'utilisateur d'avoir à fixer les paramètres de l'algorithme comme la taille de l'essaim, les coefficients de confiance c_1 et c_2 ou le nombre de particules informatrices. A l'initialisation de l'algorithme, il est seulement nécessaire de correctement décrire le problème à optimiser: c'est-à-dire les ensembles de variations des CLF, les contraintes du problème, les fonctions coût que l'on veut minimiser

Avec les OEP, deux types de fonctionnement de l'algorithme sont envisageables sans que cela nécessite des adaptations importantes: mono-objectif ou multi-objectif. Pour le mono-objectif, on introduit un coefficient de mixage (fixe au cours de l'optimisation) qui permet de pondérer les coûts relatifs du blocage et de la coupure d'un appel dans le réseau mobile. En multi-objectif, on pourra classer les performances des différents contrôleurs dans le plan coupure – blocage : plusieurs solutions de CLF optimaux pourront être disponibles en fin d'optimisation, chacune appartenant au meilleur front déterminé par OEP et fournissant un compromis différent en termes de coupure et de blocage.

Enfin l'OEP est une méthode qui permet de prendre en compte les contraintes suivantes intrinsèques à notre formulation du problème d'optimisation des CLF :

- Les paramètres à optimiser appartiennent à des intervalles de variations fixés en début de processus (sommets de matrices de décision, limites des ensembles flous).
- Les paramètres doivent respecter différentes contraintes de granularité: valeurs entières pour les coefficients des matrices de décisions, ou valeurs continues pour les limites des ensembles flous.
- Certains paramètres doivent respecter des contraintes de relations d'ordre. Ainsi, dans le système de matrices de décisions (voir chapitre 3) implémenté dans les CLF, les règles de gestion de réseau fondée sur l'analyse et l'expérience des ingénieurs radio ont conduit à introduire des relations d'ordre entre les quatre sommets de chacune des matrices de décisions (dans la représentation paramétrique par les sommets des matrices). Par exemple au paragraphe 3.3.2.2, nous avons expliqué comment tenir compte des règles d'inférences introduites au 3.3.2.1 dans la définition de la matrice de décision pour la station centrale. La conséquence principale est d'introduire des contraintes de relations d'ordre entre les éléments de la matrice $\left[m_{ij}^{Centrale}\right]$ de taille $n_{Bloc.} \times n_{Coup.}$ ($n_{Bloc.}$ étant le nombre de lignes et $n_{Coup.}$ le nombre de colonnes) :

$$\forall i \in \{1,...,n_{Bloc.}\},\ \forall j \in \{1,...,n_{Coup.}\},\ m_{1n_{\mathrm{Re}\,jet}}^{Centrale} \le m_{ij}^{Centrale} \le m_{n_{Bloc.}1}^{Centrale} \qquad (4.1)$$

Critères ou conditions de convergence

Les critères de convergence énoncés dans [Cle02] permettent de calculer les coefficients de confiance c_1 et $c_{2,max}$ en tant que coefficients de constriction à partir d'un unique paramètre φ. Dans ce mémoire, nous utilisons les coefficients de constriction introduits pour la généralisation du système d'essaim de particules dans [Cle02] au cas où deux termes vectoriels sont ajoutés à la vitesse et où les coefficients φ_1 et φ_2 sont des variables uniformes aléatoires dans $[0, \varphi_{1,max}]$ et $[0, \varphi_{2,max}]$. Pour l'optimisation par essaim de particules des CLF nous considérons le système simplifié d'essaim de particules suivant :

$$\varphi_{1,\max} = \varphi_{2,\max} = \frac{\varphi}{2} \tag{4.2}$$

$$\chi = \frac{\kappa}{\left(\varphi - 2 + \frac{\sqrt{\varphi(\varphi-4)}}{2}\right)} \tag{4.3}$$

$$v(t+1)) = \chi\left(v(t) + \varphi_1\left(p_1 - x(t)\right) + \varphi_2\left(p_2 - x(t)\right)\right) \tag{4.4}$$

$$x(t+1) = v(t+1)) + x(t) \tag{4.5}$$

Le paramètre κ permet de contrôler le degré de convergence de l'algorithme. Le but est de choisir un paramètre κ empêchant d'une part l'explosion du système de particules tout en modérant d'autre part la constriction du système pour permettre une exploration suffisante de l'espace d'évolution des particules. En prenant $\kappa = 1$ et $\varphi = 4.14$, cela produit une convergence lente du système, ce qui signifie que l'espace de définition des particules est largement exploré avant que la population ne converge vers un optimum local.

4.1.2. Description de l'algorithme d'optimisation par essaim de particules

Dans ce paragraphe, nous allons décrire la version de l'OEP adaptée au problème d'optimisation des CLF. Nous rappelons que l'OEP est une méthode d'optimisation fondée sur l'évolution d'un essaim de particules (de taille fixe) dans l'espace de recherche (de dimension D), avec évaluation régulière des performances de chaque particule à l'aide de fonctions coûts. Pour chaque particule, nous définissons donc quatre attributs qui seront mis à jour avant chaque simulation de réseau :

- La position de la particule dans l'espace de recherche $\underline{x} = (x_1,...,x_D)$, correspondant au paramétrage du contrôleur CLF (matrices de décisions et ensembles flous).
- La vitesse de la particule $\underline{v} = (v_1,...,v_D)$, correspondant aux modifications de la position (et donc du paramétrage du contrôleur CLF) entre deux simulations du réseau mobile.
- La meilleure position trouvée pour la particule considérée dans le passé $\underline{p} = (p_1,...,p_D)$.
- La liste des particules informatrices (ou groupe d'information) pour la particule considérée qui seront tirées aléatoirement avant chaque simulation.

Initialisation de l'algorithme

(1) Choix de la taille de l'essaim S et du nombre de particules informatrices K.
(2) Choix du paramètre φ permettant de calculer les coefficients de confiance c_1 et $c_{2,max}$.
(3) Détermination de la position de la particule 0 correspondant au paramétrage par défaut fourni en entrée de l'algorithme.

(4) $\forall i \in \{1,..., S-1\}$, tirage aléatoire de la position initiale de la i-ème particule. Vérification du respect des contraintes sur la position de chaque particule : contraintes d'intervalles de définition, de granularité, et d'encadrements pour toutes les dimensions x_d de la position.

(5) $\forall i \in \{0,..., S-1\}$, tirage aléatoire de la vitesse initiale de la i-ème particule.

(6) Evaluation des performances de la position initiale de chaque particule à l'aide des fonctions coûts.

Itération n de l'algorithme

(7) Tirage aléatoire des *K* liens d'informations pour chaque particule.

(8) Pour chaque particule, détermination de la nouvelle position x^n.

a. Tirage aléatoire d'une vitesse v' autour de v^{n-1} avec une densité de probabilité $R_1\left(x^{n-1}, v^{n-1}\right)$: dans notre implémentation simplifiée on a pris $v'_d = c_1 \times v_d^{n-1}$.

b. Tirage aléatoire d'une position p' autour de p^{n-1} avec une densité de probabilité $R_2\left(x^{n-1}, p^{n-1}\right)$: dans notre implémentation simplifiée, on considère une variable aléatoire c_2 de type D-rectangulaire : $p'_d = c_{2,\max} \times \delta_d \times p_d^{n-1}$ ($c_{2,max}$ étant un paramètre fixé par l'utilisateur, cf. éq. (4.7), et δ_d une variable aléatoire uniforme dans *[0, 1]*).

c. Détermination de la meilleure position dans le passé g^n de la meilleure particule informatrice.

d. Tirage aléatoire d'une position g′ autour de g^n avec une densité de probabilité $R_2\left(x^{n-1}, p^{n-1}\right)$: dans notre implémentation simplifiée, on considère une variable aléatoire c_2 de type D-rectangulaire : $g'_d = c_{2,\max} \times \chi_d \times g_d^{n-1}$ (χ_d étant une variable aléatoire uniforme dans *[0, 1]*).

e. Détermination du déplacement global de la particule v^n :

$$v^n = v' + (p' - x) + (g' - x).$$

f. Détermination de la nouvelle position de la particule x^n :

$$x^n = x^{n-1} + v^n.$$

(9) Pour chaque particule, vérification pour la nouvelle position x^n des contraintes d'intervalles, de granularité et de relations d'ordre avec mises à jour éventuelles de la position x^n et de la vitesse v^n pour respecter ces contraintes.

(10) Evaluation des performances de la nouvelle position x^n de chaque particule à l'aide des fonctions coûts : $Coût\left(x^n\right)$.

(11) Pour chaque particule, si nécessaire, mise à jour de la meilleure position :

$$Si\ Coût\left(x^n\right) < Coût\left(p^{n-1}\right)\ Alors\ p^n \leftarrow x^n,$$
$$Sinon\ p^n \leftarrow p^{n-1}.$$

(12) Pour l'ensemble des particules, si la position x_{Best}^n de la meilleure particule à l'itération *n* est meilleure que la meilleure particule stockée sur l'ensemble des itérations précédentes, alors stocker x_{Best}^n.

Critères d'arrêt de l'algorithme

Plusieurs critères d'arrêt sont possibles pour l'algorithme d'OEP :

- Un seuil sur les fonctions coûts de la meilleure particule mémorisée.
- Une limite en termes de nombre d'itérations ou de temps de calcul.
- Plus d'amélioration de la meilleure particule pendant un nombre minimal d'itérations fixé a priori.

Ces trois critères ne sont pas exclusifs et peuvent donc être utilisés simultanément.

Paramètres de l'algorithme d'OEP

Nous récapitulons ci-dessous les différents paramètres à définir en entrées de l'optimisation par essaim de particules :

- Espace de recherche : une représentation paramétrique des contrôleurs CLF a été choisie (pour le système des matrices de décision et ensembles flous). L'espace de recherche correspond à l'ensemble de variations des paramètres définissant les contrôleurs CLF: cela pourrait être par exemple un produit d'ensembles discrets et continus de dimension typique de l'ordre de 20.
- Taille de l'essaim : le nombre total de particules de l'essaim est de l'ordre 20 - 30.
- Coefficients de confiance : il a été démontré [Cle02] qu'on pouvait exprimer les coefficients de confiance c_1 et $c_{2,max}$ à partir d'un coefficient unique φ (4.1 étant par exemple une valeur typique efficace) :

$$c_1 = \frac{2}{\left(\varphi - 2 + \sqrt{\varphi^2 - 4\varphi}\right)} \tag{4.6}$$

$$c_{2,\max} = \frac{c_1 \times \varphi}{2} \tag{4.7}$$

  En calculant de cette manière les coefficients de confiance, l'algorithme d'OEP respecte des relations qui favorisent la convergence du processus (cf. paragraphe 4.1.1).

- Nombre de particules informatrices K : à chaque itération, on tire aléatoirement pour chaque particule un nombre fixe de particules informatrices. La meilleure position de la meilleure particule informatrice impactera la détermination de la nouvelle position de la particule considérée. En augmentant ce paramètre K, on accélère a priori la convergence de l'algorithme, le risque étant de converger vers des minima locaux. En diminuant ce paramètre K, on privilégie l'exploration de l'espace de recherche par l'algorithme. Une valeur typique du paramètre K est 3.

Fonctions de comparaison des particules informatrices au cours de l'OEP

En monobjectif afin de comparer l'efficacité de différentes particules informatrices, nous supposons que nous avons déterminé une fonction coût M (voir paragraphe 4.1.3.2) basée sur les indicateurs de performance de blocage et coupure (ou d'autres si nécessaire). Nous considérons deux types de comparaison dans le cadre d'une OEP :

- Comparaison directe : la particule informatrice I_1 est meilleure que la particule informatrice I_2 pour la particule x si $M(I_1) < M(I_2)$.

- Comparaison par pseudo-gradients : la particule informatrice I_1 est meilleure que la particule informatrice I_2 pour la particule x si $\frac{M(x)-M(I_1)}{\|I_1-x\|} > \frac{M(x)-M(I_2)}{\|I_2-x\|}$, où || || est une métrique sur l'espace de recherche. Cette fonction de comparaison a été introduite dans un processus d'OEP dans l'article [Vee03]. Cette fonction permet de favoriser les particules informatrices ayant de bonnes performances présentes dans le voisinage de la particule x, et aussi de regrouper les particules en sous-essaims.

En multi-objectif, nous supposons à présent que nous avons déterminé deux fonctions coût $M_{Blocage}$ et $M_{Coupure}$ (voir paragraphe 4.1.3.2) basées respectivement sur les indicateurs de blocage et de coupure. Pour comparer les performances des particules informatrices dans le plan taux de coupure – taux de blocage pour chaque fonction coût M_X (X = *blocage* ou *coupure*), nous pouvons :

- Soit comparer les fonctions coûts des particules informatrices : $M_X(I_1) \overset{?}{<} M_X(I_2)$.
- Soit comparer les pseudo-gradients : $\frac{M_X(x)-M_X(I_1)}{\|I_1-x\|} \overset{?}{>} \frac{M_X(x)-M_X(I_2)}{\|I_2-x\|}$.

4.1.3. Application de l'optimisation par essaim de particules aux CLF

Dans ce paragraphe, nous détaillons l'implémentation de l'optimisation par essaim de particules appliquée au cas des contrôleurs CLF des algorithmes RRM. La figure ci-dessous décrit l'ensemble des modules impliqués dans l'optimisation par essaim de particules des contrôleurs CLF.

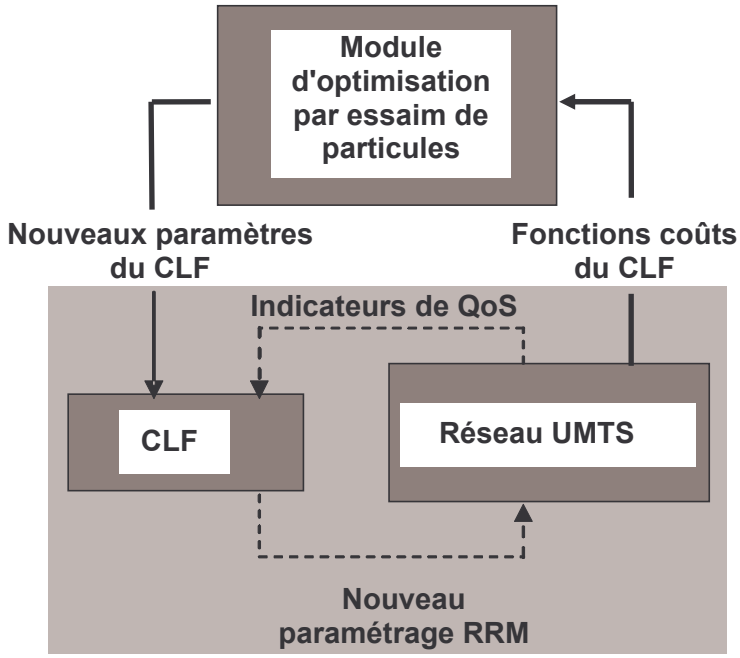


Figure 4.3 : Fonctionnement global de l'optimisation par essaim de particules des FLC

Nous expliciterons ci-après l'ensemble des paramètres et contraintes de l'OEP, ainsi que les expressions des informations échangées entre les différents modules (figure 4.3), dans le cas d'une optimisation de contrôleurs pour les réseaux mobiles UMTS.

4.1.3.1 Représentation paramétriques des contrôleurs et expression des contraintes

La position de chaque particule dans l'espace de recherche correspond à un paramétrage particulier du contrôleur de logique floue, c'est-à-dire le paramétrage des ensembles flous pour les indicateurs de performance et les corrections en sortie ainsi que le paramétrage des matrices de décisions. Nous choisissons comme représentation paramétrique des contrôleurs de logique floue :

- Des ensembles flous de type *linear gradual scaling* (voir paragraphe 3.3.1.2).
- Une représentation paramétrique par les quatre sommets (avec interpolation linéaire) pour les matrices de décisions. Les éléments des matrices de décisions correspondent à des pointeurs vers les différentes corrections possibles: ils sont donc définis dans notre représentation comme des entiers relatifs.

Ainsi, pour une optimisation du contrôleur gérant le seuil de charge cible DL (voir paragraphe 4.3.2.3), il y aura :

- 3 paramètres par indicateur de performance dont on optimise les ensembles d'appartenance : x_{min}, x_{max}, β. Les ensembles de variations de ces paramètres sont de types intervalles continus.
- 16 paramètres correspondant aux sommets des matrices de décisions dont les espaces de variations respectifs seront discrets et finis, de type *{-2, -1, 0, +1, +2}*.
- 2 paramètres pour les valeurs numériques des corrections en sortie des contrôleurs. Les ensembles de variations de ces paramètres sont de types intervalles continus: *[0.07, 0.15]* pour une correction forte, et *[0.01, 0.05]* pour une correction faible.

Au final, l'espace de recherche pour une optimisation par OEP du contrôleur gérant le seuil de charge cible DL sera dans ce cas: $\{-2, -1, 0, +1, +2\}^{16} \times [0.07, 0.15] \times [0.01, 0.05]$ et $[0, 0.3] \times [0.3, 1] \times [0, 15]$ pour les ensembles d'appartenance des indicateurs de performance.

Prise en compte des contraintes

Lors de l'évolution de l'essaim de particules dans l'espace de recherche, nous contrôlons à chaque itération que toutes les particules vérifient les contraintes intrinsèques à notre formulation du problème d'optimisation des CLF. En particulier, lors de la détermination de la nouvelle position de la particule (étape 8 de l'algorithme), il est examiné si :

- Les paramètres de l'optimisation ont bien une valeur appartenant à leurs intervalles de variation. Dans le cas contraire, le paramètre est alors pris égal à l'extrémité la plus proche. La composante correspondante de la vitesse est soit annulée soit prise égale à son opposée : Si $x_d > x_d^{\max}$ Alors $x_d = x_d^{\max}$ et $\left(v_d \leftarrow 0 \text{ ou } v_d \leftarrow -v_d\right)$.
- Les paramètres respectent leurs contraintes de granularité (valeurs discrètes). Dans le cas contraire, on ramène le paramètre en question à la valeur discrète la plus proche. On corrige la valeur de la composante correspondante de la vitesse pour respecter l'équation de déplacement : $x^n = x^{n-1} + v^n$.
- Certains paramètres vérifient entre eux des contraintes de relations d'ordre (comme celle décrite au paragraphe 4.1.1) : $m_{1n_{Coup.}}^{Centrale} \leq m_{ij}^{Centrale} \leq m_{n_{Bloc.}1}^{Centrale}$.

  Dans cet exemple, le sommet $m_{1n_{Coup.}}^{Centrale}$ doit nécessairement avoir la plus petite valeur de la matrice. Si les sommets de cette matrice dans la nouvelle position de la particule ne vérifient pas cette contrainte, la valeur de ce somment est prise égale à la valeur minimale

des sommets de la matrice. On corrige alors la valeur de la composante correspondante de la vitesse pour respecter l'équation de déplacement.

4.1.3.2 Types de fonctions coûts

Chaque position d'une particule est évaluée à l'aide de fonctions coûts construites à partir des indicateurs de performances des stations.

- En multi-objectif

Pour chaque indicateur de performance (taux de blocage et taux de coupure en général), on introduit une fonction coût élémentaire. Cette fonction est définie comme la somme pondérée des indicateurs filtrés (sur une fenêtre de taille typique entre 30 et 500 s) par station. Ainsi on définit les indicateurs de *Blocage* et *Coupure* globaux par :

$$Blocage = \sum_{i=1}^{N_{stations}} w_i b_i \,, \tag{4.8}$$

$$Coupure = \sum_{i=1}^{N_{stations}} u_i d_i \,, \tag{4.9}$$

où b_i et d_i sont les taux de blocage et de coupure de la station i, w_i et u_i les coefficients de pondération qui permettent de fixer les objectifs de l'optimisation en traitant en priorité les secteurs les plus problématiques. Par exemple, les objectifs du processus peuvent être :

- D'améliorer les performances moyennes sur l'ensemble des stations en termes d'indicateurs de blocage et coupure, en choisissant pour les coefficients de pondération :

$$\forall i \in \{1,....,N_{Stations}\} \quad w_i = u_i = 1 \,. \tag{4.10}$$

- D'améliorer les performances des pires stations du réseau en termes d'indicateurs de blocage et coupure, en choisissant par exemple pour les coefficients de pondération :

Pour les 10% pires stations du réseau, $w_i = u_i = 10$,
Pour les stations comprises entre les 10 et 20% pires stations du réseau, $w_i = u_i = 3$,
Pour le reste des stations, $w_i = u_i = 1$.

On peut aussi moduler les valeurs de ces coefficients par des indicateurs de charge par station, c'est-à-dire :

$$w_i \leftarrow w_i \times \omega_i \tag{4.11}$$

où ω_i le coefficient de modulation peut être défini par exemple par :

$$\omega_i = \frac{1/N_i^{Mob}}{\frac{1}{N_{Stations}} \times \sum_{j=1}^{N_{Stations}} 1/N_j^{Mob}} \tag{4.12}$$

où N_i^{Mob} est le nombre de mobiles servis par la station i,

$$\omega_i = \frac{\rho_i/N_i^{Mob}}{\frac{1}{N_{Stations}} \times \sum_{j=1}^{N_{Stations}} \rho_j/N_j^{Mob}} \tag{4.13}$$

où ρ_i est la charge de la station i.

En multi-objectif, on comparera l'efficacité de deux contrôleurs en les classant dans le plan coupure – blocage. Une particule i sera dite meilleure qu'une particule j (ou dominera la particule j) si : $Blocage^i \leq Blocage^j$ et $Coupure^i \leq Coupure^j$. La relation d'ordre ainsi introduite n'est pas totale. Mais elle va nous permettre de classer les particules en une série de fronts ordonnés. On mettra à jour a à chaque itération le front des particules optimales.

- En mono-objectif

Dans ce cas, on va agréger les différents indicateurs globaux de performance (décrits ci-dessus) à l'aide d'un coefficient de mixage β :

$$f = Blocage + \beta \cdot Coupure \tag{4.14}$$

Le facteur β fixé à l'initialisation du processus, permet de guider l'optimisation du contrôleur de logique floue vers différents compromis dans le plan taux de coupure – taux de blocage. En choisissant un fort coefficient β le processus va privilégier l'amélioration du taux de coupure sur l'amélioration du taux de blocage :

- Dans le cas de l'optimisation du paramètre charge cible DL (contrôle d'admission), le processus d'optimisation va donc réaliser un compromis entre le taux de coupure et le taux de blocage (voir chapitre 2).
- Dans le cas de l'optimisation du paramètre taille de fenêtre d'ajout (macro-diversité), le processus d'optimisation peut améliorer simultanément les taux de coupure et blocage (et aussi le blocage des liens crées par macro-diversité).

Nous avons testé et analysé l'impact de ces différentes fonctions coûts sur les contrôleurs optimisés par OEP (voir chapitre 6).

4.1.4. Schéma d'utilisation pour une optimisation du paramétrage des CLF

On propose le schéma suivant d'application de l'optimisation des CLF par essaim de particules, résumé par la figure 4.4 :

- Tracé des fronts de performances obtenus dans le plan taux de coupure – taux de blocage (par exemple) avec différents paramétrages uniformes des algorithmes RRM.
- Détermination d'un (ou plusieurs) paramétrage(s) initial(aux) uniforme(s) le(s) plus efficace(s), c'est à dire fournissant le(s) meilleur(s) compromis entre le taux de coupure, de blocage et de blocage des liens en MD.
- Tests de différents coefficients de mixage en mono-objectif (qui dépendront de l'environnement étudié et du paramétrage initial uniforme choisi).
- Choix d'un ou plusieurs couples (paramétrage initial, coefficient de mixage) pour les modes *off-line* et *on-line* de l'algorithme d'OEP en mono-objectif ou optimisation directe en multi-objectif.
- Comparaison des performances obtenues après optimisation par essaim de particules avec les performances initiales données par les fronts de paramétrage uniformes (taux de coupure, de blocage, délais de transmission…), en moyenne sur l'ensemble des stations ou en considérant les pires stations du réseau.

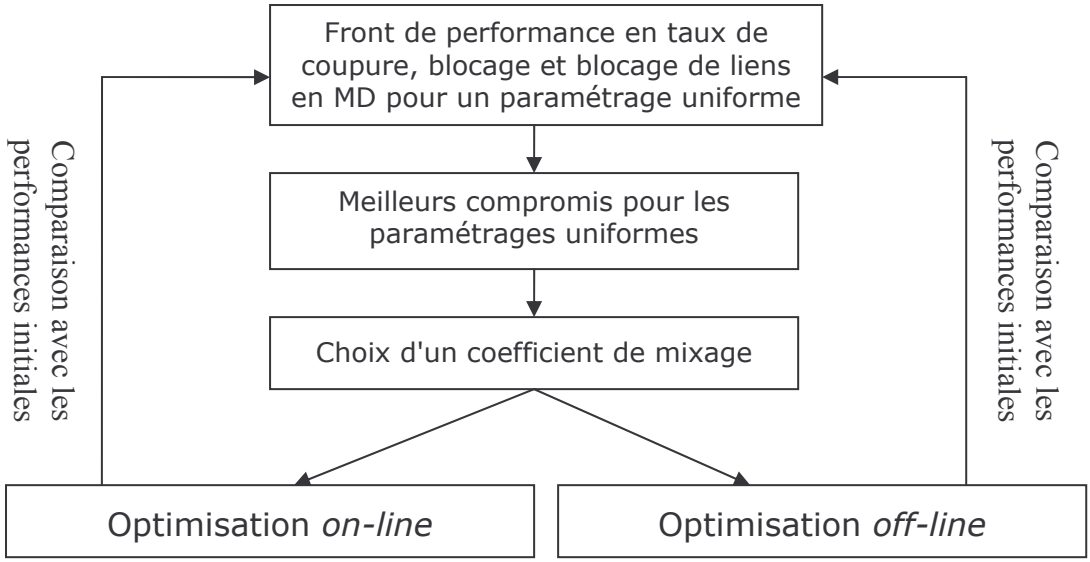


Figure 4.4 : Schéma d'utilisation de l'OEP pour l'optimisation des contrôleurs de logique floue.

Cette méthode d'optimisation a permis d'obtenir les premiers résultats d'amélioration du contrôle automatique des algorithmes RRM (voir chapitre 6). En général, les coefficients de mixage les plus efficaces en mono-objectif ne sont pas identiques dans les cas des optimisations *on-line* et *off-line*.

4.2. Optimisation par apprentissage par renforcement

Dans cette section, nous introduirons une méthode différente d'optimisation des contrôleurs de logique floue ; l'apprentissage par renforcement. Cette technique semble appropriée au problème de l'optimisation du contrôle automatique des paramètres d'un réseau UMTS, car dans ce problème, on ne connaît l'efficacité d'une modification (re-paramétrage d'un algorithme RRM) qu'après un certain délai d'application sur le réseau mobile et de manière indirecte en analysant certains indicateurs de performances. L'apprentissage par renforcement permettra de déterminer empiriquement (à l'aide d'un simulateur de réseau mobile ou de remontées de mesures terrain) les règles de décisions élémentaires qui maximiseront les critères de qualité choisis.

Par ailleurs, avec la méthode d'optimisation par apprentissage par renforcement, on préfère revenir à une description générale du problème de paramétrage automatique par un système d'inférence floue (plutôt que de se limiter à la structure du contrôleur FKBC décrit au chapitre 3) pour avoir une structure de contrôle plus générique. Il faudra certes adapter l'algorithme d'apprentissage par renforcement choisi – le *Q-learning* – au SIF, car cet algorithme n'est appliqué en principe qu'avec des processus de décisions Markoviens. Mais une fois adapté, l'algorithme de *Q-learning* permet un mode d'apprentissage plus riche que l'optimisation par essaim de particules (prise en compte parallèle des données de chaque cellule, au lieu de l'utilisation de fonctions coûts globales dans l'OEP) et utilise des notions, a priori, plus facilement transposables aux réseaux réels. Par exemple, dans les contrôleurs utilisant le *Q-learning*, deux modes de fonctionnement sont automatiquement définis: l'exploration qui permet de tester de nouvelles actions à appliquer au système et l'exploitation (utilisation du contrôleur une fois optimisée) qui choisit systématiquement la "meilleure" action pour le système. En particulier, le mode exploratoire permet la calibration du contrôleur (à l'aide de simulations ou sur un réseau expérimental), mais aussi l'adaptation d'un contrôleur à un environnement local nouveau (par exemple dans le cas d'un contrôleur ayant été optimisé sur une zone dense urbaine de Paris, et que l'on veut adapter à une zone dense urbaine de Lyon).

Enfin, avec l'apprentissage par renforcement, il est possible d'intégrer plusieurs niveaux de connaissance a priori : connaissance nulle, connaissance imprécise et connaissance précise. Au contraire, avec le CLF introduit au chapitre 3, la connaissance était introduite dans la structure même du contrôleur : au niveau de l'enchaînement des matrices de décisions du moteur d'inférence et des relations d'ordre entre les éléments des différentes matrices. L'OEP ne permettait pas d'explorer en-dehors de cette structure prédéfinie. Au contraire avec le *Q-learning*, il sera possible d'introduire des connaissances imprécises dans le SIF à l'initialisation, mais l'algorithme pourra ensuite aller explorer aléatoirement l'ensemble des règles de contrôle possibles

4.2.1. Introduction à l'apprentissage par renforcement et choix du Q-learning

Dans la communauté informatique, quatre familles de méthodes d'apprentissage co-existent. Chacune d'entre elles est caractérisée par le type d'information utilisée pour apprendre, ces familles étant classées en fonction du niveau d'information fourni à l'agent :

- L'apprentissage supervisé fournit, à chaque itération, l'action correcte que l'agent doit appliquer au système [Lin91].

- Les méthodes de type *distal teacher* n'indiquent pas l'action à appliquer, mais décrivent les effets désirés du contrôle exercé sur le système : pour exploiter ces méthodes, il est nécessaire d'introduire une fonction inverse permettant de déterminer l'action à appliquer en fonction de l'erreur estimée [Jor92].
- La mesure de performance, qui indique la qualité du contrôle exercé par l'agent pour un ensemble d'états du système ; ces méthodes d'apprentissage sont en général associées à des algorithmes évolutionnaires, comme les algorithmes génétiques [Cor96]. L'optimisation par essaim de particules appartient à cette famille de méthodes d'apprentissage (la qualité du contrôle étant mesurée à l'aide des fonctions coûts).
- Les méthodes d'apprentissage par renforcement, qui donne des récompenses ou des pénalités en fonctions des états atteints par l'agent (ou contrôleur) en train d'apprendre sur le système.

Au lieu d'utiliser des algorithmes combinatoires, on propose donc d'utiliser une méthode d'apprentissage par renforcement. L'algorithme le plus souvent utilisé mettant en oeuvre cette approche est le *Q-learning* [Glo97] [Jou98], qui sera décrit précisément au paragraphe 4.2.4. L'apprentissage par renforcement est un processus d'optimisation faiblement supervisé. Le principe de base est que si une action (modification locale du paramétrage RRM par exemple) conduit le système contrôlé (la zone considérée du réseau mobile dans notre cas) dans un nouvel état satisfaisant, alors elle doit être récompensée. En revanche, si le nouvel état est mauvais, alors cette action doit être pénalisée. L'algorithme de *Q-learning* est, quant à lui, un processus permettant d'apprendre par renforcement une fonction de qualité (voir définition au paragraphe 4.2.2.3, cette fonction estime la qualité d'un politique de contrôle) Q^* optimale, c'est à dire correspondant à une politique π^* optimale de choix des actions par l'agent. A partir de la fonction de qualité Q^* optimale, on est capable en utilisant les équations d'optimalité de Bellman de déterminer une politique déterministe optimale (cf. [Jou98] [El-A01]). Cette politique optimale sera mise en œuvre pour contrôler le paramétrage RRM du réseau mobile au cours de la phase d'exploitation.

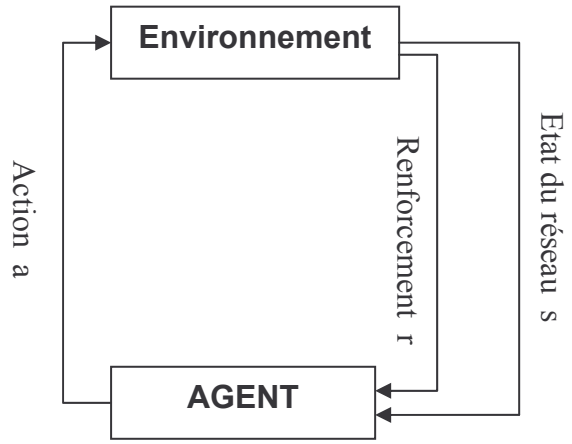


Figure 4.5 : Modèle standard d'apprentissage par renforcement.

Les états du système correspondent à des valeurs indicatrices décrivant la qualité de service du réseau mobile dans la zone, comme par exemple les taux de coupure et de blocage. Quand on utilise la logique floue, les fonctions d'appartenance aux ensembles flous permettent d'estimer l'état du système : selon les valeurs des indicateurs en entrées, le degré d'appartenance à chaque label d'entrée est estimé. En fonction de ces valeurs floues représentant l'état du système, le contrôleur détermine en utilisant le système d'inférence floue, les actions correctives à appliquer au paramétrage RRM du réseau.

Dans le cas de l'application de l'apprentissage par renforcement à un SIF, l'algorithme tient compte de l'action globale appliquée au système, de l'ancien et du nouvel état de la zone, des règles élémentaires qui ont été mises à contribution dans le SIF pour déterminer les actions élémentaires, et de la valeur de la fonction de renforcement. Le renforcement est une valeur scalaire permettant d'estimer le degré de satisfaction des critères de qualité choisis et qui est lié à la valeur (calculée à partir des indicateurs de performances) attribuée à chaque état du système. Cette valeur de renforcement permettra à l'agent d'évaluer la qualité de l'action qu'il a effectuée à l'étape précédente et d'ainsi mettre à jour son modèle de prise de décisions (les fonctions de qualité des règles de décisions élémentaires) en fonction de cette valeur de renforcement et de l'état courant du réseau.

Applications des méthodes d'apprentissage par renforcement au réseaux mobiles

Les algorithmes d'apprentissage par renforcement ont déjà été appliqués au contrôle des réseaux mobiles. Ainsi dans les documents [Sen04] et [Yu04], les problèmes à traiter respectifs ont chacun été formulés comme un processus de décision Markovien*. Dans chaque cas, un algorithme d'apprentissage par renforcement a été utilisé, qui est une version adaptée de l'algorithme du *Q-learning* décrit au paragraphe 4.2.4.2. Dans le premier cas, il s'agit de résoudre le problème du contrôle d'admission des appels (CAC) dans les réseaux mobiles. Dans le second cas, la question du *QoS provisioning* dans les réseaux mobiles est étudiée. Dans ces deux exemples, la modélisation du système contrôlant le réseau mobile est de type déterministe et non pas un système de type inférence floue.

Adaptation de l'algorithme du Q-learning

Dans les articles [Sen04] et [Yu04], en comparaison avec l'algorithme du *Q-learning* qui sera décrit au paragraphe 4.2.4.2, la principale différence est que l'équation de l'étape (7) de mise à jour de la fonction de qualité est remplacée par :

$$Q_{t+1}(s,a) \leftarrow \begin{cases} Q_t(s,a) + l_t \times \Delta Q_t(s,a) & \text{if } s = s_t \text{ and } a = a_t \\ Q_t(s,a) & \text{otherwise} \end{cases} \tag{4.15}$$

Avec

$$\Delta Q_t = r_t + \gamma \times V(s') - Q_t(s,a) \tag{4.16}$$

$$V(s') = \max_{a' \in A(s')} \left[Q_t(s',a') \right], \tag{4.17}$$

et a est l'action réalisée (donnée par la politique d'Exploration / Exploitation) dans l'état s, menant au nouvel état s', avec une valeur de renforcement r_t, γ étant le facteur de décote (ou d'atténuation) et l_t le taux d'apprentissage.

Dans [Sen04], le taux d'apprentissage est défini par $l_t = \dfrac{1}{1 + visit_t(s_t, a_t)}$ où $visit_t(s_t, a_t)$ représente le nombre total de fois que ce couple état - action a été visité au cours de l'algorithme, et dans [Yu04] le taux d'apprentissage l est aussi défini comme étant inversement proportionnel à la fréquence de visite du couple état - action jusqu'à l'itération présente.

* Un processus de décision Markovien (*Markovian Decision Process ; MDP*) discret est un modèle mathématique pour l'optimisation de la prise de décisions séquentielles dans des phénomènes (typiquement des systèmes dynamiques) où interviennent des prises de décision en général en environnement stochastique.

4.2.2. Définitions des fonctions utilisées dans les méthodes d'apprentissage par renforcement

4.2.2.1 Le modèle mathématique

Le système (ou environnement) doit être décrit par un processus de décision Markovien qui est décrit par un quadruplet (S, A, T, R). S correspond à l'ensemble des états, A à l'ensemble des actions, T à la fonction de probabilité de transition (entre les états s et s', quand l'action a est déclenchée) et R est la fonction of renforcement. $T(s,a,s')$ et $R(s,a,s')$ dépendent de l'état courant s, de l'action a, qui amène le système dans le nouvel état s', et pour les processus non déterministes aux nouveaux états s'. Les fonctions T et R sont Markoviennes car leurs valeurs ne dépendent pas des états et/ou actions antérieurs.

Dans le cas des réseaux mobiles, *le processus est non déterministe*. En effet, une même action a (modification de paramétrage) dans un état s pour une cellule peut conduire à plusieurs nouvelles situations différentes, selon l'arrivée aléatoire et l'établissement ou non de nouveaux appels, la coupure de certaines communications, la fin d'autres communications, le déplacement aléatoire des utilisateurs dans la cellule… Cependant il est impossible d'établir une expression analytique ou d'estimer la fonction de probabilité de transition T de notre problème, car le nombre d'état accessibles s' à partir d'un état s est infini (les variables d'indicateurs de performances étant continues) et nous n'avons aucun moyen pour estimer la probabilité de chaque transition. L'apprentissage par renforcement permet de déterminer les actions a qui ont, dans un état s, les plus grandes probabilité d'amener le réseau vers de nouveaux états s' aillant de meilleurs performances et/ou qualité de service, sans connaître la fonction de probabilité de transition du système.

Introduction des fonctions d'évaluation

Dans le contexte du processus de décisions Markovien, nous introduisons trois fonctions d'évaluation, qui nous permettent d'évaluer l'efficacité globale d'une politique en tenant compte des renforcements futurs.

- Modèle par somme à horizon fini

Nous introduisons le paramètre h de temps à vie (ou mémoire de l'efficacité d'une politique). Nous définissons la fonction d'évaluation comme l'espérance de la somme des renforcements reçus pendant h itérations du système :

$$E\left(\sum_{t=1}^{h} r_t\right) = E\left(r_1 + ... + r_h\right) \tag{4.18}$$

- Modèle par somme décotée à horizon infini

La fonction d'évaluation est définie comme l'espérance de la somme décotée de tous les renforcements reçus dans le futur. γ[*] est le facteur de décote utilisé pour pondérer les renforcements en fonction du temps. Plus le renforcement est lointain dans le futur et plus son poids est faible dans la fonction d'évaluation :

$$E\left(\sum_{t=1}^{\infty} \gamma^{t-1} r_t\right) = E\left(r_1 + \gamma . r_2 + \gamma^2 . r_3 + ...\right) \tag{4.19}$$

- Modèle du renforcement moyen

[*] Le facteur γ est défini dans l'intervalle [0, 1].

Avec cette fonction d'évaluation, le processus d'apprentissage essaie de maximiser la moyenne de tous les renforcements reçus :

$$\lim_{h\to\infty} E\left(\frac{1}{h}\sum_{t=1}^{h} r_t\right) \tag{4.20}$$

Le problème avec ce modèle est que deux politiques différentes peuvent avoir une évaluation globale identique, alors que la première reçoit beaucoup de récompenses en début de processus et la seconde en reçoit peu au début, mais beaucoup sur la fin du processus.

Conclusions

Dans le cas des réseaux mobiles, nous préférons utiliser le modèle par somme décotée, car il est important que le processus de contrôle fournisse des gains de performances et qualité de services à court terme dans le réseau : nous privilégions donc une certaine réactivité du contrôleur, notamment à cause de la variabilité non déterministe de l'état du réseau (à cause de multiples phénomènes aléatoires liés au comportement des utilisateurs et au canal de transport radio).

4.2.2.2 La politique de l'agent (ou contrôleur)

Le comportement de l'agent est déterminé par sa fonction politique, $\pi : S \times A \to [0,1]$. Pour chaque état, une action donnée est effectuée en appliquant une certaine politique choisie. Dans une modélisation stochastique, π représente la probabilité que dans un état donné *s*, l'action *a* soit déclenchée :

$$\forall\ s \in S \qquad \sum_{a\in A(s)} \pi(s,a) = 1 \tag{4.21}$$

Dans un modèle déterministe de politique, la fonction politique sélectionne dans un état *s*, une seule action *a* qui amènera le système à un nouvel état s' : dans ce cas, il y a une seule action qui peut être choisie par état. Dans l'optique d'un contrôleur du paramétrage RRM de nodes B, on choisit un type déterministe de politique : pour un même état, on applique toujours la même décision. En revanche, le processus reste globalement de nature non déterministe, cette décision pouvant mener à différents états postérieurs.

Fonctions d'évaluation de politiques

A partir de chaque modèle de fonction d'évaluation introduit au paragraphe 4.2.2.1, nous pouvons définir une fonction d'évaluation de politiques, $V^{\pi} : S \to \Re$ qui permette de comparer les efficacités respectives de différentes politiques. Par exemple, avec le modèle par somme décotée à horizon infini et sous les hypothèses qu'on applique toujours la même politique π au cours du processus et que l'état s est l'état initial (γ dans [0,1], étant le facteur d'atténuation), on définit la fonction dévaluation V^{π} :

$$V^{\pi}(s) = E_{\pi}\left(\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \;\middle|\; s_t = s\right) \tag{4.22}$$

La fonction V^{π} est Markovienne, car ses valeurs ne dépendent pas des états et/ou actions antérieures du processus. Dans le cas d'un processus déterministe (pas le cas des réseaux mobiles), l'expression de V^{π} se réduit à :

$$V^{\pi}(s) = \left(\sum_{k=0}^{\infty} \gamma^{k} r_{t+k+1} \;\middle|\; s_t = s \right) \tag{4.23}$$

Remarque : $\gamma = 1$ est théoriquement possible si et seulement s'il existe une probabilité non nulle d'atteindre un état terminal du système en appliquant cette politique. Or cette notion d'état terminal, n'a pas de sens pour les réseaux mobiles. En effet, en permanence de nouveaux utilisateurs peuvent demander des ressources au réseau.

4.2.2.3 Introduction des fonctions de qualité

Avec notre modèle de fonction d'évaluation, nous définissons la fonction de qualité (ou *Q-fonction*) d'une politique comme l'espérance de la somme décotée des futurs renforcements (ou récompenses) quand on commet l'action *a* dans l'état *s* en se fondant à chaque étape future sur la politique π. Donc *Q* est une fonction du couple état *s* et action *a*, $Q^{\pi}(s,a)$, associée à une politique π.

$$Q^{\pi}(s,a) = E_{\pi}\left(\sum_{k=0}^{\infty} \gamma^{k} r_{t+k+1} \;\middle|\; s_t = s, a_t = a \right) \tag{4.24}$$

Les politiques étudiées et optimisées seront déterministes, mais l'environnement du réseau mobile étant non déterministe, nous ne pouvons pas simplifier l'expression (4.24) ci-dessus.

4.2.3. Les équations de Bellman et notion de politique optimale

4.2.3.1 Enoncés des équations de Bellman

Dans le paragraphe 4.2.2.2, nous avons introduit l'expression suivante pour la fonction d'évaluation d'une politique π, dans un modèle de somme décotée à horizon infini :

$$V^{\pi}(s) = E_{\pi}\left(r_{t+1} + \gamma \sum_{k=0}^{\infty} \gamma^{k} r_{t+k+2} \;\middle|\; s_t = s \right) \tag{4.25}$$

Si la politique π est déterministe, nous déduisons, en développant la relation (4.25) en fonction des valeurs de probabilité de transition *T* et de l'espérance du renforcement $R(s,\pi(s),s') = E_{\pi}\left(r(s,\pi(s),s') \;\middle|\; s_t = s\right)$, la première équation de Bellman :

$$V^{\pi}(s) = \sum_{s' \in S} T(s,\pi(s),s') \times \left[R(s,\pi(s),s') + \gamma . V^{\pi}(s') \right] \tag{4.26}$$

La politique est déterministe, mais le processus global est aléatoire (ou non déterministe), car dans cette expression la fonction de probabilité peut ne pas être constante (*T=1* pour un processus déterministe) : ceci correspond bien à la situation des réseaux mobiles que nous voulons étudier.

Plus généralement, si la politique est non déterministe, nous obtenons la seconde équation de Bellman (cette relation pourrait nous être utile si l'on souhaite ensuite développer des politiques de contrôle non déterministes), en développant la relation (4.25) en fonction des valeurs de probabilité de transition *T*, des probabilités de choisir l'action *a* dans l'état *s* et de l'espérance du renforcement $R(s,a,s') = E_{\pi}\left(r(s,a,s') \;\middle|\; s_t = s\right)$:

$$V^{\pi}(s) = \sum_{a \in A(s)} \pi(s,a) \left(\sum_{s' \in S} T(s,a,s') \times \left[R(s,a,s') + \gamma . V^{\pi}(s') \right] \right) \quad (4.27)$$

De la même façon, nous introduisons l'équation (4.28) de Bellman pour les Q-fonctions, valable pour une politique π déterministe ou non. Cette relation est obtenue en développant l'équation (4.24) avec les valeurs de probabilité de transition T et les probabilités de choisir l'action a dans l'état s :

$$Q^{\pi}(s,a) = \sum_{s' \in S} T(s,a,s') \times \left[R(s,a,s') + \gamma . V^{\pi}(s') \right] \quad (4.28)$$

Conséquences des équations de Bellman

A partir des relations (4.26), (4.27) et (4.28), nous déduisons une relation importante entre la fonction d'évaluation V et la fonction de qualité Q d'une politique π, qui sera utilisée directement dans l'algorithme d'apprentissage du *Q-learning*.

Si la politique π est déterministe, l'action choisie vérifie $a = \pi(s)$. Et d'après les équations de Bellman (4.26) et (4.28), nous déduisons que :

$$V^{\pi}(s) = Q^{\pi}(s, \pi(s)) \quad (4.29)$$

Si la politique π est non déterministe, d'après les équations de Bellman (4.27) et (4.28), nous déduisons que :

$$V^{\pi}(s) = \sum_{a \in A(s)} \pi(s,a) Q^{\pi}(s,a) \quad (4.30)$$

4.2.3.2 Définition et existence d'une politique optimale

Dans le processus d'apprentissage par renforcement, il est important de pouvoir comparer les efficacités respectives de deux politiques. Pour cela, nous introduisons la relation d'ordre, non totale suivante :

Soit π et π' deux politiques (déterministes ou non) sur le processus MDP M. La politique π' est une amélioration de la politique π (ou est dite meilleure que π) si et seulement si :

$\forall\, s \in S \quad V^{\pi'}(s) \geq V^{\pi}(s)$ avec une inégalité stricte pour au moins un état de l'ensemble S.

Par conséquent intuitivement, une politique π^* est dite optimale, s'il n'existe pas de politique π qui puisse être considérée comme une amélioration de π^* (au sens de la relation d'ordre partielle introduite ci-dessus).

- *Définition mathématique d'une politique optimale*

Une politique π^* est optimale si et seulement pour toute autre politique π' :

$$\forall\, s \in S \quad V^{\pi^*}(s) \geq V^{\pi'}(s) \quad (4.31)$$

La définition mathématique d'une politique optimale n'est pas équivalente à l'énoncé intuitif précédent car la relation d'ordre introduite n'est pas totale.

- *Théorème d'existence d'une politique déterministe optimale pour tout processus MDP fini*

Dans le document [Gar04], il a été établi que pour tout processus de décisions Markovien fini *M*, si le facteur d'atténuation vérifie $0 \le \gamma < 1$, alors il existe au moins une politique déterministe optimale pour le processus.

Et nous déduisons des relations (4.26), (4.28) et (4.29) (valables pour une politique π déterministe), les équations de conditions d'optimalité de Bellman, vérifiée par une fonction d'évaluation optimale V^* et la fonction de qualité optimale Q^* :

$$V^*(s) = \max_{a \in A(s)} \left(\sum_{s' \in S} T(s,a,s') \times \left[R(s,a,s') + \gamma . V^*(s') \right] \right) \tag{4.32}$$

$$\text{Et } Q^*(s,a) = \sum_{s' \in S} T(s,a,s') \times \left[R(s,a,s') + \gamma . \max_{a' \in A(s)} \left(Q^*(s,a') \right) \right] \tag{4.33}$$

Les fonctions d'évaluation V^* et de qualité Q^* sont dites optimales quand ce sont les fonctions d'évaluation et de qualité d'une politique π^* elle-même optimale au sens de la relation (4.31).

- *Introduction d'un algorithme de calcul d'une politique déterministe optimale*

Nous supposons ci-dessous que nous pouvons estimer pour le processus MDP *M* considéré, la fonction de probabilité de transition *T* et la fonction de renforcement R. Nous pouvons alors construire récursivement la politique déterministe π^* avec l'algorithme suivant :

(1) Nous tirons une politique aléatoire initiale π^0.

…

(n) $\forall\ s \in S \quad \pi^{n+1}(s) = \arg\max_{a \in A(s)} \left(Q^{\pi^n}(s,a) \right)$

Condition d'arrêt:

$$\text{Si} \left[\quad \forall\ s \in S \quad \pi^{n+1}(s) = \pi^n(s) \quad \right] \tag{4.34}$$

$$\text{alors} \left[\quad \forall\ s \in S \ \ V^{\pi^{n+1}}(s) = \max_{a \in A(s)} \left(\sum_{s' \in S} T(s,a,s') \times \left[R(s,a,s') + \gamma . V^{\pi^{n+1}}(s') \right] \right) \right] \tag{4.35}$$

Et par conséquent, $\pi^* = \pi^{n+1}$ vérifie la condition de Bellman d'optimalité (4.32) pour la fonction d'évaluation *V*, et π^* est donc une politique déterministe optimale.

Remarque : Cette méthode permet de déterminer une politique déterministe optimale, mais nécessite de pouvoir estimer les fonctions *T* et *R*. Or, dans le contexte des réseaux mobiles, il n'est pas possible d'estimer ces fonctions, car le nouvel état *s'* dépend de nombreux paramètres extérieurs (arrivée de nouveaux utilisateurs, mobilité des utilisateurs, contraintes de l'environnement radio…). Il est nécessaire d'introduire des stratégies permettant, d'une part, d'apprendre la valeur de chaque politique, et d'autre part de déterminer une politique déterministe optimale, et qui n'utilisent pas les probabilités de transition.

4.2.4. Les algorithmes d'apprentissage

Ci-dessous, nous supposons qu'il n'est pas possible pour le processus de décisions Markovien considéré *M*, d'estimer les fonctions *T* et *R*. Nous allons décrire deux algorithmes permettant

respectivement: d'apprendre la valeur de toute politique (algorithme des différences temporelles) et de déterminer une politique déterministe optimale (algorithme du *Q-learning*).

4.2.4.1 Evaluation de la valeur d'une politique : algorithme des différences temporelles (TD(γ))

Dans [Gar04], l'algorithme suivant des différences temporelles permettant d'estimer la valeur de toute politique π sur le processus de décisions Markovien *M* est décrit.

(1) Initialiser $V_0(s) = 0$ pour tout $s \in S$
(2) $\alpha_t \in [0,1]$, $\gamma \in [0,1[$
(3) Choisir l'état initial *s*
(4) Répéter jusqu'à la convergence :
(5) Déterminer l'action *a* à l'aide de la politique π: $\pi(s) \to a$
(6) Réaliser l'action *a*, qui mène au nouvel état *s'*, avec la valeur de renforcement *r* (figure 4.5)
(7) $V_{t+1}(s) \leftarrow (1-\alpha_t) \times V_t(s) + \alpha_t \times (r + \gamma \times V_t(s'))$ (4.36)
(8) Si *s'* est un état terminal alors choisir un autre état initial *s*
(9) Sinon $s \leftarrow s'$
(10) Fin de la boucle

On pourra utiliser cet algorithme dans le contexte des réseaux mobiles pour évaluer la valeur d'une politique particulière ou comparer les efficacités de deux politiques. En particulier, à l'aide de l'algorithme des différences temporelles, on pourra estimer la valeur d'une politique de contrôle d'un CLF optimisé par essaim de particules, et comparer sa valeur avec une politique de contrôle obtenue avec l'algorithme du *Q-learning*.

- *Théorème de convergence*

Le théorème suivant garantit la convergence de l'algorithme des différences temporelles vers la fonction d'évaluation V^π de la politique π pourvu que les conditions 1) 2) et 3) soient remplies.

Hypothèses : M est un processus de décisions Markovien fini et non déterministe et π est la politique à évaluer.

1) Chaque état *s* du MDP *M* a été visité un nombre infini de fois.
2) $0 \le \gamma < 1$.
3) $\sum_{t=0}^{\infty} \alpha_t = \infty$ et $\sum_{t=0}^{\infty} \alpha_t^2 < \infty$.

Conclusions : La fonction *V* tend vers la fonction d'évaluation V^π de la politique π, où *V* est donnée par l'équation (4.36) de l'étape (7) de l'algorithme des différences temporelles.

Pour estimer la valeur d'une politique avec une précision acceptable, l'algorithme doit donc visiter un nombre minimal de fois les différents états du processus. Dans ce but, on pourra relancer pour la même politique π l'algorithme des différences temporelles avec différents niveaux de charge et distributions de trafic dans le réseau mobile.

4.2.4.2 Détermination d'une politique optimale : l'algorithme du Q-learning

Dans [Glo97] et [Jou98], l'algorithme de *Q-learning* est décrit dans le cadre des problèmes classiques de *cart-centering*, de guidage de bateau et de contrôle de *mountain-car*. Pour chacun de ces processus de décisions Markoviens, l'algorithme permet de déterminer une

politique optimale. Cela signifie qu'à chaque étape et pour chaque état du système, le module de contrôle est capable de choisir l'action qui permettra au système d'atteindre les meilleures performances possibles. Nous résumons ci-dessous les différentes étapes de l'algorithme du *Q-learning* pour un environnement non déterministe :

(1) Initialiser $Q_0(s,a)=0$ pour tout $s \in S, a \in A$

(2) $\alpha_t \in [0,1]$, $\gamma \in [0,1[$

(3) Choisir l'état initial *s*

(4) Répéter jusqu'à la convergence :

(5) Déterminer l'action *a* à l'aide de la politique d'exploration / exploitation EEP (*) : $EEP(s, Q_t, S, A) \rightarrow a$

(6) Réaliser l'action *a*, qui mène au nouvel état *s'*, avec la valeur de renforcement *r*

(7) (**) $$Q_{t+1}(s,a) \leftarrow (1-\alpha_t) \times Q_t(s,a) + \alpha_t \times \left(r + \gamma \max_{a' \in A(s')} Q_t(s',a') \right) \qquad (4.37)$$

(8) Si *s'* est un état terminal alors choisir un autre état initial *s*

(9) Sinon $s \leftarrow s'$

(10) Fin de la boucle

(*) $EEP(s, Q_t, S, A)$ est une politique d'exploration / exploitation qui, avec une probabilité prédéterminée ε (< *0.5*), choisit aléatoirement une nouvelle action *a'*, et avec la probabilité *(1-ε)* choisit l'action *a* qui maximise la fonction de qualité $Q(s,a)$. Ce choix au hasard permet à l'algorithme d'explorer et tester des actions différentes.

(**) Si l'environnement est déterministe, c'est à dire quand le passage d'un état au suivant n'est pas aléatoire; l'équation (4.37) de l'étape (7) est remplacée par :

$$Q_{t+1}(s,a) \leftarrow r + \gamma \max_{a' \in A(s')} Q_t(s',a') \qquad (4.38)$$

Dans le cas des réseaux mobiles, l'environnement est non déterministe (d'où on utilise l'équation (4.37)). Le système composé par le problème du contrôle du paramétrage d'une zone d'un réseau mobile ne constitue pas a priori un processus de décisions Markovien. Pour se ramener au cas d'un MDP, on utilise la modélisation de ce problème par le système d'inférence floue introduit au chapitre 3. Il sera aussi nécessaire d'adapter l'algorithme de *Q-learning* à une utilisation avec un SIF (voir le paragraphe 4.2.5). Enfin, pour que l'algorithme converge effectivement vers une fonction de qualité correspondant à une politique optimale, l'algorithme doit visiter un nombre minimal de fois les différents couples états – actions du processus (voir le théorème de convergence ci-dessous).

- *Théorème de convergence*

Le théorème suivant garantit la convergence de l'algorithme de *Q-learning* vers la fonction de qualité Q^* d'une politique optimale π^* pourvu que les conditions 1) 2) et 3) soient remplies.

Hypothèses : M est un processus de décisions Markovien fini et non déterministe

1) chaque paire état - action *(s,a)* du MDP *M* a été visité un nombre infini de fois.

2) $0 \le \gamma < 1$.

3) $\sum_{t=0}^{\infty} \alpha_t = \infty$ et $\sum_{t=0}^{\infty} \alpha_t^2 < \infty$.

Conclusions : La fonction Q tend, indépendamment de la politique EEP d'exploration / exploitation vers la fonction de qualité Q^* d'une politique optimale π^*, où Q est donnée par l'équation (4.37) de l'étape (7) de l'algorithme de *Q-learning*.

- *Détermination de la politique optimale π**

L'algorithme de *Q-learning* converge donc pour toute EEP vers une fonction de qualité correspondant à une politique optimale. On déduit de la condition d'optimalité (4.33) une moyen permettant de déterminer une politique déterministe optimale π^*, en choisissant (pour tout état s) $\pi^*(s)$ comme étant l'action qui maximise la fonction de qualité optimale $Q^*(s,a)$:

$$\pi^*(s) = \arg\max_{a \in A(s)} Q^*(s,a) \tag{4.39}$$

En conséquence, la politique π^* est bien une politique optimale, car elle vérifie l'équation suivant qui est équivalente aux équations de Bellman d'optimalité (Eq. (4.32) et (4.33)).

$$V^{\pi^*}(s) = \max_{a \in A(s)} Q^*(s,a) \tag{4.40}$$

- *Politique classique d'Exploration / Exploitation EEP*

Une politique possible d'Exploration / Exploitation, décrite dans [Glo97] [Jou98], est la politique pseudo stochastique (ou *ε-greedy*) définie par l'algorithme suivant :

$\varepsilon \in [0,1]$, $s \in S$.
$a \leftarrow \arg\max_{a \in A(s)} Q(s,a)$.
$a' \leftarrow$ choisir a avec la probabilité ε et choisir une action aléatoire dans $A(s)/\{a\}$ avec la probabilité $(1-\varepsilon)$.
Appliquer l'action $a' \rightarrow EEP(s,Q,S,A)$ en sortie.

<u>*Conclusions*</u>

L'algorithme de *Q-learning* fournit donc une méthode d'apprentissage efficace sur les processus MDP finis. Cependant, dans le cas du contrôle du paramétrage des réseaux mobiles le processus à contrôler est beaucoup plus complexe que les problèmes classiques (effets de couplage entre nodes B…). Par conséquent, il va falloir adapter le *Q-learning* au système d'inférence floue défini au chapitre 3, ramenant ainsi le problème continu étudié à un MDP fini (en tenant compte des degrés d'appartenance aux différents labels d'entrées). Enfin, un point important va concerner la phase d'exploration de l'algorithme. Pour que l'algorithme converge effectivement vers une fonction de qualité correspondant à une politique optimale, il doit visiter un nombre minimal de fois les différents couples états – actions du processus. Dans ce but, on pourra lancer plusieurs fois l'algorithme de *Q-learning* avec différents niveaux de charge et distributions de trafic dans le réseau mobile, et sans réinitialiser les fonctions de qualité élémentaires correspondant à chaque règle de décisions. De cette manière l'algorithme pourra explorer successivement différents états du système.

4.2.5. Adaptation du Q-learning aux systèmes d'inférence floue

Dans ce paragraphe nous allons décrire l'adaptation de l'algorithme de *Q-learning* à un système d'inférence floue dans le contexte d'un environnement continu. Cette adaptation permet ainsi d'utiliser le *Q-learning* dans le cas d'un système dont l'état est défini à partir de variables continues réelles telles que les indicateurs de performances des réseaux mobiles. En

effet, cet algorithme n'est normalement applicable que dans le cadre d'un processus de décision Markovien (voir paragraphe 4.2.2), qui nécessite que l'ensemble S des états du processus soit discret.

4.2.5.1 Définition d'un système d'inférence floue adapté à l'algorithme de Q-learning

Au paragraphe 3.1.1, nous avions décrit les composantes d'un système d'inférence floue (SIF). Nous rappelons qu'un SIF est basé sur N_R règles de décisions de la forme suivante :

$$R_i : \text{Si } s_1 \text{ est } L^i_1 \text{ et } \ldots s_{Ns} \text{ est } L^i_{Ns}$$
$$\text{Alors } y_1 \text{ est } O^i_1 \text{ et } \ldots y_{No} \text{ est } O^i_{No}$$

Les différents éléments suivants composent et définissent le SIF : la ième règle de la base de règles R_i, le vecteur d'état du système $\underline{s} = (s_1, s_2, \ldots, s_{Ns})$ (Ns étant le nombre d'entrées), le terme linguistique L^i_j (label flou) correspondant à la variable d'entrée s_j dans la règle R_i, la fonction d'appartenance $\mu_{L^i_j}$ pour le label flou L^i_j ; le vecteur des No variables de sortie $(y_1, y_2, \ldots, y_{No})$, le jème terme linguistique O^i_j (label flou) correspondant à la variable de sortie y_j dans la règle R_i.
Ci-dessous, nous détaillons la définition d'un système d'inférence floue permettant le contrôle du seuil de charge cible *downlink* dans un réseau UMTS. Dans la suite du paragraphe nous expliciterons l'application de l'algorithme d'apprentissage de *Q-learning* au SIF ainsi défini.

Définition détaillée d'un SIF adapté au Q-learning pour le contrôle d'admission

Au paragraphe 3.3.3 dans le cadre d'une modélisation par un processus FKBC du contrôle de l'algorithme d'admission, le moteur d'inférence a été défini par un système de quatre matrices de décisions. Sur la figure 4.6 suivante, on a simplifié ce système de matrices, en supprimant le dernier niveau (qui prenait en compte la correction effectuée à l'itération précédente) :

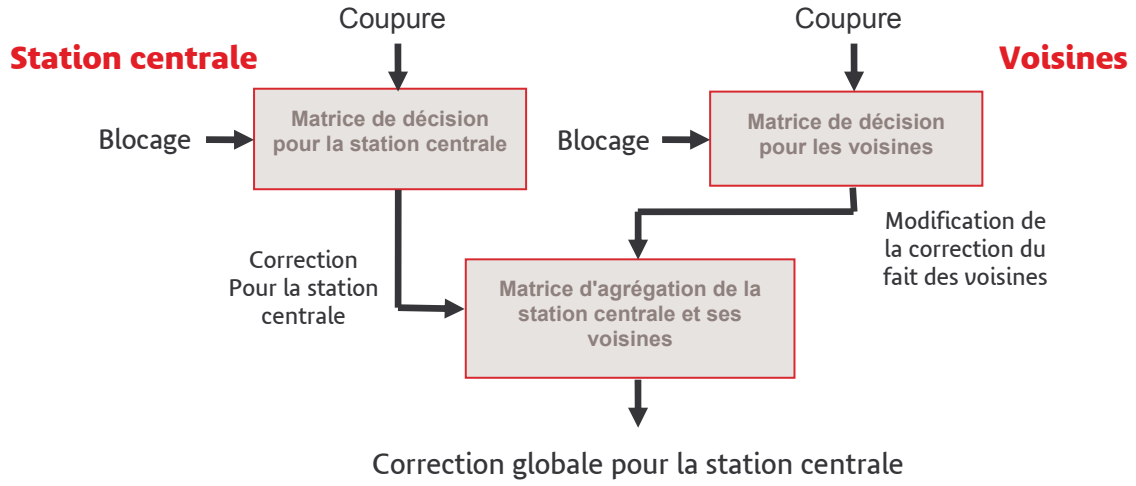


Figure 4.6 : Système de matrices de décisions pour le contrôle de l'algorithme d'admission

Nous allons expliquer ci-dessous comment transformer ce processus FKBC en un système d'inférence floue sur lequel on pourra appliquer l'algorithme du *Q-learning* adapté aux systèmes flous.

Définition du vecteur d'état en entrées

Le vecteur d'entrées est défini par un vecteur de couples (taux de blocage, taux de coupure) pour la station centrale (station 0) et ses voisines :

$$s = (b_0, d_0) \times (b_1, d_1) \times \ldots \times (b_{Ns}, d_{Ns}), \quad (4.41)$$

où *Ns* est le nombre de voisines de la station centrale. L'ensemble de définition correspondant à ce vecteur d'entrées est $[0,1]^2 \times [0,1]^2 \times \ldots \times [0,1]^2$.

Définition des labels flous d'entrées

On introduit deux labels flous possibles pour les taux de blocage b_j et les taux de coupure d_j aussi bien pour la station centrale que pour les stations voisines : faible et fort.

Définition des fonctions d'appartenance

Les fonctions d'appartenance pour les variables d'entrées de taux de blocage et taux de coupure sont définies avec des formes de types triangulaires. Les ensembles flous pour le taux de blocage de la station centrale sont pris identiques à ceux du taux de blocage moyen sur l'ensemble des voisines, et de même pour le taux de coupure (pas nécessaire avec le *Q-learning*).

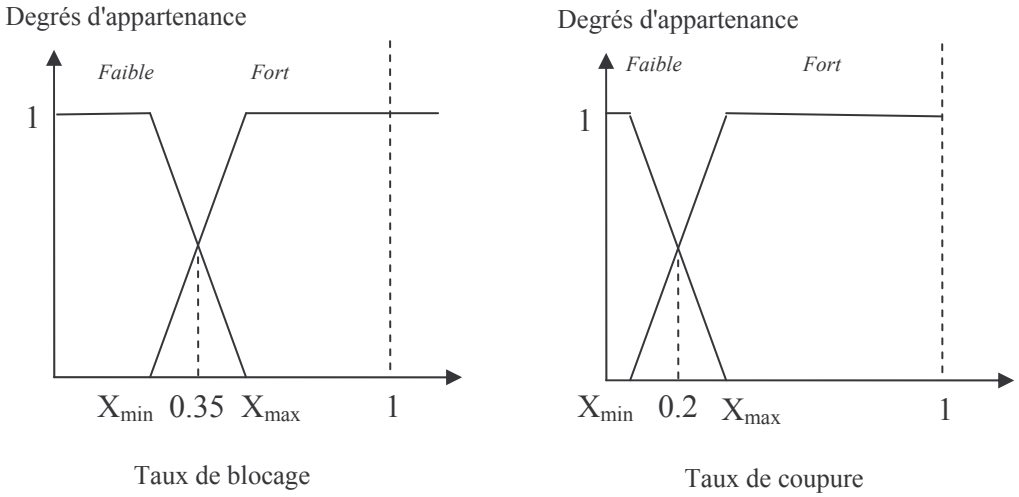


Figure 4.7 : Fonctions d'appartenance des taux de blocage et taux de coupure

Définition de la variable de sortie et de ses labels possibles

La variable de sortie Δc correspond à la modification du paramètre de seuil de charge cible downlink (paramètre du contrôle d'admission). Nous introduisons trois labels flous pour la variable de sortie de correction Δc : *baisse*, *pas de modification* et *hausse*. A chaque label flou pour Δc correspond une valeur numérique de correction. Par exemple les valeurs numériques des trois labels précédents peuvent respectivement être prises égales à: *-0.1*, *0* et *+0.1*.

Définition de l'ensemble des règles d'inférences

Nous construisons donc un SIF fondé sur 16 règles d'inférences (= 2^4 règles, pour les 4 variables d'entrée ayant chacune 2 labels flous possibles) avec une variable de sortie (par règle) pouvant correspondre à 3 labels flous différents.

Par ailleurs, nous pouvons introduire dans notre SIF de la connaissance a priori. Cette connaissance peut être de deux catégories : imprécise ou précise.

- Exemple de connaissance a priori précise :

Si $B_{Centrale}$, $D_{Centrale}$, $B_{Moy\ Vois.}$ et $D_{Moy\ Vois.}$ sont *forts*,

Alors *diminuer* le seuil de charge cible DL, c'est à dire $\Delta c = -0.10$.

- Exemple de connaissance a priori imprécise :

Si $B_{Centrale}$ est fort, $D_{Centrale}$ est faible, $B_{Moy\ Vois.}$ est faible et $D_{Moy\ Vois.}$ est *fort*,

Alors *plutôt diminuer* le seuil de charge cible DL, c'est-à-dire :

$\Delta c = -0.10$ pour le label de correction O_1^i (*baisse moyenne*),

$\Delta c = -0.05$ pour le label de correction O_2^i (*baisse faible*),

et $\Delta c = 0$ pour le label de correction O_3^i (*pas de modification*).

Nous expliquerons au paragraphe 4.2.6 comment prendre en compte cette connaissance a priori. La figure 4.8 suivante résume la structure du SIF que nous avons ainsi défini pour le paramétrage automatique du contrôle d'admission.

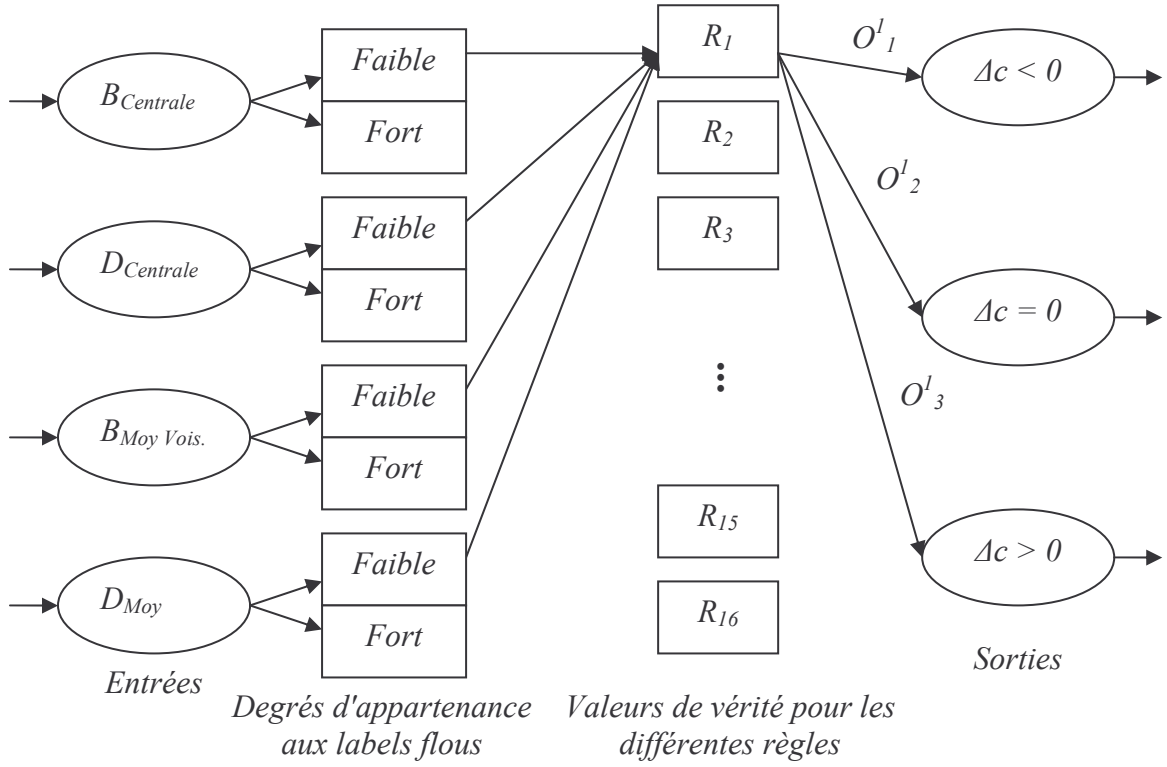


Figure 4.8 : Structure du SIF défini pour le paramétrage automatique du contrôle d'admission

L'algorithme de *Q-learning* permettra de déterminer le choix optimal de modification du seuil de charge cible DL pour la règle R_1 (par exemple) du SIF décrit sur la figure 4.8.

4.2.5.2 Introduction d'une fonction de vérité pour chaque règle élémentaire du SIF

Comme indiqué dans [Glo97], nous introduisons pour chaque règle élémentaire du SIF, une fonction $s \to \alpha_i(s)$ qui représente la valeur de vérité de la règle R_i pour le vecteur d'entrée (ou état) s. Cette fonction de vérité pour chaque règle R_i est définie à partir des degrés non nuls d'appartenance aux différents labels d'entrées de la règle [Jou98] :

$$\alpha_{R_i}(s) = T\left(\mu_{L_1^i}(s_1), \mu_{L_2^i}(s_2), ..., \mu_{L_{Ns}^i}(s_{Ns})\right) \tag{4.42}$$

Où la *T-norme* peut être définie par l'opérateur produit (en prenant seulement en compte les degrés non nuls des labels d'entrées de la règle) :

$$\alpha_{R_i}(s) = \prod_{j=1}^{Ns} \mu_{L_j^i}(s_j) \tag{4.43}$$

ou par l'opérateur minimum :

$$\alpha_{R_i}(s) = \underset{j=1}{\overset{Ns}{Min}}\left(\mu_{L_j^i}(s_j)\right) \tag{4.44}$$

Définition d'un prototype de la règle *i*

A partir de la fonction de vérité introduite ci-dessus, nous allons définir la notion de prototype P_i des règles élémentaires R_i.

Définition: un vecteur d'entrée P_i de l'ensemble des états S est appelé un prototype de la règle R_i si et seulement si il vérifie :

$$\begin{aligned} &\alpha_{R_i}(P_i) = 1 \\ &\alpha_{R_i}(P_j) = 0 \ si \ j \neq i \end{aligned} \tag{4.45}$$

Conséquences sur les variables de sortie du SIF

Les variables de sortie du SIF peuvent donc être exprimées comme des fonctions des valeurs de vérité des règles du SIF activées au cours du processus. Cette modélisation simplifiée de *Takagi-Sugeno* ([Glo97]) sera développée au paragraphe suivant.

$$\forall m \in \{1,..., No\}, \quad y_m(s) = \frac{\sum_{R_i} \alpha_{R_i}(s) \times o_m^i}{\sum_{i=1}^{N} \alpha_{R_i}(s)} \tag{4.46}$$

En particulier cette expression permet d'estimer l'action globale inférée dans l'état *s* quand le contrôleur est modélisé par un système d'inférence floue.

4.2.5.3 Cas du SIF simplifié de Takagi-Sugeno

Dans le modèle de *Takagi-Sugeno* [Yan01], l'action *a* inférée par le système dans l'état *s* est égale au centre de gravité des valeurs numériques $\left(o^i\right)_{i=1}^{N}$ des labels de sortie des règles*, pondérées par les valeurs de vérité des règles dans l'état s :

$$a(s) = \frac{\sum_{i=1}^{N} \alpha_{R_i}(s) \times o^i}{\sum_{i=1}^{N} \alpha_{R_i}(s)} \tag{4.47}$$

et la valeur correspondante de la fonction de qualité pour l'action inférée globale *a* est donnée par l'expression suivante :

* N est le nombre de règles du SIF.

$$Q(s,a) = \frac{\sum_{i=1}^{N} \alpha_{R_i}(s) \times q_i(s,o^i)}{\sum_{i=1}^{N} \alpha_{R_i}(s)} \tag{4.48}$$

où $q_i(s,o^i)$ est la fonction de qualité élémentaire pour la variable de sortie o^i de la règle i. Dans la nouvelle version du *Q-learning*, il s'agira de mettre à jour les fonctions de qualité élémentaires des règles activées à chaque itération.

Adaptation de l'algorithme du Q-learning à un SIF

A l'initialisation, les fonctions de qualité élémentaires sont toutes prises égales à zéro et n'ont donc pas des valeurs significatives au cours des premières étapes du processus d'apprentissage. Comme décrit au paragraphe 4.2.4.2, les actions sont sélectionnées en cours d'apprentissage à l'aide de la politique d'Exploration / Exploitation (EEP) (classiquement la politique pseudo stochastique). Soient o^i_{select}, la variable de sortie sélectionnée pour la règle i en appliquant la politique EEP et $o^i_{\max}$, la variable de sortie telle que :

$$o^i_{\max} = \arg\max_{o \in O^i(s)} (q_i(s,o)) \tag{4.49}$$

où $O^i(s)$ est l'ensemble des actions possibles dans l'état s pour la règle i. A partir de l'équation précédente (4.48) et du paramètre o^i_{select}, nous déduisons l'expression suivante de la fonction de qualité dans l'état s et pour l'action globale inférée a :

$$Q(s,a) = \frac{\sum_{i=1}^{N} \alpha_{R_i}(s) \times q_i(s,o^i_{Select})}{\sum_{i=1}^{N} \alpha_{R_i}(s)} \tag{4.50}$$

On peut déduire de la même manière l'expression de la valeur V de l'état s, en fonction des relations (4.48) et (4.49) :

$$V(s) = \frac{\sum_{i=1}^{N} \alpha_{R_i}(s) \times q_i(s,o^i_{Max})}{\sum_{i=1}^{N} \alpha_{R_i}(s)} \tag{4.51}$$

A partir de ces expressions, nous pouvons déduire l'adaptation de l'algorithme de *Q-learning* appliqué à un système d'inférence floue.

(1) Initialiser les fonctions de qualité élémentaires $q_i(s,o) = 0$ pour tout les règles R_i, $\forall s \in S, \forall o \in O^i$.
(2) $\gamma \in [0,1[$, $l \in [0,1]$, $\varepsilon \in [0,1]$
(3) Choisir l'état initial s
(4) Répéter jusqu'à la convergence :
(5) Pour chaque règle R_i: déterminer la sortie sélectionnée o^i_{select} par la politique d'exploration / exploitation EEP.
(6) Calculer l'action inférée *a(s)* et la valeur de qualité correspondante *Q(s,a)*
(7) Réaliser l'action a, qui mène au nouvel état s', avec la valeur de renforcement r

(8) Calculer la valeur du nouvel état *V(s')*
(9) Mettre à jour les fonctions de qualité élémentaires pour chaque règle R_i en utilisant les valeurs *Q(s,a)*, *r* et *V(s')*.
(10) Si s' est un état terminal alors choisir un autre état initial s
(11) Sinon $s \leftarrow s'$
(12) Fin de la boucle

Comme expliqué dans [Glo97] et [Jou98], la différence entre les valeurs de la fonction de qualité à l'itération t et à l'itération t+1, $Q_{t+1}(s,a)$ et $Q_t(s,a)$, peut être considérée comme une erreur temporelle, définie comme :

$$\Delta Q = r + \gamma \times V(s') - Q(s,a) \tag{4.52}$$

Nous pouvons utiliser ce signal d'erreur ΔQ pour mettre à jour les fonctions de qualité élémentaires à l'étape (9) de l'algorithme de *Q-learning* adapté à la logique floue. Par exemple, avec la méthode de descente du gradient, on peut exprimer la mise à jour de la fonction de qualité élémentaire par [Glo97] :

$$\Delta q_i\left[s, o^i_{select}\right] = l \times \Delta Q \times \frac{\alpha_{R_i}(s)}{\sum_{i=1}^{N} \alpha_{R_i}(s)} \tag{4.53}$$

où le paramètre *l* est le taux d'apprentissage. Il est aussi possible de combiner le *Q-learning* adapté à la logique floue et la méthode des différences temporelles pour mettre à jour les fonctions de qualité à l'étape (9) du *Q-learning*.

4.2.5.4 Fonction de fiabilité d'un FIS

Nous définissons maintenant l'activation maximale par règle du SIF, sur l'ensemble S_K des K premiers états du réseau par :

$$\alpha^{Max}_{R_i}(K) = \max_{s \in S_K}\left(\alpha_{R_i}(s)\right) \tag{4.54}$$

La variable $\alpha^{Max}_{R_i}$ estime quantitativement dans quelle mesure les K premiers vecteurs d'état du réseau se sont rapproché au maximum d'un prototype de la règle *i*. Nous définissons à présent l'activation moyenne par règle du SIF (en ne tenant compte que des cas où cette activation est supérieure à un certain seuil α_{Min}), sur l'ensemble S_K des K premiers états du réseau par :

$$\alpha^{Moy}_{R_i}(K) = \frac{\sum_{s \in S_K}\left(\alpha_{R_i}(s) \times 1_{\alpha_{R_i}(s) > \alpha_{Min}}\right)}{\sum_{s \in S_K} 1_{\alpha_{R_i}(s) > \alpha_{Min}}} \tag{4.55}$$

La variable $\alpha^{Moy}_{R_i}$ estime quantitativement dans quelle mesure les K premiers vecteurs d'état du réseau se sont rapproché en moyenne d'un prototype de la règle *i*. A partir de ces variables $\alpha^{Moy}_{R_i}$, il est possible de construire un indicateur sur le degré de confiance après K itérations de l'algorithme de *Q-learning* que l'on peut accorder à l'action choisie par un SIF en fonction d'un vecteur d'état *s* :

$$conf_K(s) = \sum_{i=1}^{N_R} \alpha^{Moy}_{R_i}(K) \times \alpha_{R_i}(s) \tag{4.56}$$

Si cet indicateur est faible, cela signifie que le vecteur d'état *s* appartient à une zone peu explorée de l'espace d'entrée : dans ce cas le système d'inférence floue est considéré comme peu fiable. On pourra donc calculer à chaque itération de l'algorithme la valeur de cet indicateur pour l'état courant. Et à partir d'un certain seuil, on pourra considérer que l'algorithme a suffisamment exploré les états possibles du réseau pour le niveau de charge étudié.

4.2.6. Mise en œuvre du Q-learning dans le cas de l'optimisation des CLF

Dans ce paragraphe nous décrivons la mise en œuvre de l'algorithme de *Q-learning* dans le cadre de l'optimisation des contrôleurs de logique floue, gérant de manière automatique le paramétrage des algorithmes RRM au niveau des nodes B.

4.2.6.1 Introduction de connaissances dans la phase d'exploration

Comme cela a été mentionné au chapitre précédent, il est possible d'introduire dans les algorithmes d'apprentissage par renforcement plusieurs niveaux de connaissance a priori. Ces connaissances seront utilisées dans la phase d'exploration de l'algorithme, afin de guider et de raccourcir la durée de l'apprentissage par le *Q-learning*.

Connaissance nulle

Aucune connaissance a priori n'est fournie en entrée de l'algorithme. Pour chaque règle d'inférence, on considère que l'ensemble des actions o^i sera possible dans la conséquence de la règle. On définit seulement les bornes des ensembles de définition $\left\{o^i_{Min}, o^i_{Max}\right\}$ de ces sorties et le nombre d'ensembles flous de la partition floue. Dans le cas du SIF défini en début de chapitre, cela signifie que pour chaque règle, on introduit les valeurs minimale et maximale physiquement possibles. On découpe alors cet intervalle en trois ensembles flous correspondant aux 3 corrections possibles pour Δc : *baisser*, *ne rien faire* ou *augmenter* le seuil de charge cible DL. Enfin à chaque label flou O^i_j, on attribue une valeur numérique o^i_j appartenant à l'ensemble de définition correspondant au label.

Connaissance imprécise

Dans ce cas, la connaissance que l'on introduit dans l'algorithme d'apprentissage s'exprime sous la forme d'énoncés qualitatifs. Cela implique que plusieurs labels de sortie sont possibles dans les conséquences des règles, et pour chaque variable de sortie on choisira une valeur numérique dans l'ensemble de définition du label (fourni par la connaissance a priori). Dans la phase d'exploration, l'algorithme déterminera quelle interprétation de la règle qualitative est la plus efficace. Ainsi dans l'exemple de la figure 4.8, on déterminera entre O^i_1, O^i_2 et O^i_3 la sortie ayant la meilleure fonction de qualité élémentaire.

Connaissance précise

Dans ce cas, on énonce une règle déterministe qui fixe la sortie du processus pour certains états précis du réseau. Ce type de règle peut être fondé par exemple sur une expérience forte des ingénieurs de gestion du réseau. Dans les phases d'exploration et d'exploitation, pour les règles impliquées, l'algorithme choisira systématiquement la sortie fixée par la règle.

Dans un processus global d'optimisation par l'algorithme de *Q-learning* on va combiner ces trois niveaux de connaissance a priori. Par exemple, dans le cas du contrôle du seuil de charge cible *downlink* d'admission, on pourra introduire :

- Des connaissances précises pour certains états du système correspondant à des parties des matrices de décisions du moteur d'inférence décrit par la figure 3.19.
- Des connaissances imprécises pour certains états du système correspondant à d'autres parties des matrices de décisions du moteur d'inférence décrit par la figure 3.19.
- Aucune connaissance dans les autres cas.

Par ailleurs, on peut aussi envisager une utilisation sur plusieurs couches de l'algorithme d'apprentissage par renforcement. Partant de règles de connaissances imprécises, on va lancer une première optimisation à l'aide de l'algorithme de *Q-learning*. Cela va permettre d'affiner la connaissance de notre SIF: en excluant certains labels de sortie pour certaines règles, en rétrécissant les intervalles de définition des sorties… Sur le nouveau SIF ainsi défini, on pourra lancer une nouvelle optimisation par l'algorithme de *Q-learning*, et itérer de cette manière plusieurs fois le processus pour affiner la connaissance du système.

4.2.6.2 Choix des paramètres et fonctions de l'algorithme de Q-learning

On énumère ci-dessous les différents paramètres et fonctions à fixer à l'initialisation de l'algorithme du *Q-learning* :

- γ le facteur d'atténuation a des valeurs typiques comprises entre : $0.9 < \gamma < 0.99$.
- Une politique d'exploration / exploitation typique est la politique pseudo-stochastique (ou ε-*greedy*) telle que :

  Pour chaque règle R_i nous sélectionnons soit la sortie o^i_{Best} maximisant la fonction de qualité de la règle R_i dans l'état s avec une probabilité $(1- \varepsilon)$ soit une action aléatoire dans $O^i(s) / \{o^i_{Best}\}$ avec une probabilité ε.

- ε la probabilité pour que la politique d'exploration / exploitation choisisse une action au hasard au lieu de la meilleure action o^i_{Best} a des valeurs typiques telles que : $0.01 < \varepsilon < 0.5$
- l le taux d'apprentissage peut être défini :
  - Comme une constante telle que : $0.001 < l < 0.1$.
  - Par la fonction suivante : $l_t = \dfrac{1}{1 + visit_t(s_t, a_t)}$, (4.57)

    où $visit_t(s_t, a_t)$ est le nombre total de fois où la paire état– action a été visitée. Dans le cas d'un système flou, il est moins direct d'introduire une telle fonction de visite. Nous verrons ci-après comment nous proposons de définir cette fonction.
- La fonction de vérité de chaque règle R_i est définie à l'aide de l'opérateur produit et des fonctions d'appartenance $\mu_{L^i_j}$ correspondant au j-ème label flou L^i_j de la règle R_i :

$$\alpha_{R_i}(s) = \mu_{B_{Centrale}}(s) \times \mu_{R_{Centrale}}(s) \times \mu_{B_{Voic.,Moy}}(s) \times \mu_{R_{Vois.,Moy}}(s) \quad (4.58)$$

4.2.6.3 Schéma d'utilisation du *Q-learning*

Nous considérons le SIF simplifié introduit au paragraphe 4.2.5.1. Nous avons donc défini 16 règles d'inférences élémentaires dont nous allons estimer les fonctions de qualité correspondantes à l'aide de la phase d'exploration de l'algorithme de *Q-learning* adapté à un système flou :

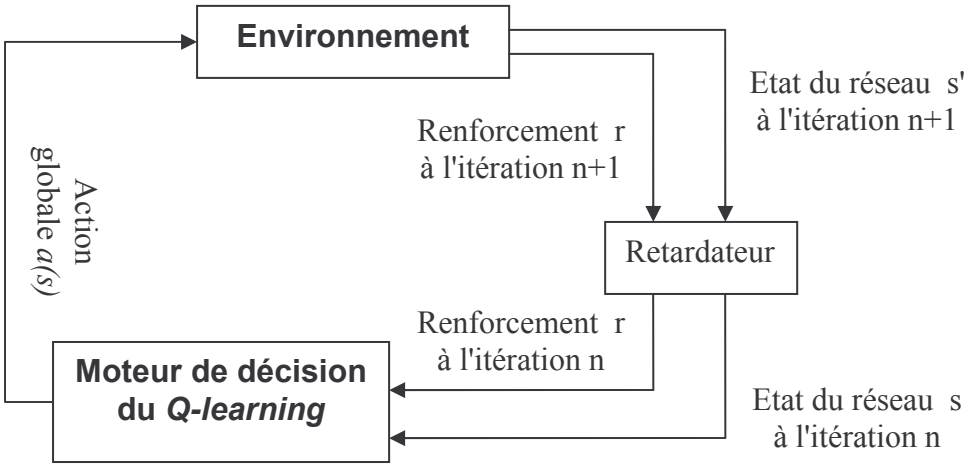


Figure 4.9 : Fonctionnement général de l'algorithme de *Q-learning*

(1) Initialiser $\forall i \in \{1, 2,..., 16\}\ \ q_i(s,o) = 0$ pour tout $s \in S, o \in O^i$

(2) $\gamma \in [0,1[$, choix du taux d'apprentissage $l \in [0,1]$ (voir paragraphe 4.2.6.2), $\varepsilon \in [0,1]$

(3) Choisir un état initial au hasard *s*

(4) Recommencer jusqu'à convergence :

(5) Pour chaque règle R_i telle que $\alpha_{R_i}(s) \neq 0$, déterminer la sortie o^i_{select} sélectionnée en utilisant la politique d'exploration / exploitation suivante : Choisir $o^i_{Best} = \arg\max_{o \in O^i(s)} q_i(s,o)$ avec la probabilité *(1- ε)* ou une action aléatoire dans $O^i(s) / \{o^i_{Best}\}$ avec une probabilité ε.

(6) Calculer l'action inférée globale : $a(s) = \dfrac{\sum_{i=1}^{N} \alpha_{R_i}(s) \times o^i_{Select}}{\sum_{i=1}^{N} \alpha_{R_i}(s)}$ et la fonction de qualité globale correspondante : $Q(s,a) = \dfrac{\sum_{i=1}^{N} \alpha_{R_i}(s) \times q_i(s, o^i_{Select})}{\sum_{i=1}^{N} \alpha_{R_i}(s)}$.

(7) Réaliser l'action *a(s)*, menant au nouvel état *s'*, avec un renforcement *r*.

(8) Calculer la valeur du nouvel état : $V(s') = \dfrac{\sum_{i=1}^{N} \alpha_{R_i}(s') \times q_i(s', o^i_{Max})}{\sum_{i=1}^{N} \alpha_{R_i}(s')}$.

(9) Mettre à jour les valeurs de qualité de chaque règle R_i en utilisant les valeurs globales *Q(s,a)*, *r* et *V(s')* :

$$\Delta Q = r + \gamma \times V(s') - Q(s,a) \tag{4.59}$$

Par exemple avec la méthode de descente du gradient, la variation de la fonction de qualité de la règle R_i s'exprime par :

$$\Delta q_i\left[s, o_{select}^i\right] = l \times \Delta Q \times \frac{\alpha_{R_i}(s)}{\sum_{i=1}^{N} \alpha_{R_i}(s)} \tag{4.60}$$

(10) Si l'état s est terminal alors choisir un autre état initial s
(11) Sinon $s \leftarrow s'$
(12) Fin de la boucle

La phase d'exploitation de l'algorithme de *Q-learning* est identique au processus décrit ci-dessus mis à part :

- La transition de l'exploration à l'exploitation qui peut se faire en cas d'efficacité avérée du contrôleur: 1000 itérations de contrôle automatique sans par exemple obtenir de valeur de renforcement négative ou sans modification significative des fonctions de qualité élémentaires.
- L'état de départ n'est pas aléatoire mais correspond à l'état initial du système à optimiser.
- La politique d'exploration / exploitation choisit systématiquement la sortie o_{Best}^i maximisant la fonction de qualité de la règle R_i dans l'état s.
- Possibilité de relancer le processus d'exploration avec différentes hypothèses de trafic (taux d'arrivée, durée des communications et distribution des mobiles).

Définition de la fonction de visite

Nous introduisons une fonction de visite élémentaire v_i pour chaque règle R_i du SIF. Au départ, l'ensemble des fonctions de visite est initialisé à 0. A chaque itération, quand on sélectionne la sortie o_{select}^i avec la politique d'exploration, on met à jour la fonction de visite de la règle R_i :

$$\Delta v_i\left[s, o_{select}^i\right] = + \frac{\alpha_{Ri_i}(s)}{\sum_{i=1}^{N} \alpha_{Ri_i}(s)} \tag{4.61}$$

A l'étape 9 de l'algorithme de *Q-learning*, on détermine alors le taux d'apprentissage l :

- Soit à l'aide d'une constante, $l = 0.1$ par exemple.
- Soit on estime l à l'aide de la fonction de visite :

$$l = \frac{1}{1 + v_i\left[s, o_{select}^i\right]} \tag{4.62}$$

Les mises à jour des fonctions de qualité élémentaires en début de processus ne sont pas fiables. En revanche après un certain nombre d'itérations, ces valeurs deviennent de plus en plus précises. C'est pourquoi, il vaut mieux utiliser un taux d'apprentissage l constant (et non pas décroissant) pour ne pas favoriser les valeurs des premières itérations (qui ne sont pas fiables) et pénaliser celles des itérations suivantes (qui deviennent de plus en plus fiables).

Définition de la fonction de renforcement

Pour l'application du *Q-learning* au contrôle du seuil de charge cible DL pour l'admission des mobiles, on peut intégrer dans la fonction de renforcement :

- Un terme sur le taux de blocage : taux de blocage mesuré – taux de blocage cible pour la cellule (par exemple 5% dans les réseaux UMTS).
- Un terme sur le taux de coupure : taux de coupure mesuré – taux de coupure cible pour la cellule (par exemple 2% dans les réseaux UMTS).
- La fonction de renforcement pourra être éventuellement adaptée en fonction des données de trafic (en ajustant les taux de blocage et de coupure cibles par exemple).

De ces considérations, nous proposons une première fonction de renforcement à tester dans le cadre d'apprentissage sur des simulations de réseaux mobiles (le coefficient de mixage ayant une valeur typique entre 3 et 10) :

$$R = \left(b_{cible} - b_{Cellule}\right) + Coeff \times \left(d_{cible} - d_{Cellule}\right) \tag{4.63}$$

Déduction de règles linguistiques après optimisation

A condition que les fonctions d'appartenance respectent la *contrainte forte de partition*, il est possible d'interpréter les contrôleurs optimisés obtenus en fin de processus d'exploration du *Q-learning* comme de nouvelles règles linguistiques que l'ingénieur de gestion de réseau pourra ensuite tester et mettre en oeuvre dans un cadre opérationnel.

4.2.7. Possibilités d'extensions de l'algorithme de Q-learning

4.2.7.1 Exemple d'utilisation itérative du Q-learning pour affiner la connaissance du système

Dans le cas du SIF introduit au paragraphe 4.2.5.1, nous supposons à présent que les ensembles de définition correspondant à chacun des trois labels pour la correction Δc sont respectivement: *[-0.45, -0.15]*, *[-0.15, 0.15]*, *[0.15, 0.45]*. Pour chaque règle et chaque label, on choisit une valeur numérique dans l'ensemble de définition correspondant au label. Par exemple pour la règle *1*, on peut prendre les valeurs numériques correspondant aux labels O_1^1, O_2^1 et O_3^1 égales respectivement à *-0.3*, *0* et *+0.3*.

Une première optimisation (en mode d'exploration) par l'algorithme de *Q-learning* permet d'apprendre les valeurs des fonctions de qualité élémentaires. Nous précisons que, lors de l'étape *5* de l'algorithme, la valeur numérique de la sortie o_{select}^i sélectionnée est égale à la valeur choisie à l'initialisation du processus pour la règle *i*. Nous supposons par exemple que la politique optimale en fin de processus sélectionne systématiquement la correction *baisser* pour la règle *1*, et *augmenter* pour la règle *2*. Alors pour affiner la connaissance du système, on peut diviser l'intervalle de définition pour la règle *1* correspondant au label *baisser* en trois nouveaux sous-intervalles: *[-0.45, -0.35]*, *[-0.35, -0.25]*, *[-0.25, -0.15]*. Nous initialisons les nouvelles fonctions de qualité élémentaires à zéro, ces fonctions de qualité correspondant aux trois nouveaux labels flous de sortie pour la règle *1*: *baisser beaucoup*, *baisser moyennement*, *baisser peu*. Pour la règle *2*, nous découpons de la même façon l'intervalle de définition du label *augmenter* et nous attribuons un intervalle par label *augmenter peu*, *augmenter moyennement* et *augmenter beaucoup*. Ainsi pour chaque règle, nous échantillonnons de cette manière l'intervalle de définition de la sortie sélectionnée par la politique gloutonne après apprentissage. Alors une valeur numérique est affectée à chaque label de sortie et l'algorithme de *Q-learning* est relancé en mode d'exploration sur le nouveau SIF ainsi défini.

Une telle utilisation en série de l'algorithme de *Q-learning* va permettre d'affiner la précision de la sortie de contrôle. De la même façon, on peut envisager un système d'apprentissage

utilisant en série plusieurs fois l'algorithme de *Q-learning* et intégrant au fur et à mesure de nouveaux indicateurs de performances.

4.2.7.2 Introduction de nouveaux indicateurs de performances

Dans le cas du SIF introduit au paragraphe 4.2.5.1, nous supposons à présent que le contrôleur ait déjà été optimisé une première fois à l'aide de l'algorithme du *Q-learning* adapté à la logique floue. Mais, on souhaite introduire dans ce SIF contrôlant le seuil de charge cible *downlink* un nouvel indicateur de performances, en plus des taux de blocage et de coupure pour la cellule centrale et ses voisines : par exemple la charge *downlink* de la station centrale.

Apprentissage en série à l'aide du Q-learning adapté aux SIF

On introduit deux labels flous possibles pour la charge *downlink* ρ_{DL} de la station: faible et fort. Par rapport au SIF décrit sur la figure 4.8, une nouvelle variable d'entrée a donc été ajoutée ; la charge DL de la cellule, deux labels flous pour cette variable ont été définis, et le nombre de règles de décisions a été multiplié par deux. Les nouveaux ensembles de définition des variables de sortie des différentes règles, peuvent être centrés sur la sortie fournie par la politique optimale déterminée par l'optimisation précédente. Comme décrit sur la figure 4.10, on peut ainsi introduire un niveau de connaissance précis (un seul label de sortie pour la règle correspondant à une valeur numérique), imprécis (plusieurs labels de sortie correspondant chacun à un ensemble de définition dans lequel on choisit une valeur numérique) ou aucune connaissance (en revenant aux intervalles de définition initiaux). Avant de lancer l'algorithme de *Q-learning* en mode exploratoire, nous initialisons les valeurs des nouvelles fonctions de qualité à zéro (au plus 96 fonctions de qualité avec un maximum de 3 sorties par règles).

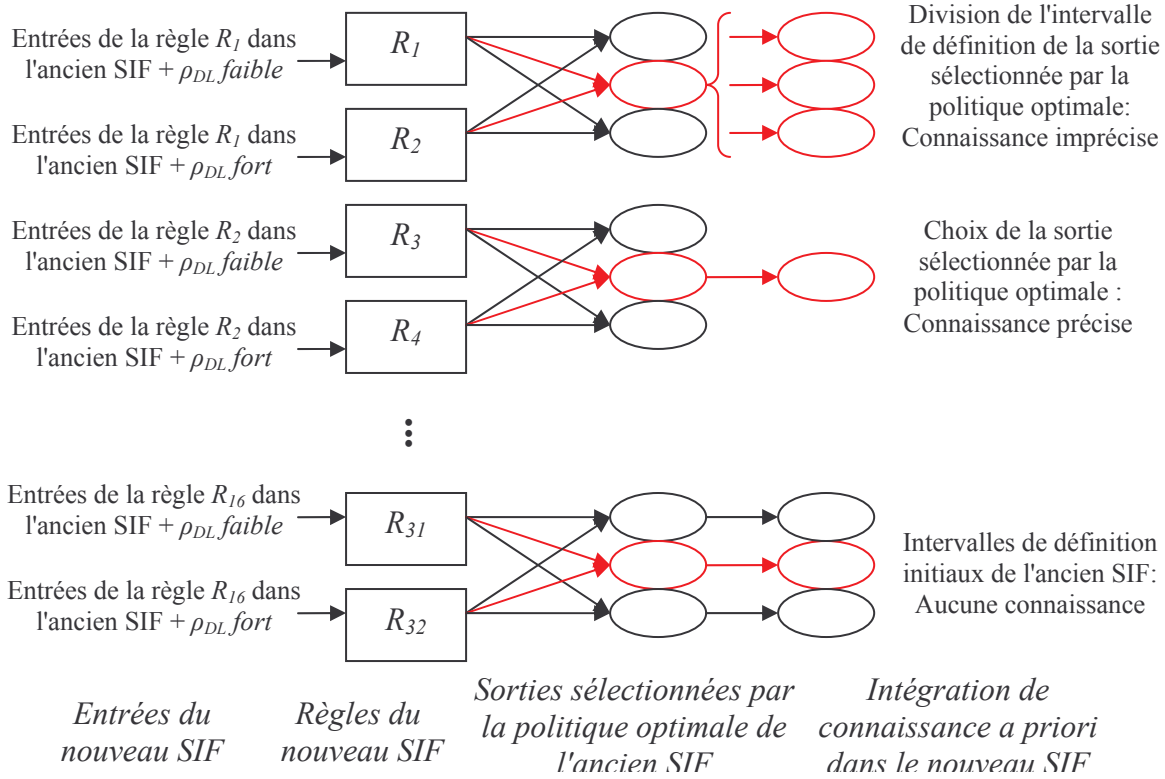


Figure 4.10 : Définition d'un SIF prenant en compte un indicateur de performance supplémentaire à partir d'un SIF optimisé avec le *Q-learning*.

Apres cette initialisation, l'algorithme de *Q-learning* est appliqué en mode exploratoire au nouveau SIF défini ci-dessus. On affine de cette manière la connaissance du système, tout en intégrant la charge DL par cellule parmi les indicateurs de performances. Les systèmes d'inférence flous permettent ainsi de transmettre assez facilement la connaissance déjà apprise. Cette caractéristique est très importante, en particulier quand on envisage d'incorporer peu à peu à notre processus de contrôle différents indicateurs de performances, différents algorithmes RRM ou d'élargir son utilisation à un fonctionnement de type multi-systèmes : à chaque introduction d'un nouveau système, il sera ainsi possible de tenir compte des apprentissages antérieurs effectués pour les autres systèmes.

4.2.8. Exploitation de systèmes d'inférence floue avec le Q-learning

Il est préférable de définir un système d'inférence floue par variable à contrôler. Par exemple, pour le couple de variables de modification du seuil de charge cible DL (paramètre pour le contrôle d'admission) et de la taille de la fenêtre d'ajout de liens de MD (paramètre de macro-diversité), il est défini un couple de SIF correspondant. Il est possible :

- Soit d'apprendre parallèlement pour chacun des SIF introduits : à chaque itération, on produit une correction de paramétrage par SIF et on met à jour simultanément toutes les fonctions de qualité impliquées de chaque SIF.
- Soit d'apprendre séquentiellement pour chacun des SIF : à chaque itération, un seul SIF est actif qui fournit la correction de paramétrage à appliquer. Les fonctions de qualité impliquées du SIF seront mises à jour à l'itération suivante. On change de SIF actif à chaque itération.
- Soit d'apprendre successivement pour chacun des SIF : on lance le processus total d'apprentissage pour un seul SIF, les autres paramètres du réseau étant figés. Une fois le processus d'exploration terminé, on choisit un nouveau SIF actif à optimiser et on recommence.

SIF global / SIF local

Dans un premier temps, on pourra définir un seul SIF global qui apprendra sur l'ensemble des cellules d'une zone homogène du réseau (homogène au niveau de l'environnement radio et de la distance inter-site, mais pas pour l'objet de trafic). Une fois ce contrôleur optimisé globalement sur la zone (calibrage global), on pourra définir node B par node B, un SIF local. Ce SIF local sera initialisé grâce à la connaissance acquise sur le SIF global. Puis on pourra relancer un processus d'apprentissage par *Q-learning*, où cette fois on mettra à jour les fonctions de qualité élémentaires de chaque SIF défini pour chaque cellule.

Apprentissage parallèle avec différents trafics

On définit différents objets de trafic (en termes de distribution et niveau de trafic). Et le même SIF va apprendre soit séquentiellement (au cours d'un seul processus d'apprentissage) soit successivement (plusieurs processus d'apprentissage consécutifs) à partir des situations fournies par les différents objets de trafic. Cette diversité des hypothèses de trafic permet d'explorer plus largement les états possibles des cellules du réseau dont on veut contrôler le paramétrage). Pour terminer le processus, la condition d'arrêt de la phase d'exploration devra être remplie pour chaque objet de trafic pris en compte.

4.3. Conclusions sur les méthodes d'optimisation

Nous avons donc présenté deux méthodes pour améliorer les performances du contrôle automatique des paramètres du réseau UMTS. La première, l'optimisation par essaim de particules, est une méthode d'optimisation itérative stochastique qui vise à améliorer l'efficacité des contrôleurs de logique floue introduits au chapitre 3. Elle permet donc d'explorer semi aléatoirement l'espace de définition des contrôleurs, en comparant leurs performances à l'aide de fonctions coûts. Deux modes de comparaison sont possibles : le mono-objectif ou le multi-objectif. Dans le premier cas, un seul indicateur synthétise les performances en termes de couverture et de capacité sur l'ensemble des cellules du réseau. Dans le second cas, on calcule un indicateur sur l'ensemble des cellules pour la couverture et un second pour la capacité, et on compare les performances des CLF dans le plan taux de coupure – taux de blocage. Les fonctions coûts permettent donc de mesurer l'efficacité globale des CLF, soit en moyennant les indicateurs sur l'ensemble des cellules soit en privilégiant certaines cellules (les pires stations du réseau par exemple). L'OEP a été adaptée aux deux modes de fonctionnement envisagés pour les CLF (chapitre 5) : le mode *on-line* et le mode *off-line*. Dans ces deux cas, nous avons obtenu des améliorations notables des performances des CLF (chapitres 6 et 7). Plusieurs caractéristiques des OEP nous ont cependant contraint à tester et étudier d'autres méthodes d'optimisation dans le contexte du contrôle automatique des paramètres du réseau UMTS :

- Il y a une perte d'information lors du calcul de la moyenne pondérée des données par cellule pour évaluer les fonctions coûts. De plus en mono-objectif, on fusionne les indicateurs de capacité et couverture, à l'aide d'un coefficient de mixage, ce qui signifie un manque de précision supplémentaire dans le processus d'optimisation. On peut dire que l'OEP prend en compte les performances globales sur la zone considérée du réseau, mais pas la situation cellule par cellule.
- L'application de l'OEP nécessite de comparer l'évolution de plusieurs dizaines de particules représentant chacune un paramétrage différent du contrôleur. En conséquence, pour pouvoir évaluer la position de chaque particule, on doit tester l'efficacité du contrôleur correspondant sur un réseau mobile de plusieurs dizaines de stations. L'utilisation de l'OEP n'est donc envisageable qu'en estimant les performances des contrôleurs à l'aide de simulations logicielles de réseau mobile. Ainsi on calibrera un CLF a priori efficace pour le mode *on-line* ou un jeu de paramètres a priori optimisés pour le mode *off-line*, que l'on testera ensuite sur un réseau expérimental voire un réseau d'exploitation.
- Le processus d'OEP est donc inadapté à une implémentation directe dans les réseaux mobiles réels.
- Le processus d'OEP n'est pas complètement générique. Au fur et à mesure qu'on prendra en compte de nouveaux indicateurs de performances, de nouveaux algorithmes RRM ou de nouveaux systèmes mobiles (dans une approche multi-système), il faudra adapter les contrôleurs. La phase d'optimisation devra ensuite être relancée, même si une certaine connaissance a priori pourra être intégrée au processus. Cette connaissance sera introduite soit au niveau des particules initiales soit en figeant certaines matrices des CLF.

Du fait des limites de l'approche par OEP, nous avons donc étudié la possibilité d'utiliser des méthodes d'apprentissage par renforcement pour contrôler le paramétrage des réseaux mobiles. Pour être le plus général et générique possible, nous avons choisi de revenir à la description de ce problème d'optimisation par le système d'inférence floue introduit au chapitre 3. Parmi les différents algorithmes d'apprentissage par renforcement, nous avons choisi d'adapter à

notre SIF, une méthode de détermination de politique déterministe optimale; le *Q-learning*. L'utilisation du SIF permet de ramener notre problème d'optimisation sur des ensembles continus à un processus de décisions Markovien. Plusieurs propriétés du *Q-learning* nous permettent d'espérer que cette méthode devrait être plus appropriée à notre problème :

- Il est possible d'apprendre simultanément et parallèlement de toutes les stations du réseau: à chaque itération, on met à jour les fonctions de qualité du contrôleur en utilisant les informations (état antérieur, correction appliquée, état présent, renforcement…) de chaque cellule.
- Utilisation directe des contrôleurs (calibrés en simulations) dans les réseaux réels quelles que soient les contraintes temporelles de remontées de mesures ou les délais d'applications des corrections de paramétrage, en mode exploitation (procédé normal) voire en mode exploration (pour permettre l'adaptation du contrôleur à un environnement local particulier) de la politique EEP.
- La méthode de contrôle par *Q-learning* est générique : introduction simple de nouveaux indicateurs de performances dans un contrôleur (avec la possibilité d'intégrer la connaissance acquise auparavant), de nouveaux algorithmes à contrôler, de nouveaux systèmes à prendre à compte…

En revanche, quelques phases du *Q-learning* sont délicates à implémenter :

- Le processus d'optimisation est très sensible au choix de la fonction de renforcement.
- Pour que les méthodes d'apprentissage par renforcement soient efficaces, il est nécessaire de visiter suffisamment de fois les différents états du système (voir les conditions de convergence de l'algorithme au 4.2.3.2). Donc, il pourra être nécessaire de lancer plusieurs fois le processus d'apprentissage avec différents niveaux et distributions de trafic.
- Choix délicat de la méthode de mise à jour des fonctions de qualité élémentaires à partir de la fonction de qualité globale : méthode de la descente du gradient…

Aux chapitres 6 et 7, nous introduirons les résultats les plus significatifs obtenus respectivement avec l'optimisation par essaim de particules et l'algorithme de *Q-learning*.

5. Fonctionnement en modes *off-line* / *on-line*

Deux types d'utilisation du processus de paramétrage automatique sont prévus; les modes *on-line* et *off-line*. Les objectifs pour les contrôleurs ne sont pas exactement les mêmes selon l'utilisation prévue :

- Dans le cas d'un fonctionnement *on-line*, il s'agira pour les contrôleurs d'adapter en permanence le réseau aux fluctuations de trafic en re-paramétrant dynamiquement les algorithmes RRM.
- Dans le cas d'un fonctionnement *off-line*, l'objectif des contrôleurs sera un re-paramétrage RRM de tout le réseau, node B par node B, pour une distribution de trafic donnée (en particulier dans le cas d'un trafic inhomogène stable).

Comme décrit au chapitre 3, le mode *on-line* est fondé a priori sur l'utilisation de contrôleurs de logique floue (CLF) intégrés à des agents distribués dans le réseau. Ce mode de fonctionnement permettra ainsi d'ajuster dynamiquement certains paramètres contrôlant des algorithmes de RRM tels que le seuil de charge cible *downlink* du contrôle d'admission ou les fenêtres d'ajout / de retrait de la macro-diversité, afin d'améliorer la qualité de service et les performances du réseau mobile et de l'adapter aux fluctuations de trafic. Le processus d'ingénierie dynamique est basé sur une boucle de régulation qui est dirigée par le contrôleur de logique floue. Ce contrôleur reçoit dynamiquement en entrées les indicateurs de qualité de la cellule et de ses proches voisines, et produit en sortie un jeu de corrections des paramètres RRM de la station centrale. On envisage deux stratégies de modifications des paramètres RRM, à chaque itération : a) modifier simultanément toutes les stations du réseau, b) modifier une seule station par groupe d'interférence de stations (ou *clusters*). Dans ce chapitre, nous présenterons une méthode systématique, permettant de partitionner le réseau en groupes d'interférence disjoints, par agrégation (ou *clustering*) de sous-groupes en utilisant une fonction distance entre stations définie à partir du niveau d'interférence mutuelle (cf. [Ben05a]).

Le deuxième type de fonctionnement de l'ingénierie automatique – le mode *off-line* (hors-ligne) – permettra le réglage local (au niveau du node B) des paramètres RRM. En effet, à l'aide d'un simulateur, semi-dynamique ou dynamique, de réseau UMTS et d'un module logiciel de paramétrage des stations, il sera possible de déterminer les modifications du paramétrage RRM les plus efficaces pour une demande en trafic donnée sur la zone géographique considérée (homogène comme inhomogène). On pourra ensuite vérifier sur le terrain l'efficacité des corrections effectuées en termes de qualité de service perçue et de performances du réseau (capacité, couverture…). Pour cela, les métriques radio avant et après la modification des paramètres RRM pour des périodes correspondant à des niveaux et types de trafic équivalents seront comparées. Cela constitue une boucle de validation de l'approche *off-line*. Avec le mode *off-line*, il sera aussi possible de déterminer (par des simulations validées ensuite par des tests) et d'utiliser plusieurs jeux de paramètres RRM typiques, node B par node B. Ces jeux de paramètres pourront dépendre du jour de la semaine, de la période de l'année (vacances), voire d'évènements exceptionnels mais récurrents (salons, grandes manifestations musicales ou sportives…).

Dans le cas des contrôleurs de logique floue utilisant l'apprentissage par renforcement, nous avons imaginé un mode de fonctionnement intermédiaire – "semi *on-line*". Pour cela, nous supposons que nous avons calibré (phase d'exploration) à l'aide de simulations logicielles un contrôleur fondé sur le *Q-learning* comme décrit au chapitre 4. Un contrôleur de ce type est alors introduit au niveau de l'OMC pour administrer un ensemble de cellules. La fréquence

des remontées de mesures ainsi que celle des corrections des paramètres RRM vont dépendre des contraintes liées aux équipements du réseau, mais aussi de la politique de l'opérateur (demi-heure, heure, demi-journée, journée). Dans le cas d'un fonctionnement "semi *on-line*", le *Q-learning* n'est plus utilisé en mode d'exploitation dans le réseau réel, mais en mode d'apprentissage. A chaque itération, le SIF va recevoir pour l'ensemble des nodes B le renforcement correspondant à l'état de la cellule et la correction appliquée (à l'itération précédente). Ainsi, le module pourra mettre à jour les fonctions de qualité correspondant aux situations des différents nodes B. Cette méthode permettra d'adapter à un environnement réel des contrôleurs calibrés seulement à partir de simulations logicielles.

Enfin dans chaque mode de fonctionnement, toute une réflexion devra être menée sur les différentes échelles de temps prises en compte par le processus : granularité des indicateurs de performances, taille des fenêtres de filtrage, intervalle inter-modifications des paramètres RRM, temps de réponse du réseau à une modification de paramétrage…

5.1. Le mode on-line en UMTS

5.1.1. Contraintes liées au mode *on-line*

Les réseaux UMTS sont des systèmes limités par l'interférence du fait de l'utilisation de la même fréquence par tous les utilisateurs. L'interférence intracellulaire est réduite grâce à l'orthogonalité des séquences de codes assignées dans une même cellule. En revanche, l'interférence intercellulaire est seulement protégée par l'affaiblissement de propagation du signal radio. Or, il est possible qu'une station éloignée interfère beaucoup la station considérée à cause de conditions radio défavorables (résurgence radio). Le partage de la même fréquence produit donc des phénomènes de couplage entre les nodes B qui vont dépendre de l'environnement radio, de la distribution de trafic et du paramétrage des stations.

Dans ce contexte, implémenter dans le réseau UMTS des contrôleurs dynamiques qui modifient en permanence et simultanément le paramétrage de toutes les stations du réseau apparaît comme risqué (effets ping-pong, couplage entre les voisines). En conséquence, développer une solution de paramétrage *on-line* nécessite de faire un compromis entre la réactivité du processus de contrôle et la réduction des effets de couplage liés à ce processus. D'une part, on est tenté de paramétrer dynamiquement à chaque itération le plus grand nombre possible de stations. Mais d'autre part, on souhaite que deux stations modifiées soient suffisamment isolées l'une de l'autre pour limiter les effets de couplage.

En conséquence, la solution est de construire des ensembles de cellules suffisamment "distants" les uns des autres, et de n'autoriser qu'une seule modification de paramétrage par ensemble. Pour construire ces groupes, on pourrait être tenté d'utiliser la distance géographique, mais celle-ci n'est pas forcément représentative des phénomènes de propagation: une cellule centrale peut être bien isolée de certaines de ses voisines géographiques (relief, canal multi-trajets) mais être par exemple brouillée par une cellule lointaine mal orientée. Il apparaît donc que le niveau d'interférence (le brouillage) est un bon indicateur de l'environnement radio; permettant en particulier d'estimer le degré de couplage entre deux nodes B. Nous allons donc introduire une distance entre stations définie à partir de l'interférence qu'elles génèrent l'une sur l'autre. Nous adapterons ensuite la méthode du *clustering* (avec différents critères d'agrégation) qui permettra de construire des groupes de stations disjoints et partiellement isolés en terme de brouillage : cette approche vise à rassembler dans le même groupe les stations qui ont une distance d'interférence avec les autres stations du groupe inférieure à la distance avec toutes les autres stations du réseau.

Nous appelons ces groupes (ou classes) de stations, les ensembles d'influence (IS ; *influence sets*).

5.1.1.1 Définition d'une distance entre stations

Intuitivement, une définition pertinente de distance entre deux stations d'un réseau UMTS (c'est-à-dire fondé sur la technologie W-CDMA qui est limitée par l'interférence), pourra se fonder sur la notion de couplage mutuel ou d'interférence mutuelle : dans un tel modèle deux stations seront considérées comme proches si elles interfèrent fortement l'une sur l'autre. Nous allons donc commencer par définir la matrice d'interférence d'un réseau UMTS.

Définition de la matrice d'interférence ([Ben04] [Ben05a] [Ben05b]*)*

Soit un réseau UMTS avec N stations. La matrice d'interférence K est une matrice carrée $N \times N$, qui décrit l'interaction radio entre toutes les paires de stations du réseau. L'élément K_{ij} indique le niveau moyen auquel la cellule j interfère la cellule i du réseau, les frontières des cellules étant définies par exemple à l'aide d'un critère de meilleure serveuse. On peut définir une matrice d'interférence sur les sens UL ou DL :

- sur le sens UL, l'élément K_{ij} est égal à la somme de toutes les puissances reçues au niveau de la station i sur l'UL émises par les mobiles servis par la station j.
- sur le sens DL, l'élément K_{ij} est égal à la moyenne, sur l'ensemble des mailles de la station i, de la puissance reçue en la maille considérée en provenance de la station j.

Les matrices ainsi définies, aussi bien sur le sens UL que sur le sens DL, ne sont donc pas symétriques. On va donc introduire une nouvelle matrice d'interférence $\widetilde{K}$ symétrique :

$$\widetilde{K}_{ij} = \frac{\left(K_{ij} + K_{ji}\right)}{2} \tag{5.1}$$

Cette matrice $\widetilde{K}$ va nous permettre de définir une distance *d(i,j)* sur l'ensemble des stations, satisfaisant les deux propriétés suivantes :

- La distance d'une station à elle-même est nulle : $d(i,i) = 0$.
- La distance est commutative : pour toutes stations i et j du réseau, $d(i,j) = d(j,i)$.

En revanche, nous n'imposons pas la propriété de l'inégalité du triangle. En effet dans les réseaux mobiles du fait des contraintes de propagation, deux stations i et j peuvent interférer très peu l'une avec l'autre, alors qu'elles peuvent chacune avoir un fort couplage avec une troisième station k. Dans ce cas, on pourrait avoir :

$$d(i,j) > d(i,k) + d(k,j) \tag{5.2}$$

Définition d'une distance entre stations

Nous rappelons que l'interférence entre deux stations i et j est une valeur positive. Du fait de la sensibilité limitée des appareils de mesure le terrain et d'un seuil de coupure pour la *path loss* estimé dans les simulations logicielles (par exemple, seuil de coupure à 300 dB), il est possible que certains termes de la matrice $\widetilde{K}_{ij}$ soient nuls. On fixe alors ces termes nuls à une valeur minimale, inférieure ou égale au plus petit élément de la matrice et on définit la distance entre deux stations, à l'aide de cette nouvelle matrice $\widetilde{K}^{P}$ dont les éléments non diagonaux sont tous strictement positifs :

$$d(i,j) = \begin{cases} 0 & si \quad i = j \\ \dfrac{1}{\widetilde{K}_{ij}^{P}} & \text{sin}\, on \end{cases} \tag{5.3}$$

D'autres définitions de distance sont possibles à conditions qu'elles respectent les deux propriétés décrites ci-dessus. La seconde contrainte sur la distance est qu'elle doit être représentative d'un certain couplage entre les stations. Par exemple, nous avons étudié deux autres distances entre stations: la distance *Euclidienne* et la distance de *Mahalanobis* :

- La distance *Euclidienne* se définit à partir d'un vecteur de mesures sur une fenêtre temporelle ($\overline{P}$ étant le vecteur des mesures entre les temps t et $t+T$ de la puissance émise par la station ou de l'interférence totale reçue) : $d(i,j) = \left[{}^{t}\left(\overline{P_i} - \overline{P_j}\right)\left(\overline{P_i} - \overline{P_j}\right)\right]^{1/2}$.
- La distance de *Mahalanobis* se définit à partir d'un vecteur de mesures sur une fenêtre temporelle et une matrice M de taille T x T d'auto-covariance des mesures, chaque terme M_{kl} estimant la corrélation entre les temps k et l : $d(i,j) = {}^{t}\left(\overline{P_i} - \overline{P_j}\right)\left[M\right]^{-1}\left(\overline{P_i} - \overline{P_j}\right)$.

5.1.1.2 Méthode de construction des groupes de stations

Le but de la méthode de *clustering* est de construire des groupes (ou classes) de stations de telle manière que la distance entre des stations appartenant au même groupe soit la plus faible possible et que, inversement, la distance entre deux classes différentes soit la plus grande possible (pour garantir une certaine isolation des classes de stations). Dans ce contexte, nous devons donc étendre la notion de distance entre stations et définir une distance entre les classes.

Définition d'une distance entre classes

Au cours du processus de *clustering*, deux classes sont fusionnées pour former une nouvelle classe dont le cardinal est la somme des cardinaux des deux classes précédentes. Soient ainsi les classes A et B comptant respectivement n_A et n_B stations, et la nouvelle classe fusionnée AB, avec $n_A + n_B$ stations. La distance de la nouvelle classe AB à une classe tierce C pourra être calculée selon différents critères de fusion des classes :

- Critère de la classe la plus proche (ou de la distance minimale) : $d(C,AB) = \min\left(d(C,A), d(C,B)\right)$. Ce critère est invariant selon les transformations monotones de la fonction distance. Il permet d'obtenir des classes "étendues" qui peuvent contenir à leurs extrémités des stations relativement distantes (voir figures 5.1 et 5.2).
- Critère de la classe la plus éloignée (ou de la distance maximale) : $d(C,AB) = \max\left(d(C,A), d(C,B)\right)$. Ce critère est invariant selon les transformations monotones de la fonction distance. Il permet d'obtenir des classes "sphériques" (c'est-à-dire plus compactes, voir figures 5.3 et 5.4).
- Critère de la moyenne pondérée : $d(C,AB) = \dfrac{1}{n_A + n_B}\left(n_A d(C,A) + n_B d(C,B)\right)$. Ce critère n'est pas invariant selon les transformations monotones de la fonction distance.

Méthodes de construction des classes de stations par clustering hiérarchique

Nous supposons que nous avons calculé une matrice de taille $N \times N$ des distances entre stations. Nous détaillons ci-dessous les différentes étapes du processus de *clustering* (méthode hiérarchique d'agrégation des *clusters*) :

1) Initialisation avec N classes correspondant aux N stations, les distances entre les classes étant prises égales aux distances entre stations.
2) Sélection des deux éléments les plus proches dans la matrice des distances et fusion en une seule classe.
3) Substitution des deux éléments sélectionnés au 2) par un nouvel élément constituant une classe. Les distances entre ce nouvel élément et les autres éléments sont calculées selon le critère choisi de la distance entre classes, et la matrice des distances est mise à jour.
4) Retourner au 2) et répéter 2) et 3) jusqu'à ce que le nombre de classes soit inférieur à un seuil fixé a priori.

On peut modifier l'étape 4) pour limiter la taille maximale des classes finales obtenues. Par exemple, à chaque itération, si une classe dépasse la taille limite, on fige cette classe qui est alors retirée du processus de *clustering*.

Remarque : nous avons étudié une deuxième méthode de construction de groupes de stations ; la méthode de partition des clusters (ou *splitting*). Dans cette méthode, à chaque itération on réalloue une station d'une classe vers une autre classe dans le but de minimiser un critère d'optimisation basé sur une matrice estimant la variabilité entre les stations d'un même groupe.

Exemple d'application du *clustering*

Le cas d'un réseau UMTS de 175 stations dans un environnement dense urbain a été étudié. La matrice d'interférence $\widetilde{K}$ a été estimée à l'aide d'un simulateur statique de réseau UMTS, le trafic étant généré en utilisant des tirages de Monte-carlo des mobiles. La méthode du *clustering* hiérarchique a tout d'abord été testée sur ce réseau, en utilisant le critère de la classe la plus proche.

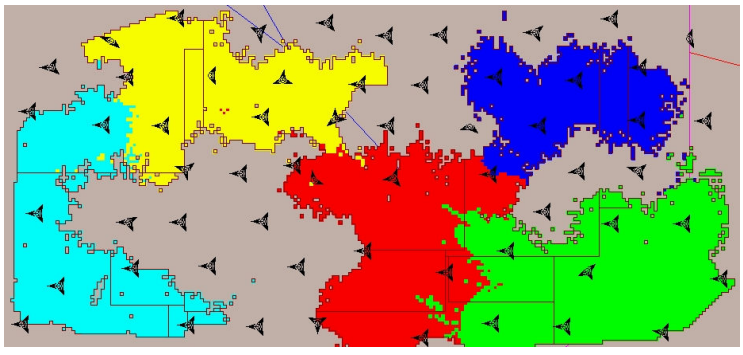


Figure 5.1 : Carte des ensembles d'influence construits en utilisant le critère de la distance minimale entre classes.

La figure 5.1 décrit les classes de stations obtenues avec ce critère, en stoppant le processus de *clustering* quand le cardinal de la classe la plus grande dépasse 15 stations. Ci-dessous, le résultat final du processus de *clustering* sur l'ensemble du réseau est illustré sur la figure 5.2.

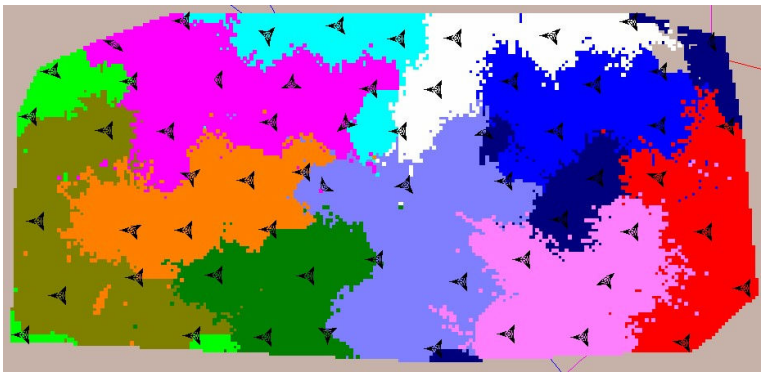


Figure 5.2 : Répartition en classes de stations pour l'ensemble du réseau en utilisant le critère de la distance minimale entre classes.

Sur la figure 5.2, il apparaît bien que le critère de distance minimale génère plutôt des classes de stations étendues qui peuvent contenir à leurs extrémités des stations relativement distantes. De plus, on remarque que plusieurs petits groupes de stations ont été laissés à part à la fin du processus de *clustering*, du fait de la limitation de la taille des classes. Pour améliorer cet aspect, on pourrait intégrer ces stations aux classes voisines les plus proches. A présent, la méthode du *clustering* hiérarchique est testée sur le même réseau, en utilisant le critère de la classe la plus éloignée.

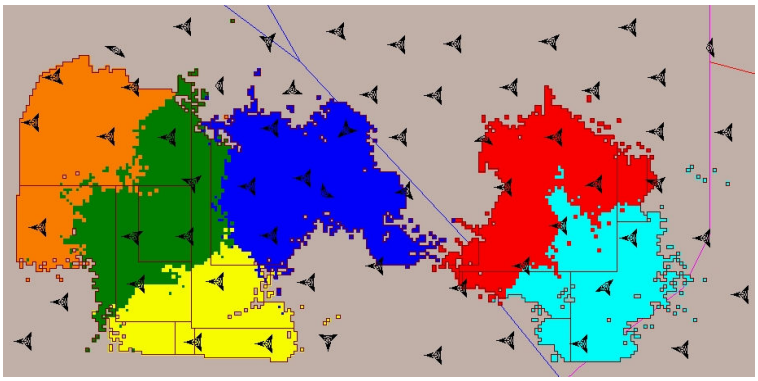


Figure 5.3 : Carte des ensembles d'influence construits en utilisant le critère de la distance maximale entre classes.

La figure 5.3 décrit les classes de stations obtenues avec ce critère, en stoppant le processus de *clustering* quand le cardinal de la classe la plus grande dépasse 15 stations. Dans cet exemple, deux classes ont fusionné et atteint un cardinal de 19 stations (classe colorée en bleu sur la figure). Les cardinaux des autres classes sont respectivement de 14 (pour trois classes), 12 et 10 stations. Si l'on veut limiter la taille des classes, la classe la plus grande pourrait être retirée avant de relancer le processus de *clustering* avec les classes restantes.

Ci-dessous, le résultat final du processus de *clustering* sur l'ensemble du réseau est illustré sur la figure 5.4. Sur cette figure, il apparaît bien que le critère de distance maximale génère plutôt des classes de stations plus compactes (ou sphériques). Par ailleurs avec ce critère, on n'a plus le problème des petits groupes de stations laissés à part du processus de *clustering* : en effet, les classes obtenues en fin de processus sont toutes connexes et de tailles relativement homogènes.

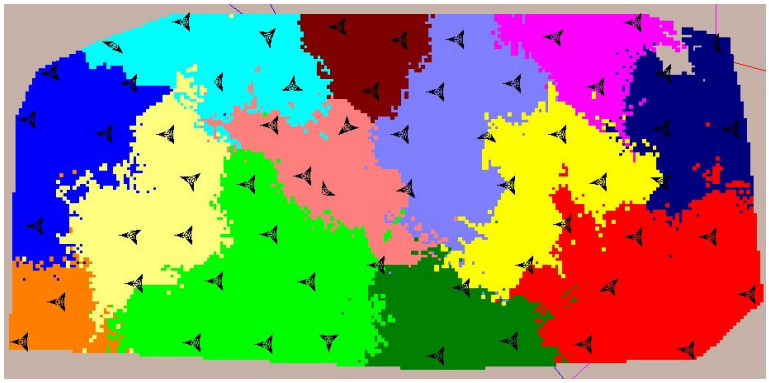


Figure 5.4 : Répartition en classes de stations pour l'ensemble du réseau en utilisant le critère de la distance minimale entre classes.

5.1.2. Application de l'ingénierie automatique

En mode *on-line*, le but est d'adapter le paramétrage du réseau sur la base de remontées régulières de mesures. Deux scénarios d'utilisation sont envisageables :

- Le premier se fonde sur des indicateurs radio de bas niveau (cellule) disponibles en temps réel : charges, délais *throughput*. Ce scénario nécessite l'introduction de modules de contrôle au niveau des composants du réseau : node B, module C-RRM, voire RNC. Ces modules ne rempliront que des tâches élémentaires et locales de modification du paramétrage.
- Le second scénario utilise des indicateurs de performances de haut niveau (comme, par exemple, les KPI calculés au niveau de l'OMC). Les mesures sont remontées et filtrées sur des intervalles de temps plus longs. Cette approche permet d'avoir d'une part des indicateurs plus globaux de l'état du réseau, et d'autre part de développer une politique de contrôle plus complète, ce qui peut être important pour résoudre certains problèmes présents sur l'ensemble d'une zone. Enfin, ces deux types de fonctionnement ne sont pas exclusifs l'un de l'autre : on peut ainsi imaginer un mode de fonctionnement permettant aux contrôleurs de haut niveau (au niveau de l'OMC) d'envoyer des commandes aux modules de contrôle distribués : mise en place et application d'une stratégie globale au niveau local ([Gaï05] [Sal05]).

Paramètres de l'ingénierie dynamique

Dans l'optique d'un déploiement opérationnel, plusieurs paramètres devront être fixés par l'opérateur. Ces valeurs pourront être tout d'abord testées et évaluées à l'aide de simulations logicielles avant d'être validées sur des réseaux expérimentaux :

- *Granularité des mesures* : ce paramètre dépend des contraintes de la norme et de celles spécifiques à chaque constructeur. Cependant à l'aide de simulations logicielles, on peut déterminer une valeur cible théorique.
- *Tailles des filtres* : en fonction de la granularité des mesures primaires, on pourra tester différentes largeurs de filtres afin de déterminer la valeur la plus efficace.
- *Temps inter-modifications* : à partir des mesures filtrées et remontées au contrôleur, le CLF va modifier régulièrement le paramétrage RRM du réseau UMTS. Ainsi en fonction de la granularité des mesures, de la taille des filtres et de l'efficacité des modifications de paramétrage, l'intervalle de temps entre deux corrections devra aussi être optimisé.

- *Paramètres contrôlés et granularité des modifications* : en fonction de l'analyse des ingénieurs radio, des contraintes de normalisation et celles liées aux constructeurs, il s'agira de déterminer quels paramètres seront modifiées par le module d'ingénierie dynamique, ainsi que la granularité optimale de ces corrections.
- *Niveau d'introduction des modules de contrôles dans le réseau* : en fonction de la disponibilité des métriques, de la standardisation des interfaces et des éléments du réseau d'accès radio, l'opérateur pourra déterminer en fonction de sa stratégie de contrôle le niveau le plus optimal où introduire les contrôleurs automatiques : nodes B, RNC, OMC…
- *Modification simultanée de toutes les cellules du réseau* : ce choix dépendra de la politique de contrôle. Le choix de modifier simultanément toutes les cellules pourra être mis en oeuvre aussi bien dans les scénarios de contrôle de type haut niveau et bas niveau.
- *Modification d'une seule cellule par ensemble d'influence et par itération* : contrairement au cas précédent, la détermination de groupes d'interférence de stations ne pourra être envisageable que dans une approche globale au niveau de l'OMC voire du RNC. Alors seulement, il sera possible de modifier une seule station par groupe d'interférence. Ce type de contrôle est donc pertinent avec un scénario de haut niveau.

5.1.3. Exemples de contrôle de paramètres RRM en mode *on-line*

La figure 5.5 illustre le contrôle dynamique du paramètre *Add_win* d'une cellule en fonction des indicateurs de blocage des liens de MD, de coupure et de la charge *downlink*. Dans ce réseau, les paramètres de seuil de charge cible DL et *Add_win* de l'ensemble des stations sont (re-)paramétrés dynamiquement et simultanément. Le contrôleur de logique floue a été optimisé par essaim de particules en multi-objectif* simultanément pour les algorithmes d'admission et de macro-diversité.

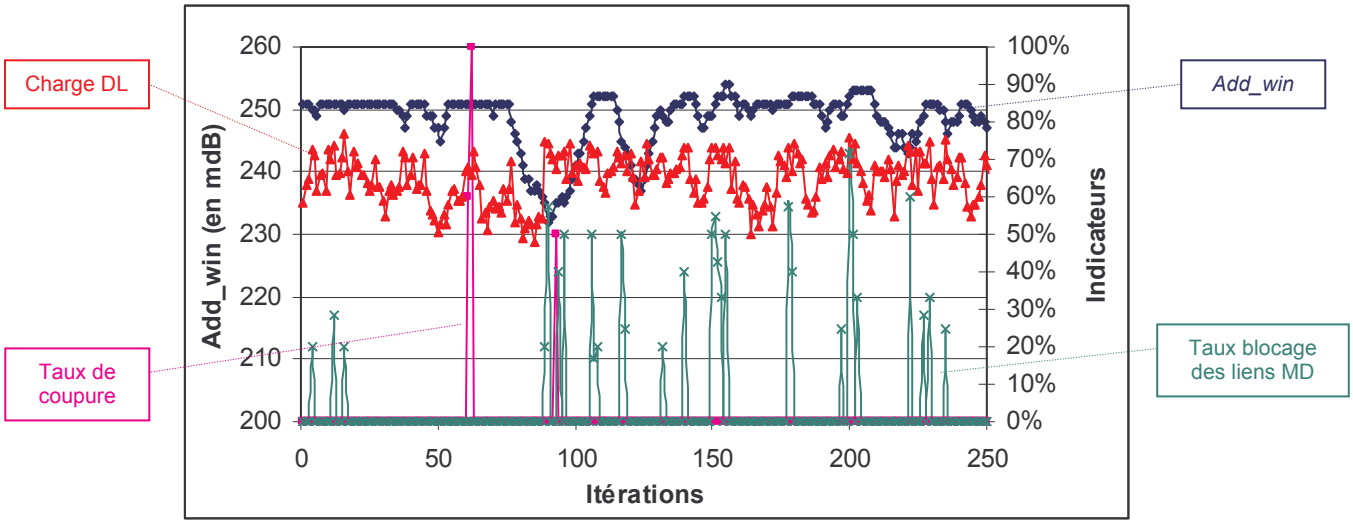


Figure 5.5 : Variations des indicateurs de charge DL (rouge), taux de coupure instantané (rose), taux de blocage des liens de MD instantané (vert) et du paramètre *Add_win* (bleu) sur une cellule chargée.

La figure 5.5 montre les variations dynamiques du paramètre *Add_win* de la cellule considérée. Jusqu'à l'itération 90, il n'y a pas de blocage des liens de MD sur la cellule et la charge DL a tendance à diminuer. Le contrôleur diminue donc la taille de la fenêtre d'ajout des liens de MD jusqu'à une valeur de 2.3 dB. Mais à partir de l'itération 90, le blocage des liens de MD et le taux de coupure augmentent brutalement sur la cellule. Cela indique que le

* Cf. paragraphe 4.1.3.2.

paramètre *Add_win* est trop faible. En conséquence, trop de liens de MD sont créés sur la cellule. Ce point est confirmé par la hausse brutale de la charge DL de 50% à plus de 70%. Pour réduire le nombre de liens en MD, le contrôleur de logique floue décide alors d'augmenter le paramètre *Add_win* de la cellule jusqu'à une valeur de 2.5 dB à l'itération 110.

Pour estimer l'efficacité du contrôle dynamique, nous indiquons dans la table 5.1 suivante les causes de rejet des mobiles avec sans ingénierie dynamique. Nous rappelons que le contrôleur de logique floue a été optimisé par essaim de particules en multi-objectif[*] simultanément pour les algorithmes d'admission et de macro-diversité.

Politiques	Paramétrage uniforme statique	CLF AC +MD[**]
Nombre de mobiles tirés	159800	160300
Nombre total de mobiles rejetés	39084 Référence	34920 **Gain =+10.7 %**
Nombre de mobiles bloqués à l'admission: CHARGE_DL	35999	30412
Nombre de mobiles bloqués à l'Admission: CHARGE_UL	2076	3483
Nombre de mobiles bloqués à l'admission: COUV_DL	668	493
Nombre de mobiles bloqués à l'admission: COUV_UL	0	0
Nombre de mobiles bloqués à l'admission: PILOTE	0	0
Total de mobiles bloqués	**38743**	**34388**
Taux de blocage global	**24.2 %**	**21.5 %**
Nombre de mobiles coupés: COUV_DL	283	358
Nombre de mobiles coupés: COUV_UL	0	0
Nombre de mobiles coupés: CHARGE_DL	0	0
Nombre de mobiles coupés: CHARGE_UL	58	174
Total de mobiles coupés	**341**	**532**
Taux de coupure global	**0.2 %**	**0.3 %**

Table 5.1 : Analyse des causes de rejet des mobiles avec et sans[***] contrôle dynamique des paramètres des algorithmes d'admission et de macro-diversité.

Par rapport à la situation initiale d'un paramétrage uniforme et statique, le contrôle dynamique a permis de réduire nettement le nombre total de mobiles rejetés (bloqués ou coupés). En adaptant les paramètres d'admission et de macro-diversité, le CLF a permis de réduire le nombre d'appels bloqués, en dégradant à peine le taux de coupure global dans le réseau. Un taux de coupure de 0.3% est acceptable pour un opérateur, la valeur cible habituelle étant 2% de coupure dans le réseau.

Comparaison des solutions optimales avec l'OEP en mono et multi-objectif

La figure 5.6 décrit le nuage des solutions explorées au cours d'une optimisation par essaim de particules en multi-objectif dans le plan fonction coût blocage – fonction coût rejet. Sur cette figure, nous avons ajouté le point correspondant à la solution optimale obtenue par OEP en mono-objectif (pour un coefficient de mixage de 4) dans les mêmes conditions. Dans cet exemple, la solution optimale obtenue en mono-objectif appartient au front optimal des solutions explorées par l'algorithme d'OEP en multi-objectif. L'avantage de la version multi-objectif est d'obtenir un ensemble des solutions optimales – le front du nuage – représentant chacune un compromis différent en termes de fonction coût de blocage et de fonction coût de coupure. Ce point sera étudié au paragraphe 6.4.3 en mode *on-line*.

[*] Cf. paragraphe 4.1.3.2.
[**] AC+MD : Contrôle d'admission et macro-diversité.
[***] Paramétrage uniforme et statique du réseau : seuil de charge cible DL=0.7 et *Add_win*=3.5 dB.

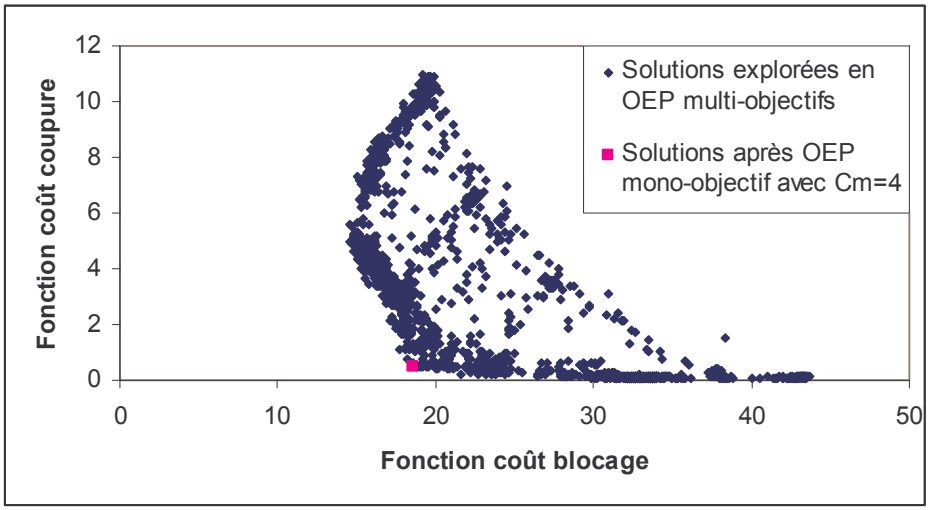


Figure 5.6 : Comparaison entre la solution optimale mono-objectif (en rose) et l'ensemble des points explorés en multi-objectif (en bleu) dans le plan coût blocage - coût coupure.

5.2. Le mode off-line en UMTS

Le deuxième mode de fonctionnement envisagé pour l'ingénierie automatique consiste à optimiser, node B par node B, le paramétrage du réseau *off-line* (hors-ligne), c'est-à-dire sans être connecté au réseau. On pourra ainsi utiliser des simulations logicielles ou des données historiques du réseau. L'intérêt est de corriger et adapter le paramétrage des algorithmes RRM localement en fonction des contraintes d'environnement et des données de trafic.

5.2.1. Objectifs du mode *off-line*

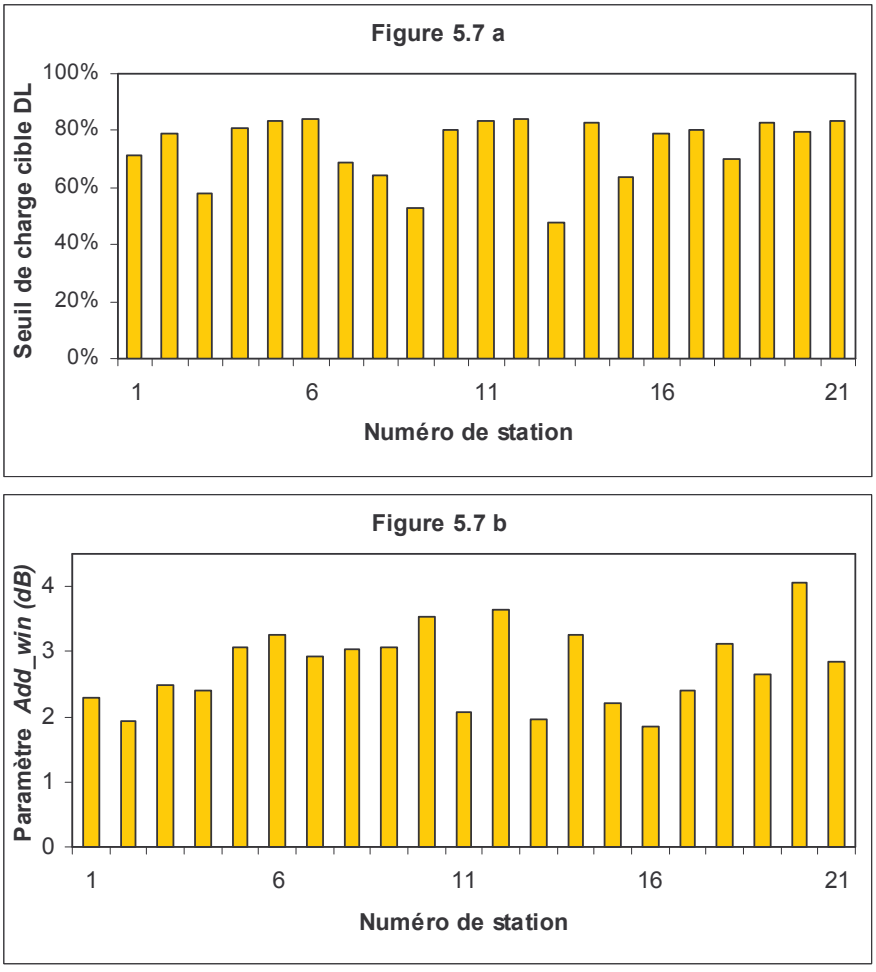
Le paramétrage initial à l'ouverture commerciale des algorithmes RRM des réseaux UMTS était uniforme, souvent conforme aux recommandations a priori des constructeurs. L'objectif du mode *off-line* est de proposer une correction de ces paramètres pour chaque node B afin d'améliorer les performances et la qualité de service dans le réseau en fonction des contraintes d'environnement et de trafic. Les modifications de paramétrage pourront être appliquées au niveau de l'OMC et être déterminées à l'aide :

- De données de trafic soit prévues par des hypothèses de trafic, ou déterminées par de mesures terrain sur la zone considérée pour certaines plages horaires représentatives.
- Des contrôleurs *off-line* optimisés soit par essaim de particules (mono et multi-objectif) soit par apprentissage par renforcement.
- Des simulations logicielles soit statiques (difficile de mettre en évidence l'optimisation de paramètres RRM), soit dynamiques (pour optimiser des algorithmes sensibles aux effets dynamiques : mobilité, apparitions et disparitions aléatoires de mobiles dans le réseau).
- Des validations sur un réseau expérimental de réseau UMTS en environnement réel.

Par rapport au mode *on-line*, plusieurs points distinguent le mode *off-line*. Tout d'abord, il n'est pas nécessaire d'isoler des groupes de stations (ou groupes d'interférence). En effet en hors-ligne, on va plutôt chercher à déterminer un paramétrage moyen efficace par node B (pour un niveau et une distribution de trafic donnés). Au contraire dans un fonctionnement dynamique, les contrôleurs adaptent régulièrement les paramètres du réseau, le paramétrage de l'itération précédente des stations voisines les plus proches (en terme de distance fondée sur la matrice d'interférence) ayant une influence sur le calcul des corrections des paramètres

pour la station centrale à l'itération courante. Cette influence réciproque est à l'origine des effets de type ping-pong. Par ailleurs, le mode *off-lin*e présente l'avantage d'être envisageable à court terme dans les réseaux réels en adaptant le paramétrage des nodes B au niveau des OMC. Au contraire l'implémentation de l'approche dynamique nécessitera d'impliquer des constructeurs (par exemple dans le projet EUREKA Gandalf [Stu05]) ainsi que de présenter des rapports techniques et d'introduire des spécifications au 3GPP. Enfin, le mode *off-line* est la seule application possible de la technique d'optimisation par essaim de particules. En effet, cette procédure ne peut être directement implémentée dans les réseaux réels d'une part à cause du temps de calcul, d'autre part du fait même de l'algorithme d'optimisation : l'OEP est basée sur l'exploration simultanée par un ensemble de particules de l'espace des contrôleurs ([Dub05]) et sur un réseau réel on ne peut tester qu'une seule particule (ou contrôleur) à la fois.

Comme indiqué ci-dessus, l'objectif du mode *off-line* est donc d'optimiser le paramétrage RRM node B par node B à partir d'un paramétrage initial uniforme. Sur les figures 5.7, nous représentons respectivement les distributions des paramétrages par secteur du seuil de charge cible DL (pour un paramétrage initial à 0.7) et de la taille de fenêtre d'ajout *Add_win* (pour un paramétrage initial à 3.5 dB) après optimisation *off-line* avec la politique *Select_multi_1* obtenue au paragraphe 7.2 sur la zone de la ville 1 (21 secteurs). La figure 5.8 représente la distribution du paramétrage par secteur de la taille de fenêtre d'ajout *Add_win* (pour un paramétrage initial à 3.5 dB) après optimisation *off-line* sur la zone de la ville 2 (109 secteurs).



Figures 5.7 : Histogrammes du paramétrage du seuil de charge cible DL et de *Add_win* après optimisation en mode *off-line* sur le réseau de la ville 1 (21 secteurs).

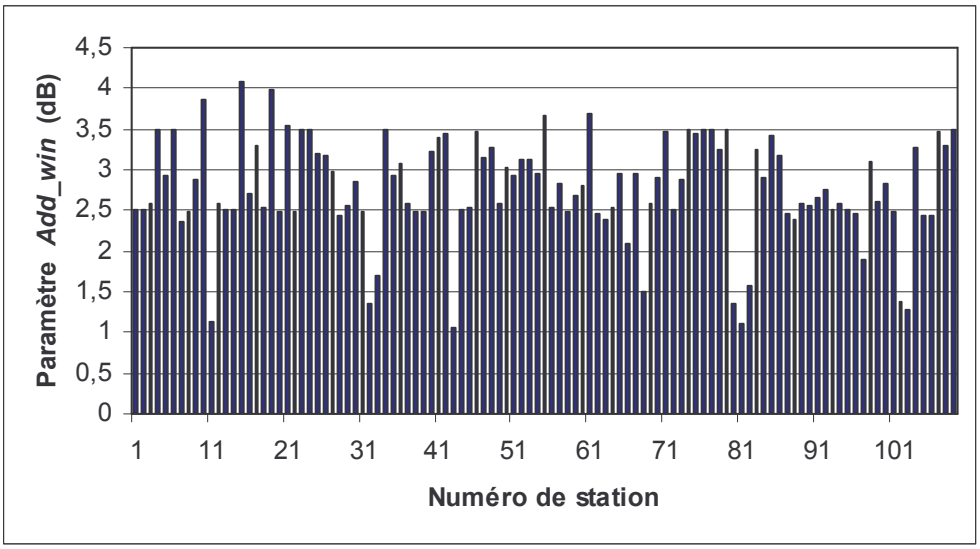


Figure 5.8 : Histogramme du paramétrage de *Add_win* après optimisation en mode *off-line* sur le réseau de la ville 2 (109 secteurs).

5.2.2. Avec l'optimisation par essaim de particules

Le but du mode *off-line* est d'ajuster node B par node B le paramétrage RRM du réseau UMTS. Au chapitre 3 nous avons introduit des contrôleurs de logique floue qui ont été plutôt conçus (à partir de règles intuitives ou expérimentales de réglage des paramètres) pour adapter de façon dynamique un ensemble de paramètres de contrôle qui régule et optimise le fonctionnement du système. L'utilisation de l'optimisation par essaim de particules telle que décrite sur la figure 4.3 permet d'optimiser les performances de ces contrôleurs dynamiques.

Au sein de France Télécom Recherche & Développement, nous avons mis au point une méthode permettant de déterminer par simulations un jeu de paramètres optimisés à implémenter sur le réseau réel. La particularité de ce procédé est de fournir au module d'optimisation des critères de qualité obtenus lors d'une évaluation *off-line* utilisant des paramètres fixes qui sont calculés grâce à une évaluation *on-line* du réseau. La figure 5.9 suivante résume le fonctionnement de ce processus d'optimisation du mode *off-line*. Dans ce processus, il y a deux niveaux de contrôle :

- Le niveau externe utilise le processus d'optimisation par essaim de particules comme décrit au chapitre 4. Le module d'optimisation fait évoluer dans l'espace des contrôleurs plusieurs particules. Il reçoit en entrées les fonctions coûts globales fournies par une simulation avec un paramétrage fixe pour la particule considérée. En fonction des performances des particules, le module d'optimisation propose en sortie une modification de paramétrage des CLF dynamiques.
- Le niveau de contrôle interne permet de déterminer à partir d'un CLF dynamique (correspondant à chaque particule du processus d'OEP) un paramétrage fixe du réseau qui sera ensuite testé et évalué. Pour cela, on lance une simulation de réseau avec contrôle dynamique du paramétrage RRM. On stocke les valeurs successives de ce paramétrage adaptatif. A partir de ces valeurs stockées, on détermine un paramétrage fixe des algorithmes RRM pour chaque node B (par exemple la moyenne des paramètres RRM stockés). Les performances de ce paramétrage fixe vont être évaluées à l'aide d'une simulation de réseau. A partir des métriques estimées au cours de la simulation, on calcule les fonctions coûts correspondantes qui seront fournies en entrée du module d'optimisation.

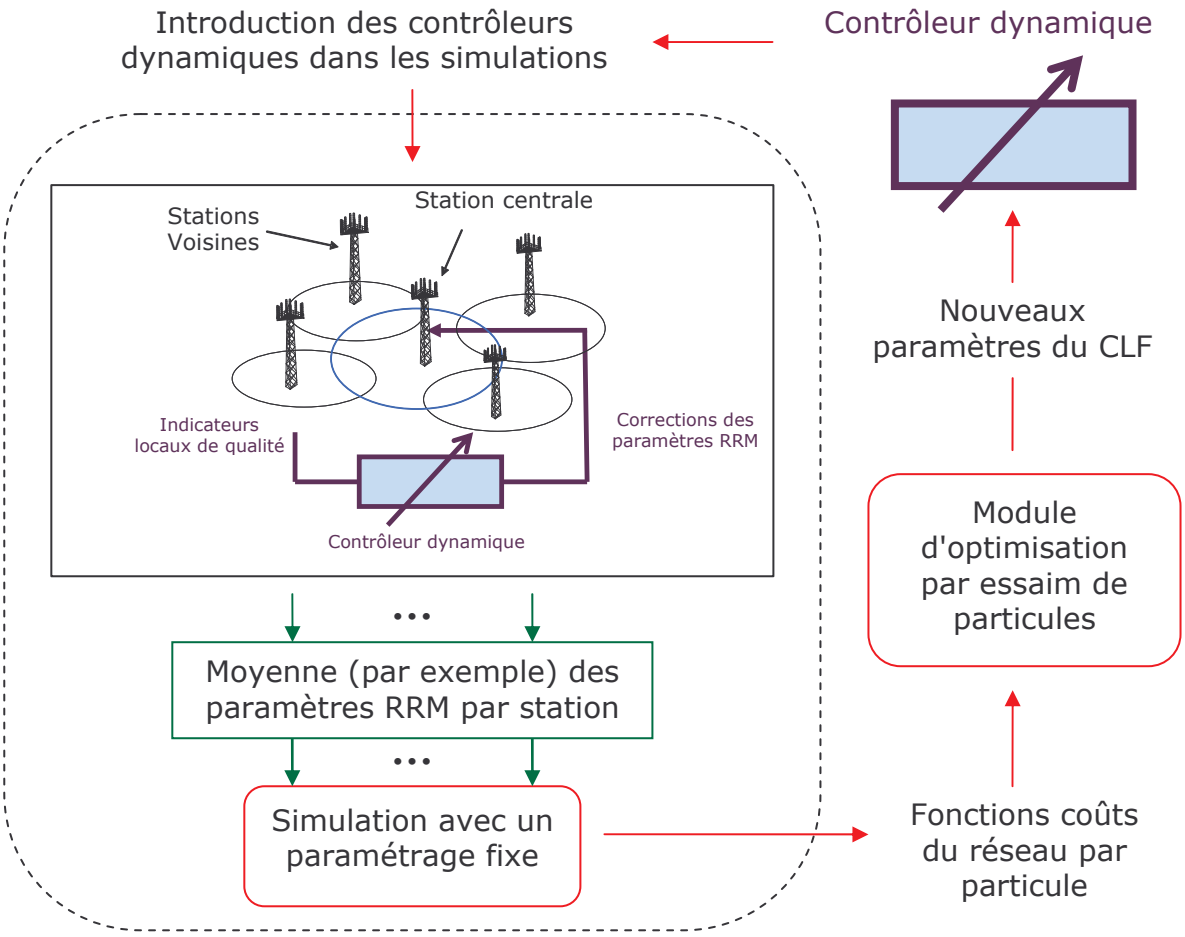


Figure 5.9 : Processus d'optimisation du mode *off-line* fondé sur l'OEP et un CLF dynamique.

L'avantage de cette méthode est que le module d'optimisation reçoit en entrées les fonctions coûts déduites d'une simulation du réseau avec un paramétrage fixe (donc représentatif de la situation terrain). Le module adapte donc le paramétrage des CLF dynamiques en fonction des performances fournies par le paramétrage fixe. Par ailleurs, dans le niveau de contrôle externe, il serait possible de remplacer le module d'optimisation par essaim de particules par un autre processus d'apprentissage (fondé sur la mesure de performances comme mentionné au paragraphe 4.2.1). Enfin le procédé d'optimisation du mode *off-line* décrit par la figure 5.9 pourrait également être utilisé directement sur un réseau réel, pour optimiser des paramètres qui ne nécessitent pas d'être contrôlés en continu.

5.2.3. Avec l'apprentissage par renforcement

Ci-dessous, la possibilité d'utiliser la modélisation par un SIF avec apprentissage par renforcement (comme décrit au paragraphe 4.2.6) pour le problème d'optimisation en mode *off-line* est étudiée. Nous rappelons que le but du mode *off-line* est d'ajuster node B par node B le paramétrage RRM du réseau UMTS en utilisant des simulations logicielles ou des données historiques collectées dans le réseau. Avec l'algorithme de *Q-learning* nous envisagerons aussi la possibilité de d'introduire une phase d'apprentissage / réglage dans un réseau réel : avec remontées des métriques toutes les heures par exemple sur l'ensemble des contrôleurs d'une zone.

- Phase d'initialisation : transmission de connaissance d'un SIF optimisé pour le mode *on-line*

A partir d'un contrôleur dynamique optimisé par apprentissage par renforcement à l'aide de simulations logicielles, nous voulons proposer une méthode permettant de guider l'apprentissage par renforcement d'un contrôleur *off-line* du paramétrage du réseau mobile. La méthode d'apprentissage guidée qui est envisagée se fonde sur les mêmes idées que l'apprentissage en série utilisant l'algorithme du *Q-learning* (pour affiner la connaissance du système ou introduire de nouveaux indicateurs de performances) tel que décrit au paragraphe 4.2.7 : on va restreindre le champ des actions possibles en sortie des règles de décisions élémentaires du nouveau SIF $S_{Off\text{-}line}$ pour le mode *off-line* en introduisant de la connaissance a priori précise ou imprécise apportée par le SIF $S_{Dyn.}$ dynamique optimisé. La figure 5.10 ci-dessous décrit comment déduire du SIF dynamique $S_{Dyn.}$ le SIF équivalent $S_{Equiv.}$ pour le mode *off-line* : transmission de connaissance du mode *on-line* vers le mode *off-line*. En début de processus, les fonctions de qualité élémentaires du SIF $S_{Equiv.}$ sont initialisées à zéro.

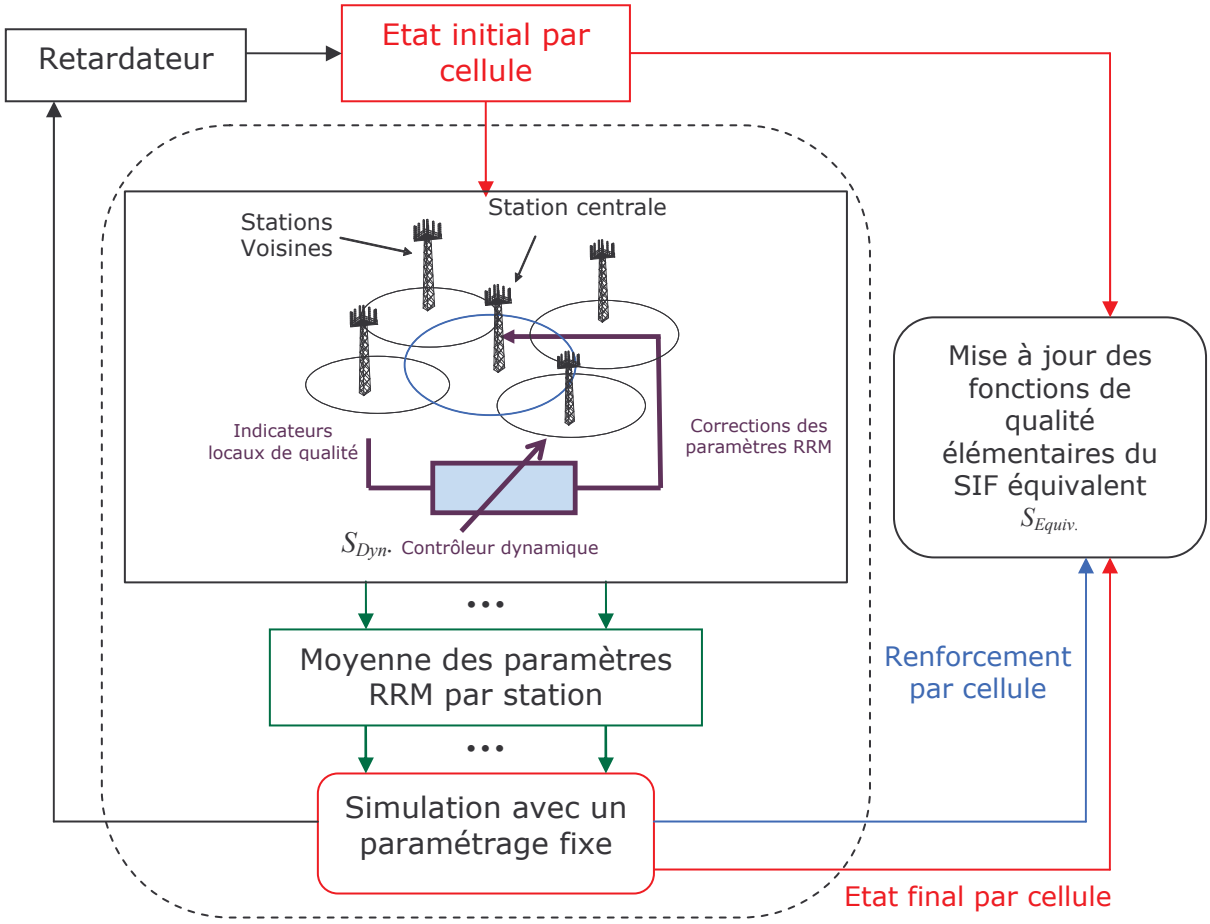


Figure 5.10 : Déduction du SIF équivalent $S_{Equiv.}$ pour le mode *off-line* à partir du SIF dynamique $S_{Dyn.}$

La structure du SIF $S_{Equiv.}$ a été préalablement définie à l'aide de l'expérience de paramétrage des ingénieurs radio ainsi que des règles déduites du mode dynamique. Comme pour le processus d'optimisation du mode *off-line* avec l'OEP, on lance une simulation de réseau avec contrôle dynamique du paramétrage RRM par le SIF $S_{Dyn.}$. On stocke les valeurs successives de ce paramétrage adaptatif. A partir de ces valeurs stockées, on détermine un paramétrage fixe des algorithmes RRM pour chaque node B (par exemple la moyenne des paramètres RRM stockés). Les performances de ce paramétrage fixe vont être évaluées à l'aide d'une simulation de réseau. A partir des métriques estimées au cours de la simulation, on calcule

l'état final de chaque cellule ainsi que les valeurs de renforcement. Ces données ainsi que l'état initial des cellules sont fournies en entrées du module de mise à jour des fonctions de qualité élémentaire du SIF $S_{Equiv.}$. On itère le processus à partir de l'état final du réseau. Il est aussi possible de répéter le processus de transmission de connaissance pour différents niveaux et différentes distributions de trafic, afin d'explorer suffisamment l'ensemble des états possibles de la cellule.

Une fois que l'on considère que le processus a convergé (plus de modification significative des fonctions de qualité élémentaires), il s'agit de déduire du SIF équivalent $S_{Equiv.}$ la structure initiale du nouveau SIF $S_{Off\text{-}line}$ pour le mode *off-line*. Pour cela, on va définir les nouveaux ensembles de définition des variables de sortie des différentes règles de $S_{Off\text{-}line}$, en fonction des sorties fournies par la politique optimale déterminée par $S_{Equiv.}$: de la même manière que sur la figure 4.10, on peut introduire un niveau de connaissance précis (un seul label de sortie pour la règle correspondant à une valeur numérique), imprécis (plusieurs labels de sortie correspondant chacun à un ensemble de définition dans lequel on choisit une valeur numérique) ou aucune connaissance (en revenant aux intervalles de définition initiaux) dans $S_{Off\text{-}line}$. Sur le SIF ainsi défini, nous pouvons maintenant mettre en oeuvre la phase d'exploration de l'algorithme du *Q-learning*.

- Phase d'exploration de l'apprentissage par renforcement

A présent, on applique au SIF $S_{Off\text{-}line}$ défini ci-dessus l'algorithme du *Q-learning* dans sa phase d'exploration. Comme décrit sur la figure 5.11, à chaque itération les paramètres RRM par cellule sont corrigés à l'aide des modifications $\Delta c_{RRM}^{Pol.\,PS}$ déterminées par la politique P.S. (pseudo-stochastique). Le paramétrage est ensuite figé lors de la simulation du réseau UMTS. Enfin on met à jour les fonctions de qualité élémentaires du SIF en fonction de l'état des cellules initial et final (indicateurs de performances et paramétrage), des corrections appliquées et des valeurs de renforcement. Le contrôleur tient compte simultanément des informations de chaque cellule pour la mise à jour des fonctions de qualité.

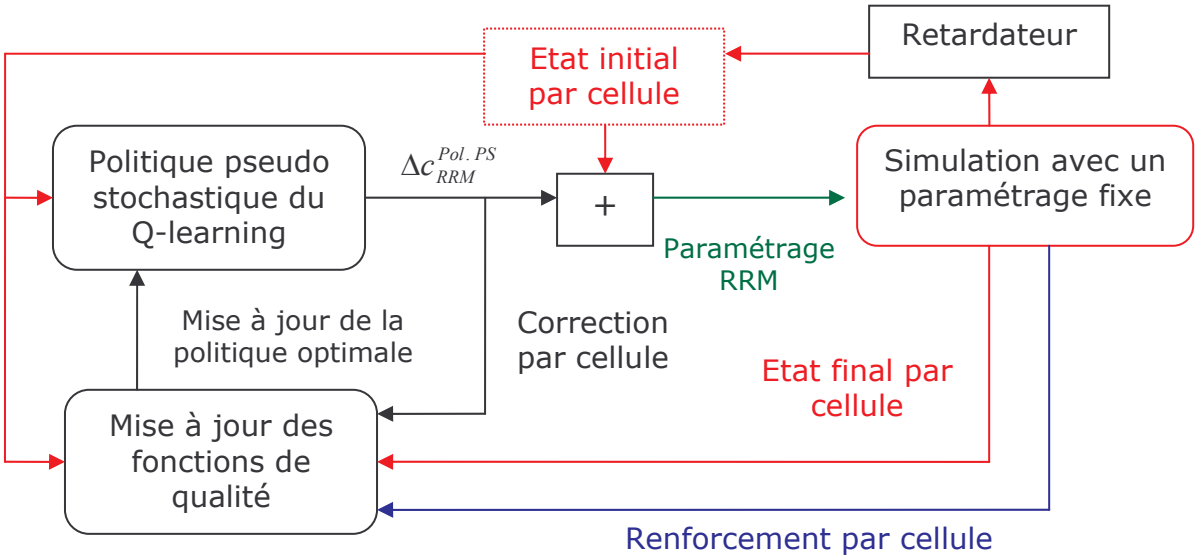


Figure 5.11 : Phase exploration de l'apprentissage par renforcement d'un contrôleur *off-line* du paramétrage du réseau mobile.

Une fois que l'on considère que l'algorithme a suffisamment "appris" (plus de modification significative des fonctions de qualité élémentaires), on peut répéter le processus d'exploration avec des niveaux et distributions de trafic modifiés graduellement. On est prêt à utiliser le

contrôleur en phase d'exploitation, dés lors que l'ensemble des règles du SIF a été visité un nombre suffisant de fois.

- Phase d'exploitation et validation par des simulations / expérimentations

Le contrôleur *off-line* a donc été initialisé à l'aide de la connaissance transmise d'un SIF optimisé pour le mode *on-line* et calibré avec des simulations logicielles de réseau. Comme décrit sur la figure 5.12, nous proposons de le tester dans le cas d'une situation de paramétrage initial uniforme. Pour cela, un réseau UMTS est simulé avec un certain niveau de charge réaliste. A partir des indicateurs de performances (pouvant être une synthèse des résultats de simulations et de mesures du terrain, grâce à des outils de type *super-prédiction*[*]) on détermine les corrections des paramètres RRM à appliquer par cellule $\Delta c_{RRM}^{Pol.Opt.}$ en utilisant la politique optimale de l'algorithme de *Q-learning* (phase d'exploitation). Enfin, on lance une nouvelle simulation de réseau mobile avec les valeurs RRM corrigées par cellule et pour le même niveau de charge :

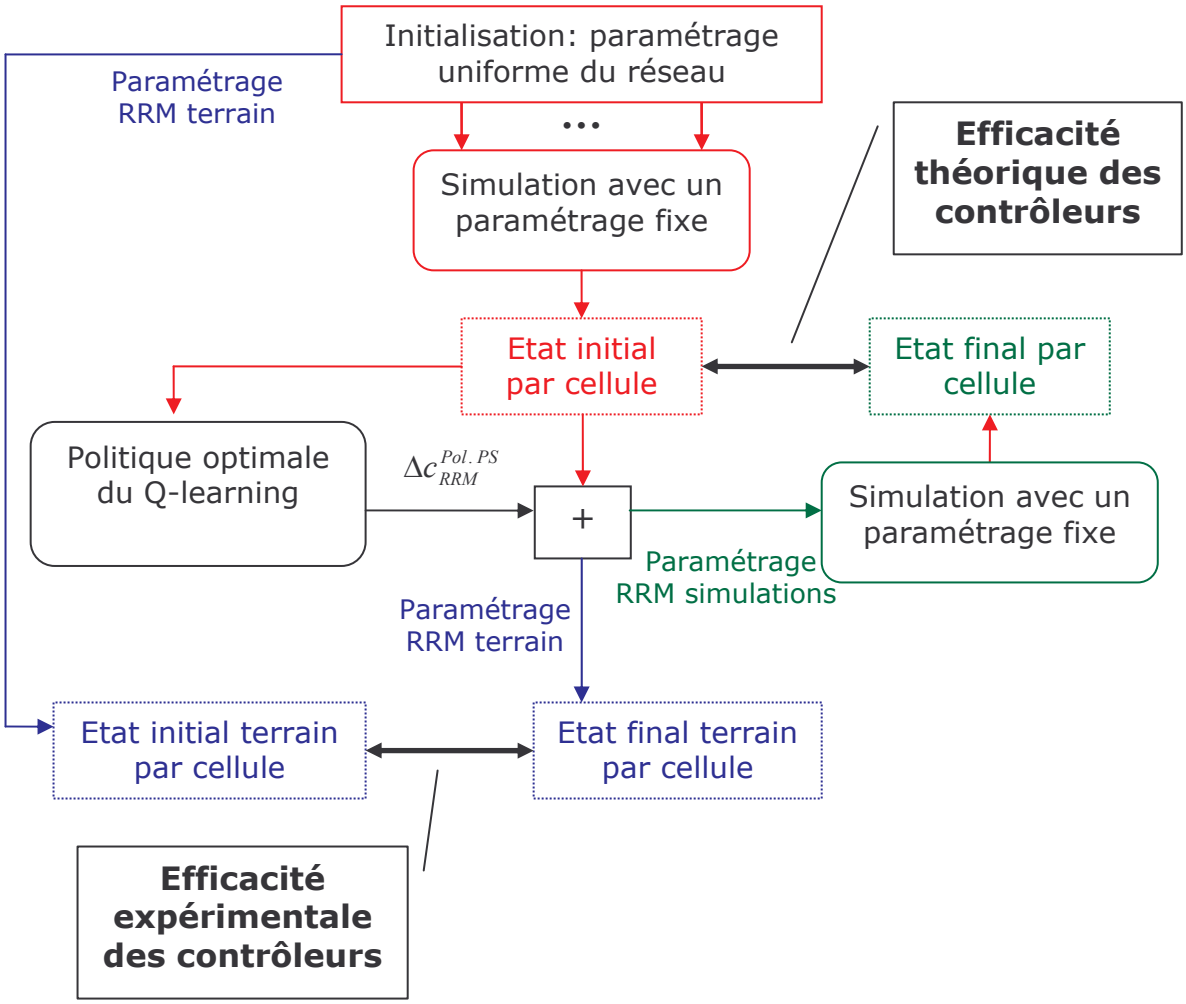


Figure 5.12 : Phase exploitation de l'apprentissage par renforcement d'un contrôleur *off-line* de réseau mobile et validation par simulations / expérimentale.

[*] Nouvelle méthodologie développée en interne par France Télécom Recherche & Développement pour améliorer les prédictions faites par les simulateurs de réseaux mobiles à l'aide des métriques et KPI mesurés dans les réseaux mobiles opérationnels ou expérimentaux.

Pour tester l'efficacité du contrôleur *off-line* construit avec les méthodes décrites dans les paragraphes précédents :

- On compare tout d'abord les indicateurs de performance calculés par le simulateur avant et après modification des paramètres RRM, pour vérifier l'efficacité théorique du contrôleur.
- On pourra aussi corriger les paramètres RRM (en fonction des recommandations du contrôleur) sur la partie du réseau correspondant à la zone simulée. On comparera alors les indicateurs de performances pour des niveaux de trafic équivalents avant et après modification des paramètres RRM pour valider l'efficacité expérimentale des contrôleurs.

L'étude de la validation expérimentale des contrôleurs de type *off-line* nous a permis d'envisager un nouveau type de fonctionnement intermédiaire entre les modes *on-line* et *off-line*. En tenant compte des contraintes temporelles liées aux remontées de mesures et aux modifications du paramétrage, le contrôleur (implémenté par exemple au niveau de l'OMC) qui aura été calibré à partir de simulation pourra être utilisé en mode d'exploration dans le réseau mobile réel (les corrections n'étant pas complètement aléatoires, pour éviter les modifications aberrantes). Ainsi cela permettrait, dans une phase d'adaptation à l'environnement réel par exemple, au contrôleur d'apprendre en ligne dans le réseau.

5.2.4. Le mode intermédiaire "semi-dynamique" avec l'apprentissage par renforcement

Dans le cas des contrôleurs de logique floue utilisant l'apprentissage par renforcement, nous avons donc imaginé un mode de fonctionnement intermédiaire – "semi *on-line*". Nous supposons tout d'abord avoir calibré (phase d'exploration) à l'aide de simulations logicielles un contrôleur (appelé SIF théorique) fondé sur l'algorithme de *Q-learnin*g comme décrit ci-dessus au paragraphe 5.2.3. Un module de contrôle du paramétrage RRM de type SIF est alors introduit par exemple au niveau de l'OMC pour administrer un ensemble de nodes B. La fréquence des remontées de mesures ainsi que celle des corrections des paramètres RRM vont dépendre des contraintes liées aux équipements du réseau, mais aussi de la politique de l'opérateur (demi-heure, heure, demi-journée, journée). La nouveauté est qu'au lieu d'appliquer le *Q-learning* dans son mode d'exploitation, on continue de l'utiliser en mode d'exploration (en vérifiant que les corrections aléatoires ne soient pas aberrantes). Ainsi à chaque itération, le module continuera à mettre à jour les fonctions de qualité du SIF en tenant compte simultanément des informations des différents nodes B :

- Etat antérieur et présent de la cellule (indicateurs de performances et paramétrage).
- Valeur de renforcement de la cellule.
- Corrections des paramètres RRM appliquées au node B.

Ainsi cette méthode permettra d'adapter aux contraintes du réseau mobile réel (environnement radio et prise en compte de l'ensemble des services et classes de trafic) des contrôleurs calibrés seulement à partir de simulations logicielles. On résume le fonctionnement du contrôleur "semi *on-line*" sur la figure 5.13 suivante.

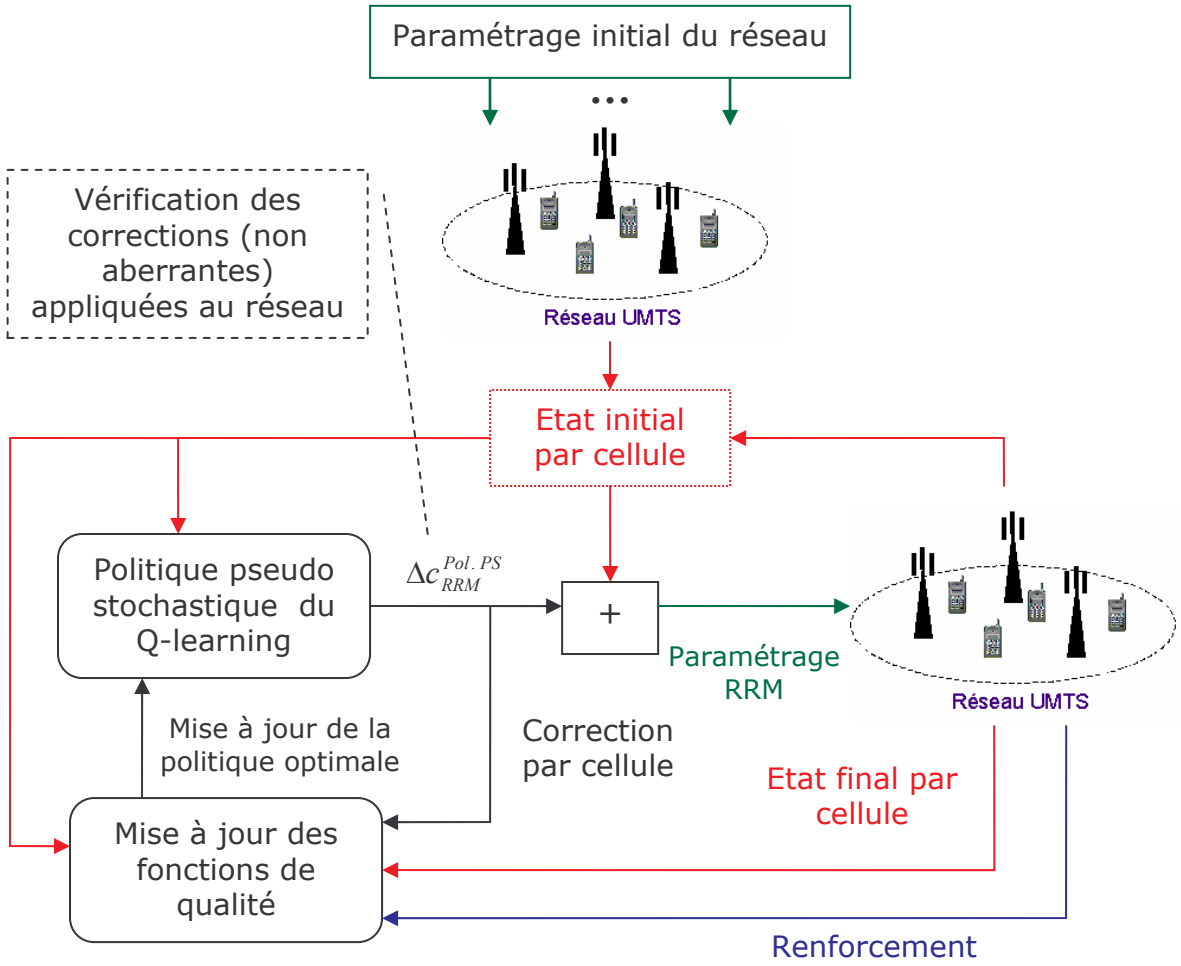


Figure 5.13 : Mode de fonctionnement "semi *on-line*" d'un contrôleur automatique de réseau mobile

Lors de la mise en ligne dans le réseau mobile réel du contrôleur calibré avec des simulations logicielles, nous avons le choix :

- Soit de garder la même structure que le SIF théorique en conservant les valeurs des fonctions de qualité élémentaires estimées à partir des simulations de réseau.
- Soit de restreindre le champ des actions possibles en sortie des règles de décisions élémentaires en introduisant de la connaissance a priori (déduite du SIF théorique) dans le nouveau SIF (comme décrit au paragraphe 4.2.7) mis en ligne dans le réseau mobile réel. Dans ce cas, les fonctions de qualité élémentaires du nouveau SIF sont initialisées à zéro.

Dans les deux cas, il sera essentiel de prévoir un module de vérification qui contrôlera que la correction de paramétrage déterminée par le SIF en environnement réel n'est pas aberrante.

5.3. Conclusions

Dans ce chapitre, nous avons donc introduit et décrit les deux modes de fonctionnement de l'ingénierie automatique des réseaux UMTS. Le mode *on-line* vise à adapter en permanence le réseau aux fluctuations de trafic et aux contraintes d'environnement radio en re-paramétrant

dynamiquement les algorithmes RRM par l'intermédiaires de contrôleurs soit situés au niveau d'un nœud central de management de réseau (comme l'OMC) soit distribués dans le réseau. Le mode *off-line* permet de déterminer les modifications du paramétrage RRM les plus efficaces pour une demande en trafic donnée sur la zone géographique considérée qui pourront être appliquées au niveau des nœuds de management du réseau. Ces adaptations de paramètres pourront dépendre du jour de la semaine et/ou de la période de l'année. Aux chapitres 6 et 7, nous allons tester ces deux modes de fonctionnement à l'aide de notre simulateur de réseau mobile. Enfin, le développement des contrôleurs de logique floue utilisant l'apprentissage par renforcement a permis d'envisager un mode de fonctionnement intermédiaire – "semi *on-line*" qui vise à adapter localement ces contrôleurs, calibrés à l'aide de simulations logicielles, en apprenant directement dans le réseau réel déployé. Les méthodes de contrôle décrites aux chapitres 3 et 4 ont été développées tout d'abord pour l'ingénierie automatique du réseau mobile UMTS.

Néanmoins nous avons commencé à envisager l'adaptation de ces méthodes au contexte de l'optimisation des algorithmes de management des ressources radio dans le cadre des réseaux d'accès radio hétérogènes [Nas05]. Ainsi le projet coopératif EUREKA - CELTIC Gandalf vise à développer des méthodes de surveillance (*monitoring*), d'analyse et de gestion des pannes (*troubleshooting*) et de contrôle automatique des performances et de la qualité de service dans le contexte de réseaux radio multi-systèmes coopératifs GSM/UMTS/WLAN [Stu05]. Sur ce thème, différents scénarios de gestion de la ressource radio conjointe (ou *Joint-RRM* [Agu04]) ont commencé à être étudiés au 3GPP [3GPP1] [3GPP2]. Par exemple, le module *Combined Radio Ressource Management (C-RRM)* permettra la gestion de la mobilité, de la sélection de réseau ou de la répartition de la charge (*load balancing*) inter-systèmes. Ainsi, une première approche (*Multiple Attribute Decision Making* ou MADM) décrite dans [Zha04] définit un cadre de fonctionnement permettant la prise de décisions pour le *hand-over* multi-systèmes, en fonction d'indicateurs de performances des réseaux (*Key Performance Indicators* ou KPI) et d'un vecteur de priorisation entre les systèmes et les services. Une seconde approche décrite dans [Tol02a] [Tol02b] vise à utiliser la matrice de critères pour contrôler le paramétrage des seuils de *hand-over* multi-systèmes, et non pas la prise de décision de resélection de réseau.

Dans le deux cas, un processus d'apprentissage par renforcement adapté (*on-line* ou *off-line*) pourra permettre d'ajuster le vecteur de priorisation entre les services et les systèmes afin d'optimiser les performances des réseaux mobiles et la qualité de service perçue par les utilisateurs.

6. Résultats obtenus avec l'ingénierie automatique en mode *on-line*

Comme indiqué au chapitre 3, l'objectif principal de l'ingénierie automatique est d'améliorer les performances du réseau mobile. Les performances initiales du réseau correspondent dans l'espace des indicateurs de performances[*] à des fronts de paramétrage uniforme, c'est-à-dire les fronts de performances donnés par un paramétrage uniforme et constant des algorithmes RRM. Avec l'adaptation dynamique des paramètres RRM, on souhaite obtenir des performances (pour des niveaux de trafic équivalents) en deçà de ces fronts de paramétrage uniforme, en moyenne dans le réseau et/ou pour les pires stations du réseau (selon la politique de l'opérateur). En effet, du point de vue de l'opérateur, il pourra être nécessaire dans certaines situations de faire un compromis entre l'impact global et l'impact local de l'amélioration des performances du réseau (en particulier pour les pires stations du réseau). On pourra accepter ainsi une légère dégradation des performances moyennes du réseau, à condition d'améliorer nettement les cellules les plus défavorisées. Dans une autre situation, l'ingénierie dynamique améliorera aussi bien les performances moyennes du réseau que celles de des pires stations.

Dans ce chapitre, nous allons appliquer les concepts d'ingénierie dynamique[**] aux algorithmes de contrôle d'admission et de macro-diversité. Précisément nous allons contrôler dynamiquement, node B par node B, les paramètres de seuil de charge cible DL et/ou de *Add_win* au cours de simulations semi-dynamiques de réseau UMTS sur le logiciel développé en interne par France Télécom Recherche & Développement. Dans les environnements considérés, les contrôleurs de logique floue (CLF) introduits au chapitre 3 vont être optimisés par optimisation par essaim de particules (OEP). En particulier, nous vérifierons que le contrôle dynamique et simultané des algorithmes d'admission et de macro-diversité est plus efficace que le contrôle seul de l'admission. L'OEP sera aussi testée dans les modes mono et multi-objectif. Le front optimal du nuage des solutions explorées en multi-objectif sera ainsi comparé avec les meilleures solutions obtenues en mono-objectif. En fin de chapitre, on fournira aussi les premiers résultats d'ingénierie dynamique optimisée par apprentissage par renforcement. Pour tester ces méthodes d'optimisation, deux environnements sont utilisés : un petit réseau (21 secteurs) de la ville 1 et un réseau plus grand (109 secteurs) sur la ville 2. Dans ces deux zones, on définira respectivement un objet de trafic inhomogène (Ville 1) et homogène (Ville 2). Plusieurs tests permettront de mettre à l'épreuve l'efficacité des CLF :

- Des contrôleurs optimisés par OEP pour un certain niveau de trafic restent-ils efficaces même si l'on fait varier le niveau de trafic sur la zone ?
- Des contrôleurs optimisés par OEP sur la région de la ville 1 seront directement utilisés et testés sans adaptation dans le réseau de la ville 2. Restent-ils performants, même si l'environnement radio et la distribution du trafic sont différents ? Cela permettrait d'utiliser dans un premier temps dans les réseaux réels des contrôleurs automatiques génériques par type d'environnement (dense urbain, urbain, suburbain, rural) avant d'envisager à plus long terme de réellement les adapter localement.
- Sur la zone de la ville 1, les meilleurs contrôleurs obtenus par OEP en mono-objectif seront comparées aux solutions explorées en multi-objectifs par le même algorithme.
- Sur la zone de la ville 2, on comparera ensuite les performances d'un contrôleur optimisé par OEP sur la ville 2 avec celles d'un contrôleur optimisé sur la ville 1.

Nous mettrons ainsi en œuvre le schéma d'application de l'optimisation des CLF par essaim de particules décrit par la figure 4.4 du paragraphe 4.1.4

---

[*] Plan taux de coupure – taux de blocage, par exemple.
[**] Ou mode *on-line.*

6.1. Premiers résultats d'optimisation de l'ingénierie dynamique par OEP

Nous considérons tout d'abord le cas d'un trafic inhomogène sur un réseau de petite taille (7 sites, 21 secteurs, voir figure 6.1) dans la région de la ville 1. En effet, on commence par l'optimisation de l'ingénierie dynamique par essaim de particules sur un petit réseau pour éviter des temps de calcul trop importants. De plus, le trafic défini en entrée est inhomogène et élevé (charge moyenne* de 42 communications voix par secteur sur la zone) car on pressent que c'est dans de tels cas que l'ingénierie dynamique sera la plus efficace.



Figure 6.1 : Réseau sur la zone de la ville 1.

Dans ce chapitre, nous commençons donc par déterminer les fronts de paramétrage uniforme (fronts P.U.) dans le plan taux de coupure – taux de blocage correspondant aux performances du réseau pour différents paramétrages uniformes du contrôle d'admission sur l'ensemble des stations. A partir d'un même réseau de départ nous allons lancer successivement des optimisations par OEP du contrôle d'admission puis de la macro-diversité, pour différents coefficients de mixage en mono-objectif. Nous comparons alors l'efficacité des politiques ainsi optimisées avec les résultats obtenus pour un paramétrage uniforme (fronts P.U.).

6.1.1. Détermination des fronts de paramétrage uniforme

Les fronts de paramétrage uniforme seront la référence en termes de performances pour le taux de coupure / taux de blocage moyennés (obtenus à l'aide d'un paramétrage uniforme du contrôle d'admission et de la macro-diversité). Les fronts P.U. correspondent aux moyennes des taux de blocage et taux de coupure sur l'ensemble des stations, pour les 10% pires et pour les 5% pires stations dans le plan coupure / blocage. Dans le cas suivant (figure 6.2), le seuil du contrôle d'admission DL, *downlink load target threshold*, varie dans l'ensemble {0.5, 0.6, 0.7, 0.8, 0.825} pour l'ensemble des stations. Pour un paramétrage uniforme du réseau, le meilleur compromis initial en termes de taux de coupure et taux de blocage a été obtenu sur la figure 6.2 pour une valeur de 0.7 du seuil de charge cible DL (pour une charge moyenne de 42 communications voix par secteur).

* La charge totale moyenne est égale au produit du taux d'arrivée total moyen sur la zone (égal 9 mobiles voix par seconde répartis de façon non homogène) par la durée moyenne des communications (égale à 100s). La charge moyenne par secteur est estimée en divisant par le nombre de secteurs.

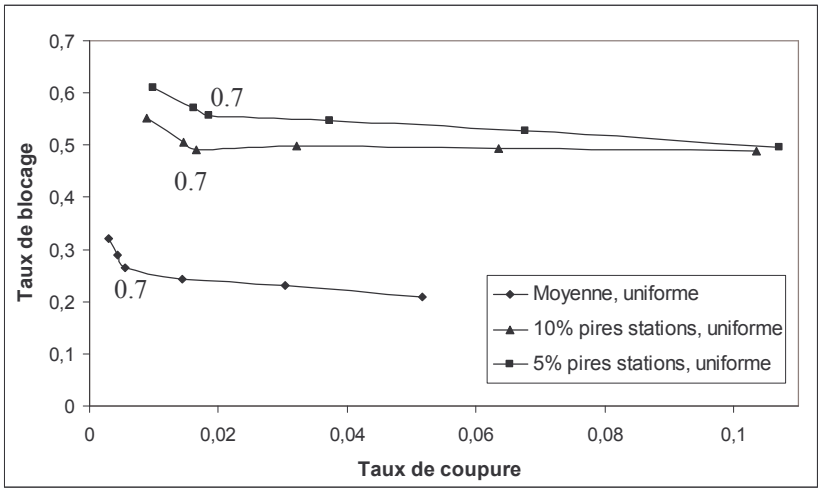


Figure 6.2 : Fronts P.U. (moyenne des stations, des 10% pires et des 5% pires stations) dans le plan taux de coupure / taux de blocage.

6.1.2. Optimisation de l'ingénierie dynamique pour le contrôle d'admission

6.1.2.1 Tests de différents coefficients de mixage lors de l'optimisation de l'ingénierie dynamique pour le contrôle d'admission

Comme introduit au paragraphe 4.1.3.2, le coefficient de mixage* permet en mono-objectif de calculer la fonction coût globale par station :

Fonction coût globale = Coût blocage + Coefficient de Mixage * Coût coupure (6.1)

Pour avoir un point de référence, nous considérons la politique initiale d'ingénierie dynamique, correspondant au paramétrage par défaut du CLF, à partir de laquelle va être lancée l'OEP en mono-objectif. La figure 6.3 illustre les performances de cette politique initiale (donnée par le paramétrage initial du CLF avant optimisation) en termes de taux de blocage et taux de coupure après simulation de l'ingénierie dynamique avec notre simulateur de réseau mobile UMTS (pour une charge moyenne de 42 mobiles voix par secteur).

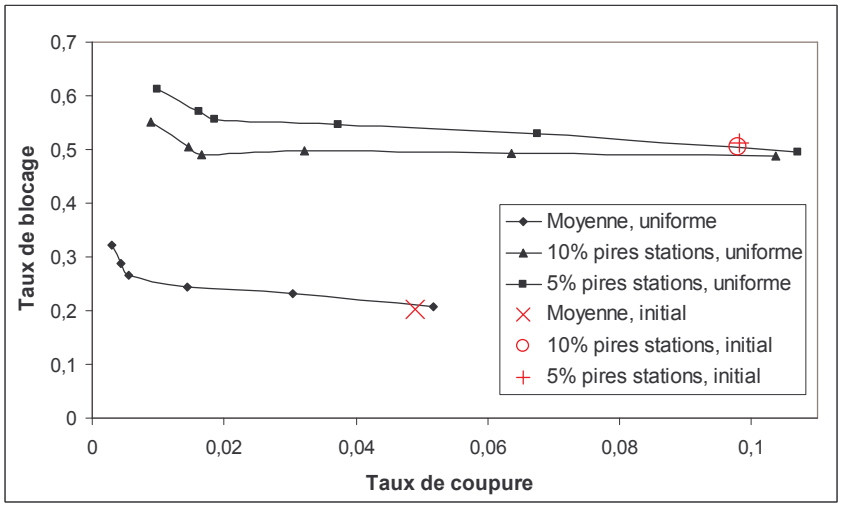


Figure 6.3 : Performances de la politique initiale (moyenne des stations, des 10% pires et des 5% pires stations) dans le plan taux de coupure / taux de blocage.

* En général dans les réseaux mobiles, le coefficient de mixage est fixé dans l'intervalle *[5, 10]*.

A partir de cette politique initiale, on lance une OEP en mono-objectif pour différents coefficients de mixage *{2.5, 3.5, 5, 10}*. Pour chaque coefficient de mixage, on sélectionne la meilleure politique obtenue par OEP en terme de fonction coût (sauf pour le coefficient 5, pour lequel deux politiques optimales *5.I* et *5.II* sont sélectionnées). La figure 6.4 illustre les performances des politiques sélectionnées pour chaque coefficient de mixage en termes de taux blocage et de taux de coupure après simulation de l'ingénierie dynamique avec notre simulateur de réseau UMTS (pour une charge moyenne de 42 mobiles voix par secteur) :

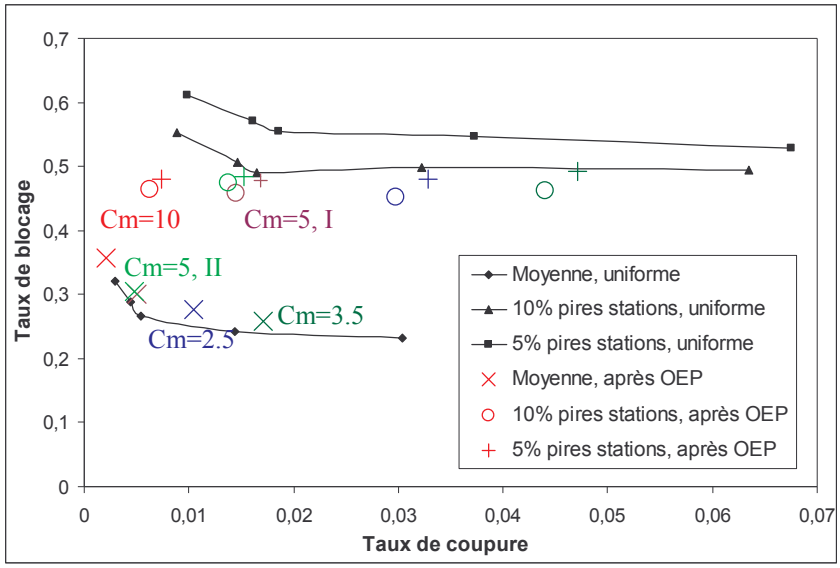


Figure 6.4 : Performances des politiques sélectionnées (moyenne des stations, des 10% pires et des 5% pires stations) optimisées pour le contrôle d'admission* par essaim de particules pour les différents coefficients de mixage dans le plan taux de coupure / taux de blocage.

Chaque politique optimisée améliore les performances en taux de blocage et de coupure par rapport aux fronts P.U. pour les 10 % et 5 % pires stations du réseau. En revanche, il y a une légère dégradation des performances moyennes par rapport au front P.U. moyenné sur l'ensemble des stations. Ce phénomène n'est pas surprenant, car la fonction coût choisie (voir le paragraphe 6.1.2.2) privilégie nettement l'optimisation des pires stations du réseau. Par ailleurs, le choix du coefficient de mixage permet d'obtenir des compromis différents en termes de taux de coupure et taux de blocage moyens, sur les 10% pires et les 5% pires stations du réseau. Typiquement sur la figure 6.4, si on considère le paramétrage uniforme du seuil de charge cible DL à 0.7 comme étant la meilleure situation de départ en terme de compromis coupure - blocage, les deux politiques *5.I* et *5.II*, optimisées avec un coefficient de mixage de 5, sont les plus efficaces pour cette zone de fonctionnement.

6.1.2.2 Remarques sur la fonction coût

La fonction coût (introduite au paragraphe 4.1.3.2) utilisée dans l'optimisation par essaim de particules effectuée au paragraphe précédent se concentre sur les "pires stations" du réseau :

Fonction coût blocage = 10*taux blocage sur les 5% pires + 3* taux blocage sur les 5-10% pires+1* taux blocage sur les autres stations

La fonction coût de coupure est similaire et donc se focalise aussi sur les "pires stations" du réseau. En conséquence, il apparaît logique que la moyenne globale sur le réseau des performances des CLF optimisés ait été pénalisée au cours du processus d'optimisation et soit

* Sur la figure 6.4, les politiques optimisées pour le contrôle d'admission sont indiquées par Cm =… (Coefficient de mixage).

légèrement moins performante que le front P.U correspondant. En essayant différentes fonctions coûts, il sera possible de moins pénaliser la moyenne des indicateurs de qualité par rapport aux pires stations du réseau. Par ailleurs, avec un réseau de petite taille, une telle fonction coût est pénalisante dans le cadre d'un processus d'optimisation, car statistiquement elle ne permet pas d'avoir toujours des résultats d'optimisation très fiables sur l'échantillon réduit des pires stations du réseau: la politique est optimisée en tenant compte de manière privilégiée du comportement de quelques stations du réseau.

6.1.3. Optimisation de l'ingénierie dynamique pour la macro-diversité

A partir des contrôleurs optimisés (obtenus au paragraphe 6.1.2.1) pour la gestion du paramètre de seuil de charge cible DL et pour chaque coefficient de mixage considéré, on souhaite à présent appliquer l'OEP de l'ingénierie dynamique en mono-objectif au contrôle du paramètre de macro-diversité *Add_win*. Pour chaque coefficient de mixage, la (ou les) solution(s) optimisée(s) pour le seuil de charge cible DL (par exemple sur la figure 6.4, la politique *5.I*) est prise comme politique initiale. On fixe alors le même coefficient de mixage avant de lancer l'optimisation par essaim de particules de la politique de gestion dynamique (ou autonome) du paramètre *Add_win* (exemple sur la figure 6.5 : la politique *5.I+MD*). Dans le paragraphe 6.1, les CLF pour le contrôle d'admission et la macro-diversité auront donc été optimisés séquentiellement par deux utilisations successives de l'algorithme d'OEP.

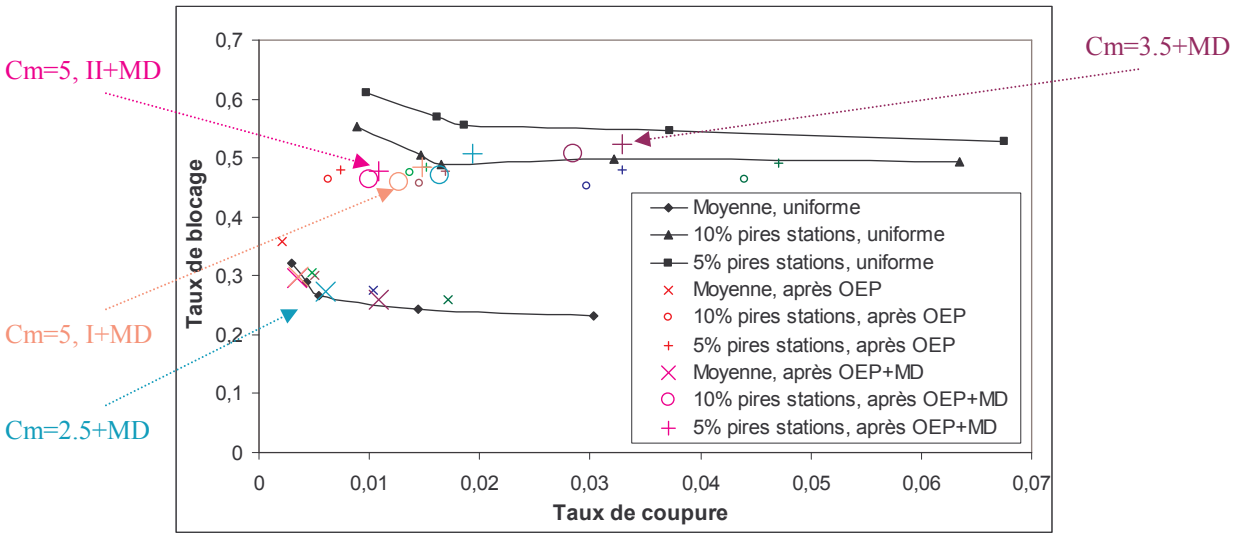


Figure 6.5 : Performances des politiques (moyenne des stations, des 10% pires et des 5% pires stations) optimisées par essaim de particules successivement pour le contrôle d'admission* et la macro-diversité** pour différents coefficients de mixage dans le plan taux de coupure / taux de blocage.

Pour chaque coefficient de mixage, l'optimisation du contrôle dynamique du paramètre *Add_win* de l'algorithme de macro-diversité a permis d'améliorer les performances du réseau par rapport à la situation antérieure où seulement le contrôle du paramètre de charge cible DL avait été optimisé, sauf pour les coefficients 3.5 et 2.5 où l'amélioration du taux de coupure s'est accompagnée d'une faible dégradation du taux de blocage. De plus, pour les deux contrôleurs optimisés avec un coefficient de mixage égal à 5, les politiques *5.I* et *5.II*, les performances ont été améliorées pour les taux de coupure et de blocage aussi bien pour la moyenne, que pour les 10% et les 5% pires stations du réseau (pour ces 3 indicateurs, les 2 politiques sont sous le front P.U. moyenné sur l'ensemble des stations).

---

* Sur la figure 6.4, les politiques optimisées pour le contrôle d'admission sont indiquées par Cm =… (Coefficient de mixage).

** Sur la figure 6.5, les pol. optimisées pour le contrôle d'admission puis la macro-diversité sont indiquées par Cm =…+MD.

6.2. Applications des politiques optimisées par OEP à un réseau différent

Nous allons maintenant considérer un réseau de grande taille (37 sites, 109 secteurs, voir figure 6.6) sur une partie de la ville 2. Comme au paragraphe 6.1, nous déterminerons tout d'abord les fronts P.U. pour un trafic homogène et différents paramétrages uniformes du contrôle d'admission et de la macro-diversité. Sur ce réseau de départ, nous allons tester l'efficacité des politiques optimisées (pour le contrôle d'admission seul ainsi que séquentiellement* pour le contrôle d'admission puis la macro-diversité comme décrit au 6.1.3) par OEP en mono-objectif au paragraphe 6.1 sur un réseau de petite taille de la ville 1. Par la suite, nous étudierons la réponse de ces différentes politiques à une variation de la densité de trafic en entrée (le taux moyen d'arrivée des appels variant de 20 à 30 mobiles par seconde pour une durée moyenne des communications de 100s, la charge moyenne variant donc de 18 à 27 communications voix par secteur sur la zone). Enfin nous comparerons les performances des politiques optimisées sur le réseau de la ville 1 avec celles d'une politique directement optimisée par OEP (simultanément** pour le contrôle d'admission et la macro-diversité dans ce cas) en mono-objectif sur le réseau de la ville 2 pour la même fonction coût et un coefficient de mixage égal à 5.

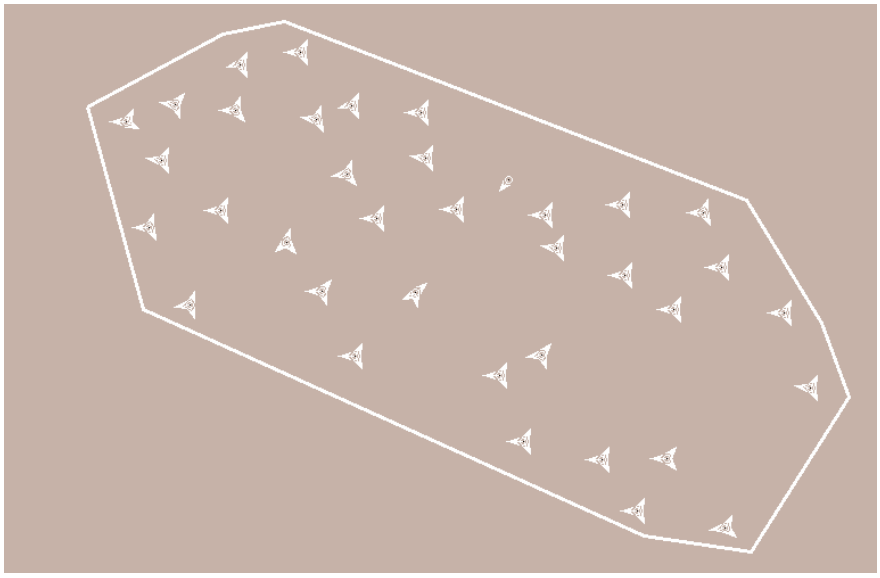


Figure 6.6 : Réseau sur la zone de la ville 2.

6.2.1. Applications des politiques optimisées pour le contrôle d'admission

6.2.1.1 Détermination des fronts P.U. sur la zone de la ville 2

Les fronts de paramétrage uniforme sont la référence des performances en termes de coupure / blocage (obtenus à l'aide d'un paramétrage uniforme du contrôle d'admission et de la macro-diversité sur l'ensemble des stations du réseau). Les fronts P.U. correspondent aux moyennes des taux de blocage et taux de coupure sur l'ensemble des stations, pour les 10% pires et pour les 5% pires stations dans le plan coupure / blocage. Sur la figure 6.7, le seuil du contrôle d'admission DL, *downlink load target threshold*, varie dans l'ensemble {0.6, 0.65, 0.7, 0.8, 0.825, 0.85}, le trafic en entrée étant uniforme et de densité constante sur la zone de la ville 2

* Les CLF ont été optimisés par deux utilisations successives de l'algorithme d'OEP pour le contrôle d'admission (C.A.) puis pour la macro-diversité (M.D.).

** Les CLF sont optimisés par une seule utilisation de l'algorithme d'OEP simultanément pour le C.A. et la M.D.

considérée (le taux moyen d'arrivée des appels étant fixé à 25 et la durée moyenne des communications à 100s, la charge moyenne est donc de 23 mobiles voix par secteur).

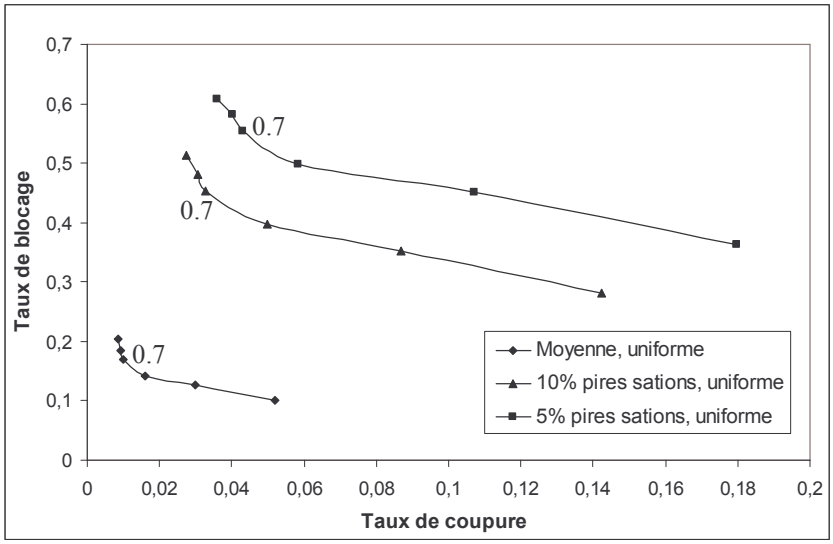


Figure 6.7 : Fronts P.U. (moyenne des stations, des 10% pires et des 5% pires stations) dans le plan taux de coupure / taux de blocage.

Pour un paramétrage uniforme du réseau, le meilleur compromis en termes de taux de coupure et taux de blocage a été obtenu pour un seuil de charge cible DL entre 0.7 et 0.8.

6.2.1.2 Tests sur le réseau de la ville 2 des contrôleurs optimisés par OEP sur la ville 1

Pour chaque coefficient de mixage {2.5, 3.5, 5}, on teste sur le réseau de la ville 2 la politique la plus efficace (pour la gestion du contrôle d'admission) optimisée par OEP en mono-objectif sur la ville 1. La figure 6.8 illustre les performances sur la ville 2 de chaque politique en termes de taux de blocage et de coupure après simulations de l'ingénierie dynamique avec notre simulateur de réseau UMTS (la charge moyenne étant de 23 mobiles voix par secteur) :

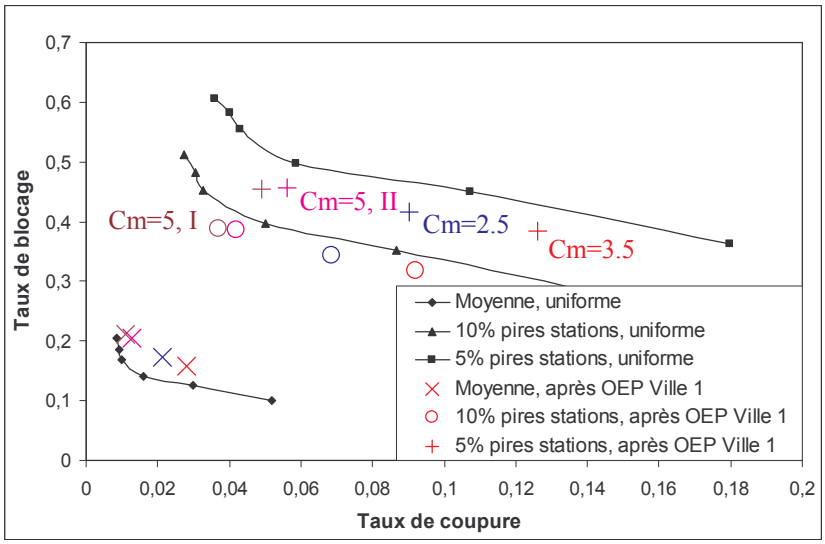


Figure 6.8 : Performances des politiques (moyenne des stations, des 10% pires et des 5% pires stations) optimisées sur la zone de la ville 1 par essaim de particules pour le contrôle d'admission* et pour différents coefficients de mixage (*Cm*) dans le plan taux de coupure / taux de blocage.

* Sur la figure 6.8, les politiques optimisées pour le contrôle d'admission sont indiquées par Cm =… (Coefficient de mixage).

Chaque politique optimisée sur la zone de la ville 1 améliore les performances sur la zone de la ville 2 par rapport aux fronts P.U. des 10% et 5% pires stations du réseau. En revanche, on constate une légère dégradation des performances moyennes par rapport au front P.U. moyenné sur l'ensemble des stations. Ce phénomène n'est pas surprenant, car la fonction coût (voir paragraphe 6.1.2.2) privilégie nettement l'optimisation des pires stations du réseau.

Par ailleurs, le choix du coefficient de mixage permet d'obtenir différents compromis en termes de taux de coupure et taux de blocage moyens, sur les 10% pires et les 5% pires stations du réseau. Typiquement sur la figure 6.8, si on considère un paramétrage uniforme du seuil de charge cible DL entre 0.7 et 0.8 comme étant la meilleure situation de départ en terme de compromis coupure - blocage, les deux politiques *5.I* et *5.II*, optimisées avec un coefficient de mixage égal à 5, sont les plus adaptées à cette zone de fonctionnement.

6.2.2. Applications des politiques optimisées pour le contrôle d'admission et la macro-diversité

Pour chaque coefficient de mixage {2.5, 3.5, 5}, on teste sur le réseau de la ville 2 la politique la plus efficace optimisée par OEP (séquentiellement pour le contrôle d'admission puis la macro-diversité comme décrit au 6.1.3) en mono-objectif sur une zone de petite taille de la ville 1. La figure 6.9 illustre les performances de chaque politique en termes de taux de blocage et taux de coupure après simulations de l'ingénierie dynamique avec notre simulateur de réseau UMTS (la charge moyenne étant de 23 mobiles voix par secteur). Il est important de noter que le coefficient de mixage a été conservé entre les deux optimisations successives.

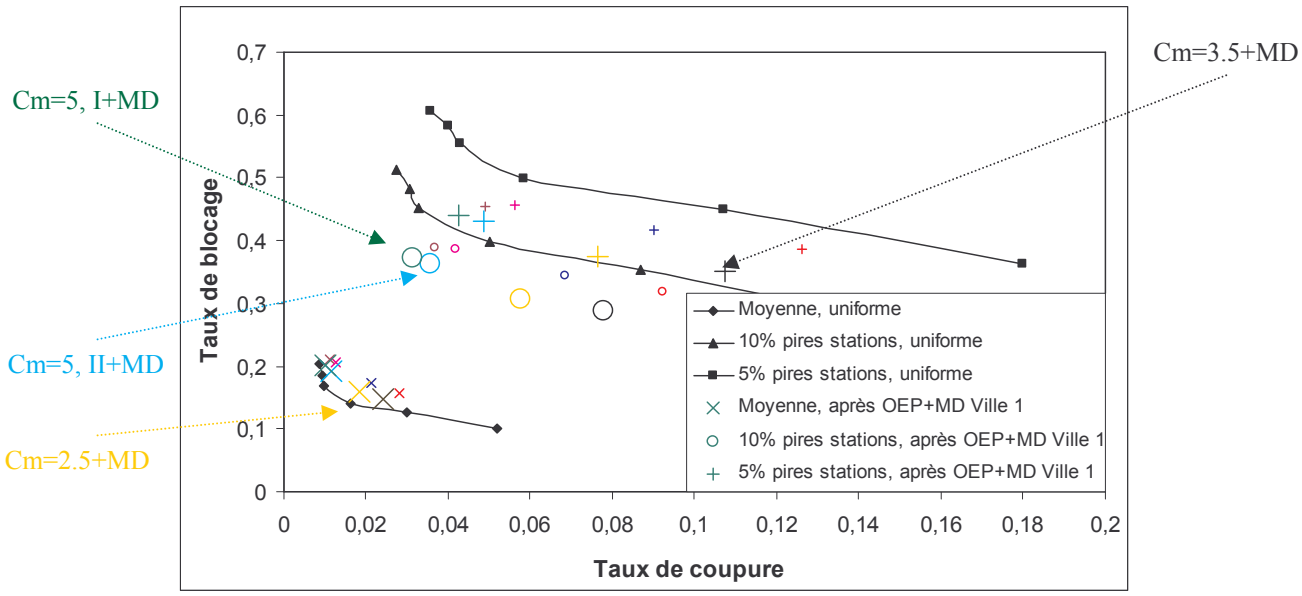


Figure 6.9 : Performances des politiques (moyenne des stations, des 10% pires et des 5% pires stations) optimisées sur la zone de la ville 1 par OEP successivement pour le contrôle d'admission* et la macro-diversité** et pour différents coefficients de mixage dans le plan taux de coupure / blocage.

Pour chaque coefficient de mixage, les politiques optimisées sur un petit réseau de la ville 1 successivement pour le contrôle d'admission et la macro-diversité ont permis d'améliorer les performances sur la zone de la ville 2 en termes de taux de coupure et taux de blocage (pour les moyennes sur l'ensemble des stations, des 10% pires et des 5% pires stations) par rapport à la situation antérieure (cf. figure 6.8) où seulement le paramètre de charge cible DL avait été

* Sur la figure 6.8, les politiques optimisées pour le contrôle d'admission sont indiquées par Cm =… (Coefficient de mixage).
** Sur la figure 6.9, les pol. optimisées pour le contrôle d'admission puis la macro-diversité sont indiquées par Cm =… +MD.

optimisé. Par rapport aux fronts P.U., ces politiques permettent d'améliorer nettement les performances moyennes des 10% et 5% pires stations du réseau, mais aucune ne permet de franchir le front P.U. moyenné sur l'ensemble des stations. Ce fait est cohérent avec la fonction coût qui avait été choisie lors du processus d'optimisation sur la zone de la ville 1.

Globalement ces résultats sont encourageants, car des politiques optimisées pour un "petit réseau", avec un trafic inhomogène, ont donné de bonnes performances en termes de blocage et coupure sur un "réseau de plus grande taille" dans un environnement urbain complètement différent et pour un trafic homogène. Ainsi cela pourrait permettre de présélectionner selon l'environnement (urbain, suburbain, rural) des politiques optimisées sur des petits réseaux et de les tester voire les optimiser sur des grands réseaux. On pourrait ainsi gagner un temps considérable de simulations aussi bien en mode *off-line* (en présélectionnant les politiques candidates) qu'en mode *on-line* (en partant d'un paramétrage de départ favorable).

6.2.3. Variations de l'efficacité des contrôleurs en fonction de la densité de trafic

6.2.3.1 Détermination des fronts P.U. sur la zone de la ville 2 pour différentes densités de trafic

Dans ce paragraphe, on considère toujours le réseau de 37 sites et 109 secteurs sur la zone de la ville 2. Pour chaque paramétrage uniforme et chaque politique optimisée, on fait varier la densité de trafic en entrée (le nombre de mobiles présents au tirage initial variant de 2000 à 3000 sur l'ensemble du réseau, le taux moyen d'arrivée des appels étant fixé proportionnellement entre 20 et 30 et la durée de communication moyenne étant fixée à 100s). Ci-dessous, on trace sur la figure 6.10, les fronts de paramétrage uniforme obtenus en moyenne sur l'ensemble des stations, pour les 10% pires et les 5% pires stations du réseau, le paramètre charge cible DL étant pris dans {0.7, 0.8, 0.825}, le nombre de mobiles présents au tirage initial (et proportionnellement le taux moyen d'arrivée des appels) variant dans l'ensemble {2000, 2500, 3000}, et la charge moyenne de 18 à 27 mobiles voix par secteur :

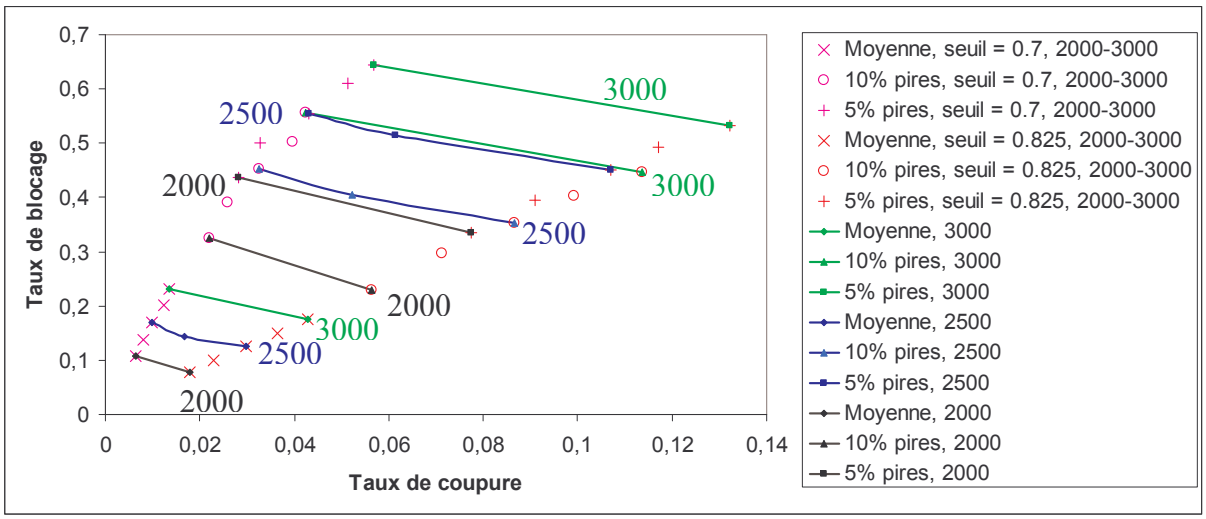


Figure 6.10 : Fronts P.U. (moyenne des stations, des 10% pires et des 5% pires stations) dans le plan taux de coupure / taux de blocage, le paramètre charge cible DL variant dans {0.7, 0.8, 0.825} et le nombre initial de mobiles dans le réseau dans l'ensemble {2000, 2500, 3000}.

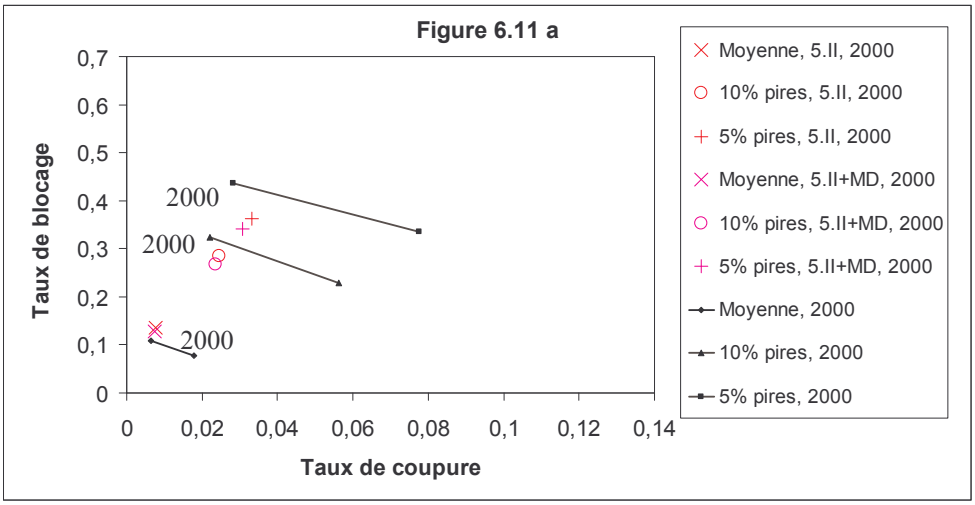
Avec l'augmentation de la densité de trafic, les fronts P.U. (pour des charges cibles entre 0.7 et 0.825) sont naturellement décalés vers des taux de blocage et de coupure plus élevés. On souhaite ensuite estimer la robustesse des performances de deux solutions optimisées, la *5.II* (pour le contrôle d'admission) et la *5.II+MD* (pour le contrôle d'admission et la macro-

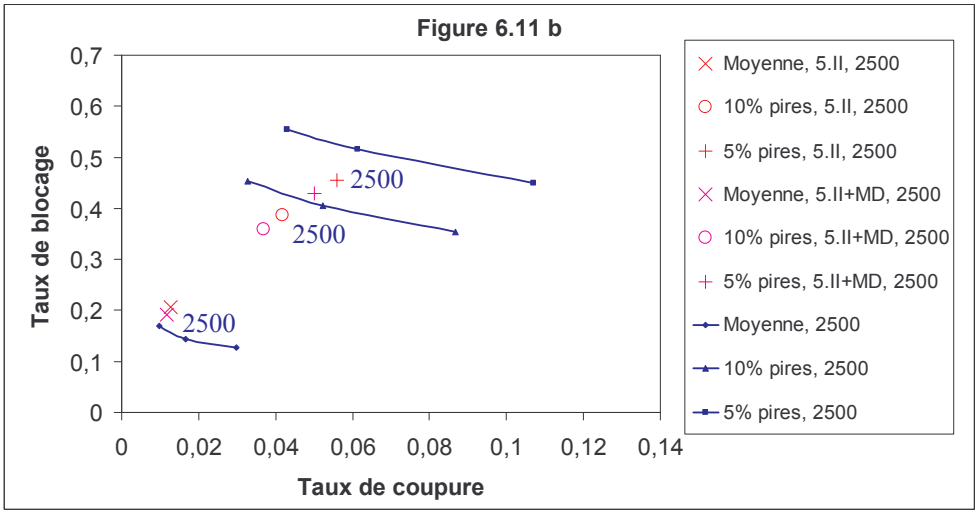
diversité), par rapport aux fronts P.U. tracés ci-dessus, quand le nombre initial de mobiles varie entre 2000 et 3000 (le taux moyen d'arrivée des appels variant proportionnellement).

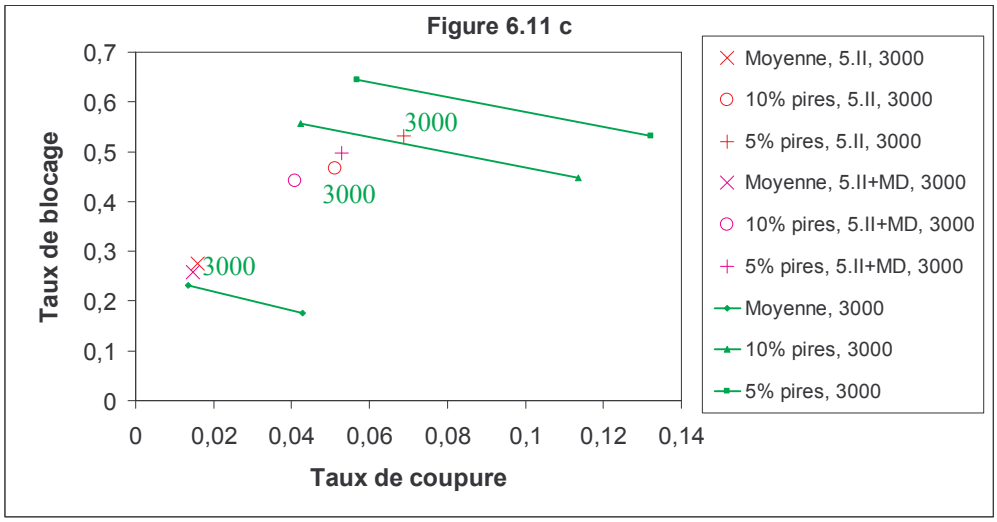
6.2.3.2 Tests pour différentes densités de trafic sur le réseau de la ville 2 de deux contrôleurs optimisés sur la ville 1

Nous étudions ci-dessous la réponse des politiques *5.II* et *5.II+MD* à une variation de la densité de trafic en entrée (variant entre 2000 et 3000 sur l'ensemble du réseau, le taux moyen d'arrivée des appels étant fixé proportionnellement).

Rappel : La politique *5.II* a été obtenue par optimisation (par OEP en mono-objectif) du contrôle du paramètre de seuil de charge cible DL (voir paragraphe 6.1.2.1) et la politique *5.II+MD* par optimisations successives (par OEP en mono-objectif) du contrôle du paramètre de seuil de charge cible DL et du paramètre *Add_win* de macro-diversité, sur un petit réseau de la zone de la ville 1 et pour un coefficient de mixage égal à 5. A partir du graphe des fronts P.U. (figure 6.10), on trace sur les figures 6.11, les résultats obtenus dans le plan taux de coupure – taux de blocage pour les politiques *5.II* et *5.II+MD* en moyenne sur l'ensemble des stations, pour les 10% pires et les 5% pires stations du réseau, la densité de trafic (et proportionnellement le taux moyen d'arrivée des appels) variant dans l'ensemble {2000, 2500, 3000}, et donc la charge moyenne de 18 à 27 mobiles voix par secteur.







Figures 6.11 : Performances des politiques *5.II* et *5.II+MD* (moyenne des stations, des 10% pires et des 5% pires stations) optimisées sur la ville 1 par essaim de particules pour un coefficient de mixage égal à 5 dans le plan taux de coupure / taux de blocage, le nombre initial de mobiles dans le réseau variant dans {2000, 2500, 3000}.

Dans tous les cas, les performances de la politique *5.II+MD* sont meilleures aussi bien en taux de blocage qu'en taux de coupure que celles de la politique *5.II*, la densité de trafic variant dans l'ensemble {2000, 2500, 3000}. Donc les optimisations successives par OEP du contrôle de l'admission et de la macro-diversité fournissent des contrôleurs dynamiques plus efficaces, ce résultat étant stable pour une densité de trafic variant dans une certaine plage [2000, 3000] autour de la densité moyenne de 2500. Par rapport aux fronts P.U., les deux politiques de contrôle dynamique donnent de meilleures performances en taux de blocage et taux de coupure qu'un paramétrage uniforme pour les 10% pires et les 5% pires stations du réseau. Cependant les performances moyennées sur l'ensemble des stations sont légèrement moins bonnes pour les politiques *5.II* et *5.II+MD* que pour un paramétrage uniforme à 0.7 pour le paramètre de seuil de charge cible DL. Ce résultat n'est pas surprenant car la fonction coût privilégie nettement l'optimisation des pires stations du réseau (voir paragraphe 6.1.2.2).

Dans le paragraphe suivant, on se propose d'optimiser, directement sur le grand réseau de la ville 2, par essaim de particules en mono-objectif, le contrôle dynamique de l'admission et de la macro-diversité. Du fait de la taille du réseau (109 secteurs), cette optimisation des contrôleurs de logique floue est beaucoup longue en temps de calculs. Cependant cela permettra de comparer les performances entre les politiques obtenues précédemment (par OEP sur la zone de ville 1) et une politique directement optimisée (à la fois pour l'admission et la macro-diversité) sur le réseau de la ville 2. Nous ferons ensuite varier les niveaux de trafic en entrée {2000, 2500, 3000} pour comparer la stabilité des performances entre les politiques précédentes et celle directement optimisée sur la zone de la ville 2.

6.2.4. Performances d'une politique directement optimisée sur le réseau de la ville 2

Dans ce paragraphe, nous allons donc optimiser par OEP en mono-objectif directement sur la zone de la ville 2 simultanément le contrôle dynamique de l'admission et de la macro-diversité pour un coefficient de mixage égal à 5 (la charge moyenne étant de 23 mobiles voix par secteur). Nous comparerons alors l'efficacité de la politique ainsi optimisée avec les résultats obtenus pour un paramétrage uniforme (fronts P.U.) et pour les politiques précédemment optimisées sur la zone de la ville 1 (paragraphes 6.2.1 et 6.2.2). Par la suite, nous étudierons la

réponse de la nouvelle politique optimisée sur le réseau de la ville 2 à une variation de la densité de trafic en entrée (le taux moyen d'arrivée des appels variant de 20 à 30 mobiles par seconde pour une durée moyenne des communications de 100s, la charge moyenne variant donc de 18 à 27 communications voix par secteur sur la zone).

6.2.4.1 Optimisation de l'ingénierie dynamique du contrôle d'admission et de la macro-diversité sur la ville 2

En choisissant la politique *5.II+MD* comme paramétrage par défaut du CLF et à partir d'un état de référence à 0.7 pour la charge cible DL, on lance une OEP* (simultanément** pour le contrôle d'admission et la macro-diversité) en mono-objectif avec un coefficient de mixage égal à 5 sur la zone de la ville 2. On sélectionne la politique optimisée la plus efficace en terme de fonction coût. On compare avec les fronts de paramétrage uniforme obtenus au paragraphe 6.2.1.1 (figure 6.7) et aux performances (figures 6.8 et 6.9) des politiques *2.5* et *5.II* (pour le contrôle d'admission) et des politiques *2.5+MD* et *5.II+MD* (pour le contrôle d'admission et la macro-diversité) optimisées sur la ville 1. La figure 6.12 illustre les performances de la politique optimisée sélectionnée pour un coefficient de mixage égal à 5 en termes de taux de blocage et taux de coupure après simulation de l'ingénierie dynamique avec notre simulateur de réseau UMTS (la charge moyenne étant de 23 mobiles voix par secteur) :

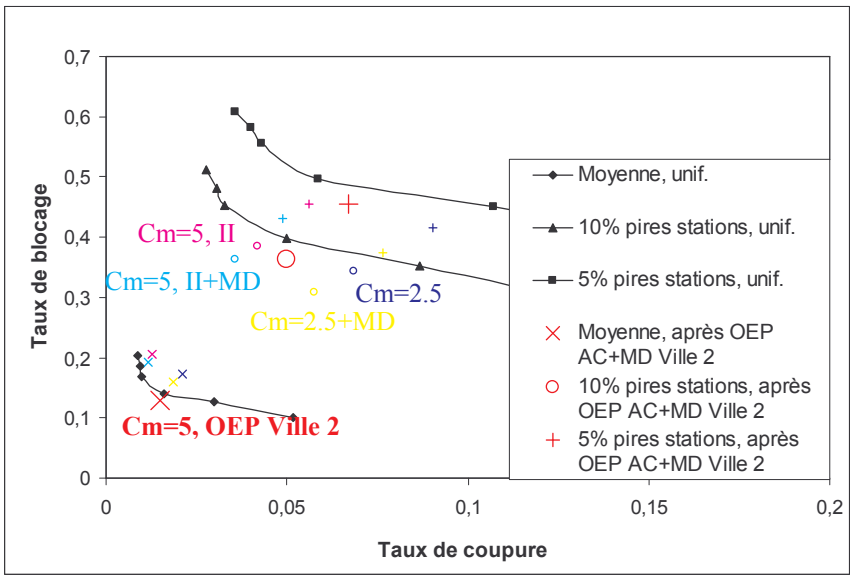


Figure 6.12 : Performances de la politique (moyenne des stations, des 10% pires et des 5% pires stations) optimisée sur la zone de la ville 2 par OEP pour le contrôle d'admission et la macro-diversité et comparaisons avec les politiques *2.5*, *2.5+MD*, *5.II* et *5.II+MD* dans le plan taux de coupure / taux de blocage.

La politique *5 Opt. Ville 2* optimisée sur la zone de la ville 2 avec un coefficient de mixage égal à 5 améliore les performances par rapport aux fronts P.U. moyen sur l'ensemble des stations, des 10% et 5% pires stations du réseau. Par rapport à la politique *5.II+MD*, la politique *5 Opt. Ville 2* donne des performances moins bonnes en taux de blocage et taux de coupure pour les 10% pires et les 5% pires stations du réseau. En revanche, les performances moyennes sur l'ensemble des stations de la politique *5 Opt. Ville 2* sont meilleures que celle de la politique *5.II+MD*. Ces résultats pourraient s'expliquer par le fait que la politique

* On a utilisé la même fonction coût qu'au paragraphe 6.1.2.2, par exemple pour le blocage: Fonction coût blocage = 10*taux blocage sur les 5% pires + 3* taux blocage sur les 5-10% pires+1* taux blocage sur les autres stations.

** Les CLF sont optimisés par une seule utilisation de l'algorithme d'OEP simultanément pour le C.A. et la M.D.

5.II+MD a été optimisée sur un réseau avec un trafic inhomogène alors que la politique *5 Opt. Ville 2* a été optimisée sur un réseau avec un trafic homogène (dans les deux cas on utilise la même fonction coût). En conséquence, la politique *5.II+MD* semble être plus adaptée à l'amélioration des performances des pires stations du réseau, alors que la politique *5 Opt. Ville 2* privilégie une amélioration moyenne des performances.

Performances sur l'ensemble des stations:

Sur les figures 6.13, 6.14 et 6.15, on a classé par ordre décroissant les performances moyennes par station respectivement en termes de taux de blocage, taux de coupure et taux de blocage des liens de macro-diversité pour les politiques uniforme, *5 Opt. Ville 2*, *5.II* et *5.II+MD*:

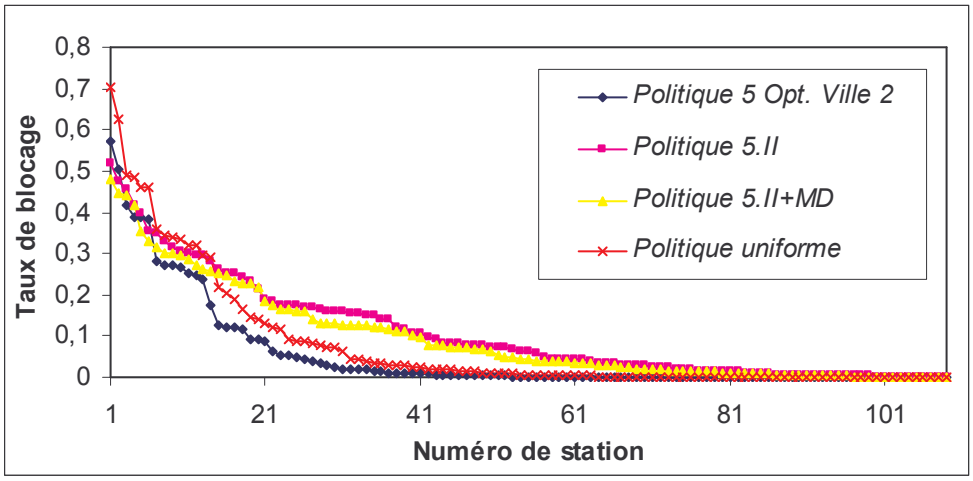


Figure 6.13 : Taux de blocage moyen par station (rangé dans l'ordre décroissant) pour les politiques uniforme, *5 Opt. Ville 2*, *5.II* et *5.II+MD* sur la zone de la ville 2.

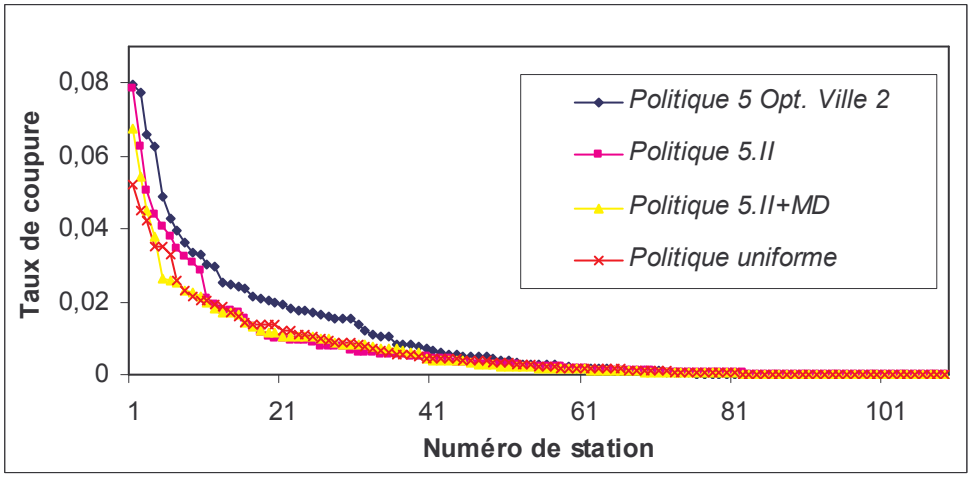


Figure 6.14 : Taux de coupure moyen par station (rangé dans l'ordre décroissant) pour les politiques uniforme, *5 Opt. Ville 2*, *5.II* et *5.II+MD* sur la zone de la ville 2.

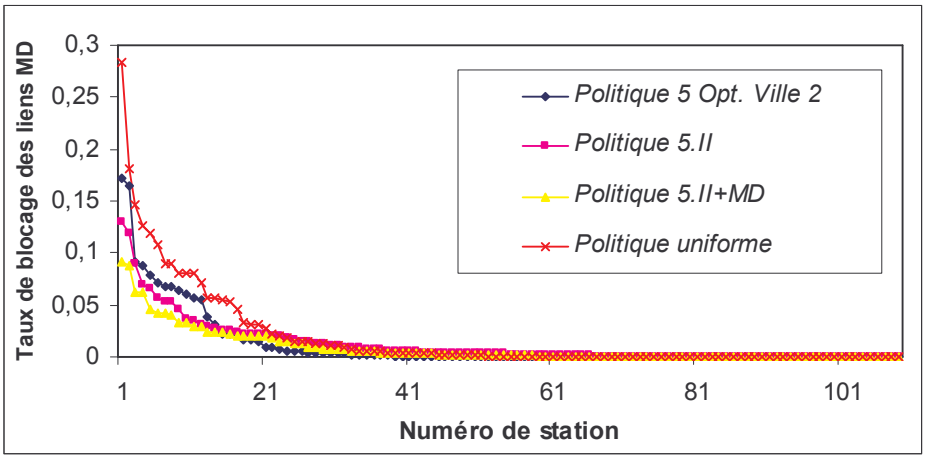


Figure 6.15 : Taux de blocage des liens MD moyen par station (rangé dans l'ordre décroissant) pour les politiques uniforme, *5 Opt. Ville 2*, *5.II* et *5.II+MD* sur la zone de la ville 2.

La politique *5 Opt. Ville 2* a permis d'améliorer les performances en taux de blocage et taux de blocage des liens de MD sur l'ensemble des stations du réseau, par rapport aux performances données par un paramétrage uniforme du réseau. Cette politique permet aussi d'améliorer le taux de blocage pour la plupart des stations en comparaison avec les politiques *5.II* et *5.II+MD*, aux prix de légères dégradations de performances pour le taux de blocage des liens de MD (seulement pour les pires stations du réseau) et pour le taux de coupure (dégradation faible mais quasi uniforme sur l'ensemble des stations du réseau).

6.2.4.2 Variations de l'efficacité des contrôleurs en fonction de la densité de trafic

Nous étudions ci-dessous la réponse de la politique *5 Opt. Ville 2* à une variation de la densité de trafic en entrée (variant entre 2000 et 3000 sur l'ensemble du réseau, le taux moyen d'arrivée des appels étant fixé proportionnellement) et comparons ses performances aux fronts P.U. (figure 6.10) et aux performances des politiques *5.II* et *5.II+MD* (figures 6.11). A partir des figures 6.11, on trace sur la figure 6.16, les résultats obtenus pour la politique *5 Opt. Ville 2* dans le plan taux de coupure – taux de blocage en moyenne sur l'ensemble des stations, pour les 10% pires et les 5% pires stations du réseau, la densité de trafic variant dans l'ensemble {2000, 2500, 3000}, et la charge moyenne de 18 à 27 mobiles voix par secteur.

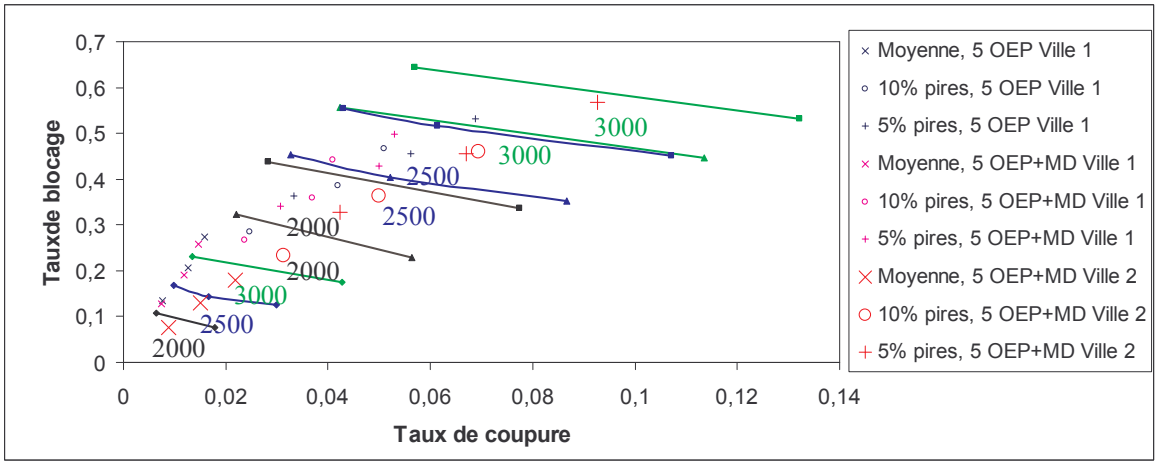


Figure 6.16 : Performances des politiques *5 Opt. Ville 2*, *5.II* et *5.II+MD* (moyenne des stations, des 10% pires et des 5% pires stations) dans le plan taux de coupure / taux de blocage.

Par rapport aux fronts P.U., la politique *5 Opt. Ville 2* de contrôle dynamique donne de meilleures performances en taux de blocage et taux de coupure qu'un paramétrage uniforme

aussi bien en moyenne sur l'ensemble des stations que pour les 10% pires et les 5% pires stations du réseau, la densité de trafic variant dans l'ensemble {2000, 2500, 3000}. Par ailleurs, par rapport aux politiques *5.II* et *5.II+MD*, la politique *5 Opt.* reste plus efficace en termes de taux de blocage et taux de coupure moyennés sur l'ensemble des stations pour une densité de trafic variant dans l'ensemble {2000, 2500, 3000}. En revanche, on constate que plus la densité de trafic augmente et plus les performances de la politique *5 Opt.* se dégradent pour les 10% pires et 5% pires stations du réseau, par rapport aux politiques *5.II* et *5.II+MD*: ce point tend à renforcer notre commentaire du paragraphe 6.2.4.1 émettant l'hypothèse que la politique *5.II+MD* semble être plus adaptée à l'amélioration des performances des pires stations du réseau (dont la situation se dégrade avec l'augmentation de la densité de trafic), alors que la politique *5 Opt. Ville 2* privilégie une amélioration moyenne des performances. Dans notre compréhension, cette différence de comportement entre ces deux politiques provient du fait que la première a été optimisée sur une zone avec un trafic fortement inhomogène, alors que l'optimisation de la seconde a été faite pour un trafic homogène.

6.3. Analyse des premiers résultats d'optimisation des CLF

Les paragraphes précédents nous ont permis d'illustrer (à l'aide de simulations avec notre simulateur de réseau mobile UMTS) les gains en performances (pour les pires stations ou en moyenne sur l'ensemble des stations selon les objectifs de l'opérateur) que l'on peut espérer avec l'ingénierie automatique en mode *on-line*. Le chapitre 7 étudiera et testera de la même façon l'application du mode *off-line* de l'ingénierie automatique. Par ailleurs, la méthode d'optimisation par essaim de particules a permis d'améliorer par une approche systématique les gains en performances et laisse la possibilité à l'opérateur de choisir les objectifs principaux de ses contrôleurs de logique floue (par l'intermédiaire du choix des fonctions coût). En conclusions de cette partie nous indiquons dans les tableaux suivants 6.1 et 6.2 les gains (ou pertes) de performances obtenus à l'aide de l'ingénierie dynamique avec les différentes politiques obtenues par optimisation par essaim de particules (en moyenne sur l'ensemble des stations, ainsi que pour les 10% pires et les 5% pires stations du réseau). Nous commençons par le tableau des gains de performances obtenus sur la ville 1 par rapport à un paramétrage uniforme par défaut (seuil de charge cible DL=0.7 et *Add_win*=3.5 dB) :

Politique optimisée par OEP	% gain moyen en blocage	% gain moyen en blocage pour les 10% pires BTS	% gain moyen en blocage pour les 5% pires BTS	% gain moyen en coupure	% gain moyen en coupure pour les 10% pires BTS	% gain moyen en coupure pour les 5% pires BTS
Cm = 2.5	**- 4 %**	**+ 8 %**	**+ 14 %**	**- 91 %**	**- 80 %**	**-76 %**
Cm = 3.5	**+ 3 %**	**+ 6 %**	**+ 11 %**	**- 214 %**	**- 166 %**	**- 153 %**
Cm = 5.II	**- 15 %**	**+ 3 %**	**+ 13 %**	**+ 13 %**	**+ 17 %**	**+ 18 %**
Cm = 10	**- 35 %**	**+ 5 %**	**+ 13 %**	**+ 62 %**	**+ 62 %**	**+ 60 %**
Cm = 2.5 + MD	**- 3 %**	**+ 4 %**	**+ 9 %**	**-12 %**	**+ 1 %**	**- 4 %**
Cm = 3.5 + MD	**+ 3 %**	**- 4 %**	**+ 6 %**	**- 98 %**	**- 72 %**	**- 77 %**
Cm = 5.II + MD	**- 11 %**	**+ 5 %**	**+ 14 %**	**+ 37 %**	**+ 40 %**	**+ 42 %**

Table 6.1 : Pourcentages des gains en taux de blocage et taux de coupure sur un petit réseau de la zone de la ville 1 par rapport à un paramétrage uniforme (seuil charge cible DL=0.7 et *Add_win*=3.5 dB).

Les gains de performances pour les taux de blocage et taux de coupure de la table 6.1 viennent confirmer les commentaires des figures 6.4 et 6.5 :

- Il est plus efficace d'optimiser séquentiellement par essaim de particules le contrôle dynamique des algorithmes d'admission et de macro-diversité plutôt que d'optimiser seulement le contrôle dynamique de l'algorithme d'admission.

- Typiquement, si on considère le paramétrage uniforme du seuil de charge cible DL à 0.7 comme étant la meilleure situation de départ en terme de compromis taux de coupure – taux de blocage, les politiques optimisées avec un coefficient de mixage de 5 sont les plus adaptées à cette zone de fonctionnement.

Ci-dessous, nous fournissons le tableau équivalent des gains de performances obtenus sur la zone de la ville 2 par rapport à un paramétrage uniforme par défaut (seuil de charge cible DL=0.7 et *Add_win*=3.50 dB) :

Politique optimisée par OEP	% gain moyen en blocage	% gain moyen en blocage pour les 10% pires BTS	% gain moyen en blocage pour les 5% pires BTS	% gain moyen en coupure	% gain moyen en coupure pour les 10% pires BTS	% gain moyen en coupure pour les 5% pires BTS
Cm = 2.5	**- 3 %**	**+ 24 %**	**+ 25 %**	**- 116 %**	**-110 %**	**- 110 %**
Cm = 3.5	**+ 7 %**	**+ 30 %**	**+ 31 %**	**- 186 %**	**- 182 %**	**- 193 %**
Cm = 5.II	**- 22 %**	**+ 15 %**	**+ 18 %**	**- 30 %**	**- 28 %**	**- 30 %**
Cm = 2.5 + MD	**+ 6 %**	**+ 32 %**	**+ 33 %**	**-89 %**	**- 77 %**	**- 78 %**
Cm = 3.5 + MD	**+ 13 %**	**+ 36 %**	**+ 37 %**	**-145 %**	**- 138 %**	**-150 %**
Cm = 5.I + MD	**- 19 %**	**+ 18 %**	**+ 21 %**	**- 4 %**	**+ 4 %**	**+ 1 %**
Cm = 5.II + MD	**- 14 %**	**+ 20 %**	**+ 22 %**	**- 17 %**	**- 9 %**	**- 13 %**
Cm = 5 Opt. Ville 2	**+ 23 %**	**+ 20 %**	**+ 18 %**	**- 54 %**	**- 53 %**	**- 56 %**

Table 6.2 : Pourcentages des gains en taux de blocage et taux de coupure sur un réseau de la zone de la ville 2 par rapport à un paramétrage uniforme (seuil de charge cible DL=0.7 et *Add_win*=3.5 dB).

Aux vues des résultats de la table 6.2 et des figures 6.7, 6.8 et 6.9, il semble plus pertinent, sur la zone de la ville 2, de comparer les performances obtenues pour les différentes politiques optimisées à celles d'un paramétrage uniforme avec un seuil charge cible DL=0.8 (ce paramétrage fournit une meilleure situation de départ en terme de compromis taux de coupure – taux de blocage) :

Politique optimisée par OEP	% gain moyen en blocage	% gain moyen en blocage pour les 10% pires BTS	% gain moyen en blocage pour les 5% pires BTS	% gain moyen en coupure	% gain moyen en coupure pour les 10% pires BTS	% gain moyen en coupure pour les 5% pires BTS
Cm = 5.II	**- 46 %**	**+ 3 %**	**+ 8 %**	**+ 21 %**	**+ 16 %**	**+ 4 %**
Cm = 5.I + MD	**- 43 %**	**+ 6 %**	**+ 12 %**	**+ 37 %**	**+ 37 %**	**+ 28 %**
Cm = 5.II + MD	**- 36 %**	**+ 9 %**	**+ 13 %**	**+ 29 %**	**+ 29 %**	**+ 17 %**
Cm = 5 Opt. Ville 2	**+ 8 %**	**+ 9 %**	**+ 9 %**	**+ 6 %**	**+ 0 %**	**- 14 %**

Table 6.3 : Pourcentages des gains en taux de blocage et taux de coupure sur un réseau de la zone de la ville 2 par rapport à un paramétrage uniforme (seuil de charge cible DL=0.8 et *Add_win*=3.5 dB).

De nouveau, sur la table 6.3, on constate qu'il est plus efficace de contrôler dynamiquement à la fois le paramétrage des algorithmes de contrôle d'admission et de macro-diversité plutôt que le contrôle d'admission seul. Dans le cas du réseau de la ville 2, on considère qu'un paramétrage uniforme du seuil de charge cible DL proche de 0.8 (cf. figures 6.7, 6.8 et 6.9) fournit une meilleure situation de départ en terme de compromis coupure – blocage. Et il ressort de notre étude que les politiques optimisées avec un coefficient de mixage de 5 sont les plus adaptées à cette zone de fonctionnement. Enfin, il apparaît sur la table 6.3, que la politique *5 Opt. Ville 2* améliore les performances des stations du réseau de manière homogène, alors que les politiques *5.I+MD* et *5.II+MD* favorisent plutôt l'amélioration des performances des pires stations du réseau.

6.4. Méthode d'OEP en mode on-line

6.4.1. Rappel de la méthode

Dans ce paragraphe, nous proposons de tester le schéma d'optimisation des CLF en mode *on-line* par essaim de particules, introduit au paragraphe 4.1.4. On commence donc par tracer les fronts de performances obtenus dans le plan taux de coupure – taux de blocage avec différents paramétrages uniformes des algorithmes RRM. Sur ce graphique, on détermine le (ou les) paramétrage(s) initial(aux) uniformes le(s) plus efficace(s), c'est-à-dire donnant le meilleur compromis entre le taux de coupure, de blocage et de blocage des liens en MD. A partir de là, on teste différents coefficients de mixage en mono-objectif (qui dépendront de l'environnement étudié et du paramétrage initial uniforme choisi) pour l'optimisation par essaim de particules. Un ou plusieurs couples (paramétrage initial / coefficient de mixage) sont sélectionnés pour le mode *on-line* de l'algorithme d'OEP en mono-objectif. On pourra aussi choisir de lancer une optimisation en multi-objectifs (voir paragraphe 6.4.3). Enfin, les performances obtenues après optimisation par essaim de particules sont comparées avec les performances initiales données par les fronts de paramétrage uniformes, en moyenne sur l'ensemble des stations ou en considérant les pires stations du réseau.

6.4.2. Mises en applications sur le réseau de la ville 1

La méthode d'optimisation des contrôleurs de logique floue proposée est mise en applications sur la zone de la ville 1, afin d'éviter des délais de simulations trop importants.

Détermination des paramétrages initiaux donnant les meilleurs contrôleurs

Différents paramétrages initiaux du seuil de charge cible DL dans {0.6, 0.65, 0.7, 0.75} ont été testés ainsi que différents coefficients de mixage dans {2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6}. Les CLF ont été optimisés (simultanément* pour le contrôle d'admission et la macro-diversité) par OEP, avec la fonction coût du 6.1.2.2, en mono-objectif pour différents couples de paramétrage initial et coefficient de mixage. La figure 6.17 compare l'efficacité des meilleures politiques obtenues par rapport aux fronts P.U (pour une charge moyenne de 42 mobiles voix par secteur) :

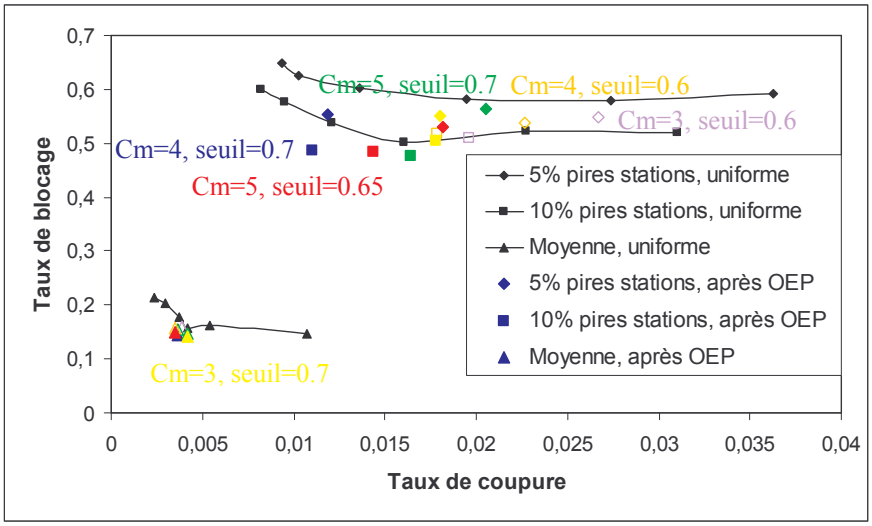


Figure 6.17 : Efficacité des politiques optimisées pour différents couples (paramétrage initial / coefficient de mixage) par rapport aux fronts P.U. dans le plan taux de coupure / taux de blocage.

* Les CLF sont optimisés par une seule utilisation de l'algorithme d'OEP simultanément pour le C.A. et la M.D.

Toutes les politiques de la figure 6.17 ont des performances moyennes en termes de taux de coupure et taux de blocage légèrement meilleures que le front de paramétrage uniforme. Pour les 5% et 10% pires stations, les deux politiques les plus clairement efficaces sont celles obtenues à partir des couples (*Cm=4, seuil init. DL=0.7*) et (*Cm=5, seuil init. DL=0.65*). Nous sélectionnons donc ces deux politiques *Select_4_0.7* et *Select_5_0.65* dont nous allons analyser ci-dessous les performances de contrôle dynamique.

Tests des différentes fonctions coûts

Sur la figure 6.18, il s'agit de comparer les différentes fonctions coûts introduites (équations 4.4 et 4.5) au paragraphe 4.1.3.2 pour l'optimisation des CLF par OEP (simultanément pour le contrôle d'admission et la macro-diversité), pour un même paramétrage initial de l'algorithme (*Cm=4, seuil init. DL=0.7*), et pour une charge moyenne de 42 mobiles voix par secteur :

- la fonction *Coût (10,3,1)* privilégiant l'optimisation des performances des pires stations: $w_i = u_i = 10$ pour les 10% pires stations, $w_i = u_i = 3$ pour les 10% stations suivantes et $w_i = u_i = 1$ pour le autres stations.
- la fonction *Coût (1,1,1)* privilégiant l'optimisation des performances moyennes du réseau: $w_i = u_i = 1$ pour toutes les stations.
- La fonction *Coût** qui module les coefficient de pondération par station conformément aux équations 4.7 et 4.8 en fonction du nombre de mobiles servis par la station *i*.
- La fonction *Coût*** qui module les coefficient de pondération par station conformément aux équations 4.7 et 4.9 en fonction du nombre de mobiles servis par la station *i* et de la charge de la station.

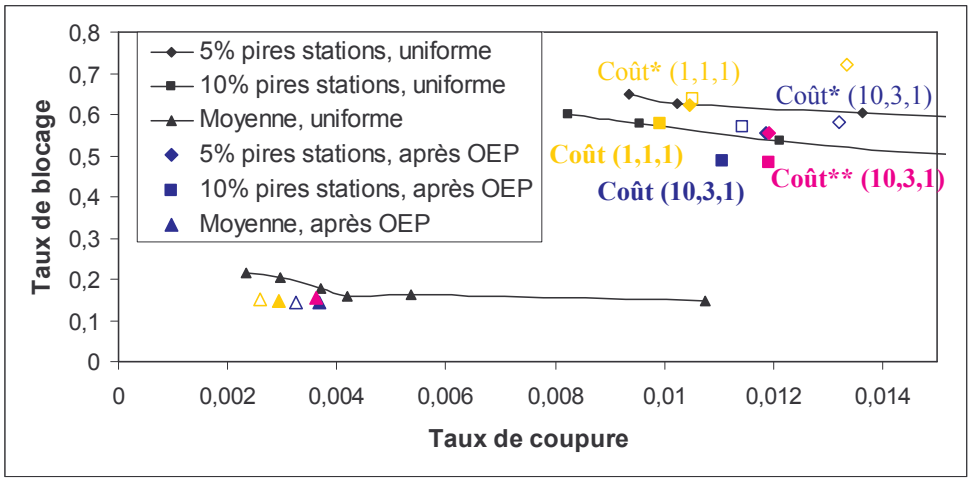


Figure 6.18 : Efficacité des politiques optimisées pour différentes fonctions coûts à partir du couple initial (*Cm=4, seuil init. DL=0.7*), par rapport aux fronts de paramétrage uniformes dans le plan taux de coupure / taux de blocage.

Le meilleur compromis entre l'optimisation des performances moyennes et celle des pires stations du réseau a été obtenu pour la fonction *Coût (10,3,1)*, qui permet aussi un gain notable des performances par rapport aux fronts de paramétrage uniforme. La fonction *Coût (1,1,1)*, même si elle apparaît moins efficace sur la figure 6.18, pourra être aussi utilisée en fonction des objectifs d'amélioration de l'opérateur (Amélioration globale / amélioration locale).

Analyse des résultats des politiques sélectionnées

A partir de simulations de réseau mobile UMTS avec contrôle dynamique des paramètres RRM dont les hypothèses sont résumées dans la table suivante, nous allons analyser l'impact de l'ingénierie dynamique sur les causes de coupure et blocage des appels et comparer avec la situation initiale où le paramétrage était uniforme. La table 6.4 suivante résume les hypothèses de ces simulations d'ingénierie dynamique :

Paramètres	Valeurs	Commentaires
Politiques testées	*Select_4_0.7* *Select_5_0.65*	Seuil de charge cible UL statique égal à 0.7 ou 0.715 Seuil de charge cible UL statique égal à 0.7
Nombre de simulations	10	Résultats moyennés sur 10 simulations
Nombre d'itérations	3000 itérations	Avec un intervalle de temps de 4 s
Paramètres contrôlés dynamiquement	Seuil de charge cible DL Taille de la fenêtre d'ajout *Add_win*	Contrôle d'admission Macro-diversité
Paramétrage de référence	Seuil de charge cible DL =0.7 *Add_win* = 3.5 dB Seuil de charge cible UL =0.7	Référence pour estimer les gains de performance
Nombre d'objets de trafic	3	-
Nombre total initial de mobiles	900 mobiles	600 sur l'objet de trafic 1, et 150 sur les objets de trafic 2 et 3
Service	Voix	Durée moyenne des communications = 100s
Taux d'arrivée des mobiles par objet de trafic	6 s^{-1}, 1.5 s^{-1}, 1.5 s^{-1}	Respectivement pour les objets de trafic 1, 2 et 3
Charge moyenne par secteur	42 mobiles voix par secteur	En faisant l'hypothèse d'une distribution de trafic homogène

Table 6.4 : Hypothèses de simulations d'ingénierie dynamique avec notre simulateur de réseau UMTS.

Sur la figure 6.19 suivante, on représente les performances (sur la zone de la ville 1) moyennées sur 10 simulations de la politique *Select_4_0.7* pour une valeur de seuil de charge cible UL statique (noté *seuil UL* sur la figure suivante) égale à 0.7 ou 0.715, et de la politique *Select_5_0.65* pour une valeur de seuil de charge cible UL statique égale à 0.7.

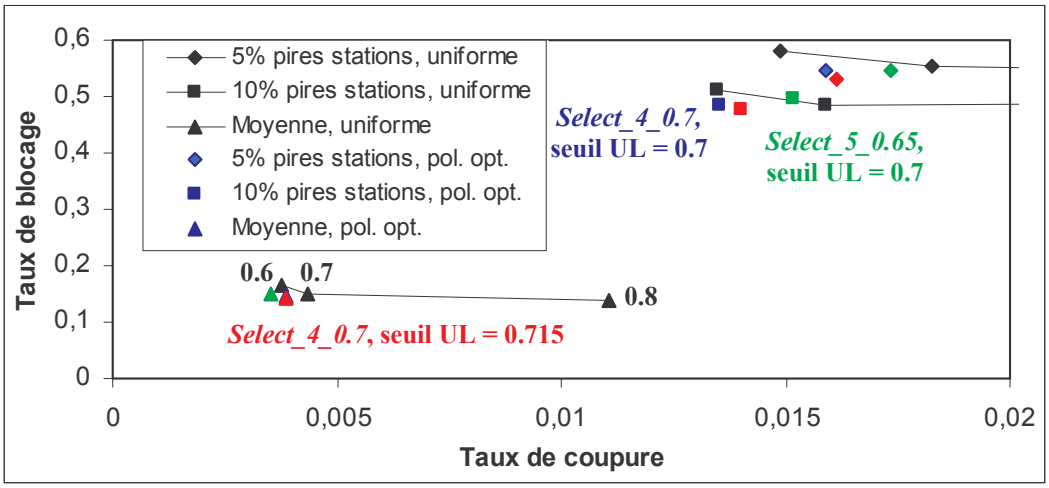


Figure 6.19 : Efficacité des politiques *Select_4_0.7* (bleu, rouge) et *Select_5_0.65* (vert) pour un niveau de trafic de 900 mobiles, par rapport aux fronts P.U. dans le plan taux de coupure / taux de blocage.

Ci-dessous, nous fournissons, dans la table 6.5, les gains de performances obtenus avec les politiques *Select_4_0.7* et *Select_5_0.65* sur la zone de la ville 1 par rapport à un paramétrage uniforme par défaut (seuil de charge cible DL=0.7 et *Add_win*=3.50 dB) :

Politique optimisée par OEP	% gains en coupure			% gains en blocage			% gains en blocage des liens de MD		
	5% pires	10% pires	Moyenne	5% pires	10% pires	Moyenne	5% pires	10% pires	Moyenne
Politique *Select_4_0.7* avec un seuil de charge cible UL = 0.7	**+13%**	**+15%**	**+11%**	**+1%**	**+0.3%**	**+2%**	**+36%**	**+29%**	**+21%**
Politique *Select_4_0.7* avec un seuil de charge cible UL = 0.715	**+12%**	**+12%**	**+11%**	**+4%**	**+1%**	**+4%**	**+37%**	**+31%**	**+20%**
Politique *Select_5_0.65* avec un seuil de charge cible UL = 0.7	**+5%**	**+5%**	**+19%**	**+2%**	**-2.5%**	**-0.5%**	**+57%**	**+57%**	**+49%**

Table 6.5 : Pourcentages des gains en taux de blocage, de coupure et de blocage des liens de MD pour les politiques *Select_4_0.7* et *Select_5_0.65* par rapport à un paramétrage uniforme (seuil de charge cible DL=0.7 et *Add_win*=3.5 dB).

La politique *Select_4_0.7* améliore la situation par rapport aux fronts P.U, notamment pour les taux de coupure et taux de blocage des liens en MD. En comparant l'efficacité de cette politique pour des seuils de charge cible UL de 0.7 et 0.715, on s'aperçoit que la deuxième valeur permet une amélioration légère des gains pour le taux de blocage. La politique *Select_5_0.65* améliore la situation par rapport aux fronts P.U pour le taux de coupure et surtout pour le taux de blocage des liens en MD. La table 6.6 suivante permet de mettre en évidence les causes de blocage et coupure des appels pour chaque politique de paramétrage :

Politiques	Paramétrage uniforme statique	Politique *Select_4_0.7*		Politique *Select_5_0.65*
Seuil de charge cible UL statique	0.7	0.7	0.715	0.7
Nombre de mobiles tirés	108529	108888	108679	108955
Nombre total de mobiles rejetés	**24127** Référence	**23817** **Gain = 1.3%**	**23583** **Gain = 2.3%**	**24432** **Perte = -1.3%**
Nombre de mobiles bloqués à l'admission: CHARGE_DL	18894	17229	18096	17293
Nombre de mobiles bloqués à l'Admission: CHARGE_UL	4883	6268	5168	6853
Nombre de mobiles bloqués à l'admission: COUV_DL	25	20	17	9
Nombre de mobiles bloqués à l'admission: COUV_UL	11	11	10	10
Nombre de mobiles bloqués à l'admission: PILOTE	1	3	2	1
Total de mobiles bloqués	**23814**	**23530**	**23293**	**24166**
Nombre de mobiles coupés: COUV_DL	212	186	183	129
Nombre de mobiles coupés: COUV_UL	65	70	67	64
Nombre de mobiles coupés: CHARGE_DL	0.7	1	0,4	0
Nombre de mobiles coupés: CHARGE_UL	35	30	40	72
Total de mobiles coupés	**313**	**286**	**290**	**265**

Table 6.6 : Analyse des causes de rejet des mobiles pour les politiques *Select_4_0.7* et *Select_5_0.65* par rapport au cas d'un paramétrage uniforme (seuil charge cible DL=0.7 et *Add_win*=3.5 dB).

Sur la table 6.6, la gestion dynamique du paramètre seuil de charge cible DL permet à la politique *Select_4_0.7* de diminuer le nombre de mobiles bloqués à l'admission à cause du

niveau de charge DL. Cependant cette baisse du nombre de mobiles bloqués sur le lien DL s'est accompagnée d'une hausse du nombre de mobiles bloqués à l'admission à cause du niveau de charge UL. Pour réduire cet effet, on a utilisé la même politique *Select_4_0.7* mais avec un nouveau seuil de charge cible UL statique de 0.715. On obtient ainsi un compromis intéressant qui a permis de réduire le nombre total de mobiles bloqués et de mobiles coupés, de 2.3% par rapport au cas du paramétrage statique et uniforme. La politique *Select_5_0.65* fournit un compromis moins intéressant pour ce niveau de trafic: les diminutions du nombre de mobiles bloqués à cause de la charge DL et du nombre de mobiles coupés s'est accompagnée d'une nette hausse du nombre de mobiles bloqués à cause de la charge UL. Globalement le pourcentage de mobiles rejetés (bloqués ou coupés) par rapport à la politique uniforme de référence s'est dégradée de 1.3%.

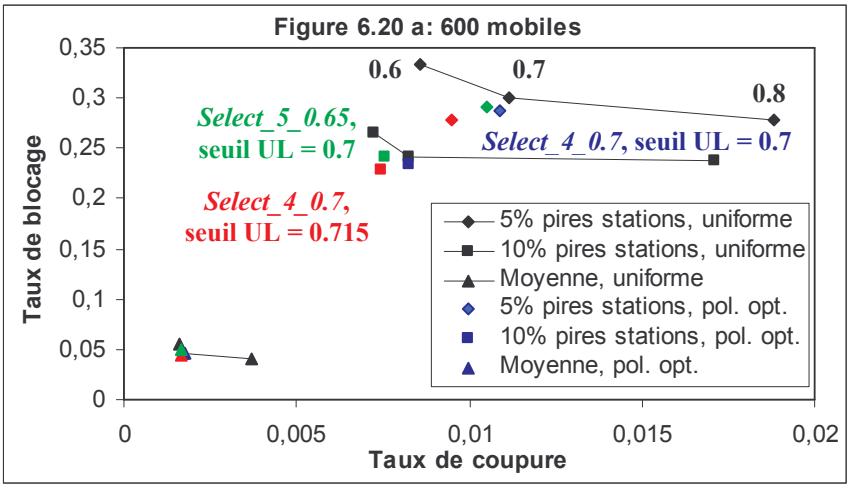
Efficacité des politiques optimisées par OEP pour différents niveaux de trafic en entrées

On souhaite à présent étudier les variations de l'efficacité des politiques *Select_4_0.7* et *Select_5_0.65* par rapport à un paramétrage statique et uniforme, pour différents niveaux de trafic. La table 6.7 suivante résume les hypothèses des simulations d'ingénierie dynamique pour différentes densités de trafic :

Paramètres	Valeurs	Commentaires
Politiques testées	*Select_4_0.7* *Select_5_0.65*	Seuil de charge cible UL statique égal à 0.7 ou 0.715 Seuil de charge cible UL statique égal à 0.7
Nombre de simulations	5	Résultats moyennés sur 5 simulations
Nombre d'itérations	3000 itérations	Avec un intervalle de temps de 4 s
Paramètres contrôlés	Seuil de charge cible DL Taille de la fenêtre d'ajout *Add_win*	Contrôle d'admission Macro-diversité
Paramétrage de référence	Seuil de charge cible DL&UL=0.7 *Add_win* = 3.5 dB	Référence pour estimer les gains de performance
Nombre d'objets de trafic	3	-
Nombre total initial de mobiles	600 / 1200 mobiles	400 / 800 sur l'objet de trafic 1, et 100 / 200 sur les objets de trafic 2 et 3.
Service	Voix	Durée moyenne des appels = 100s
Taux d'arrivée des mobiles par objet de trafic	4 / 8 s^{-1}, 1 / 2 s^{-1}, 1 / 2 s^{-1}	Respectivement pour les objets de trafic 1, 2 et 3
Charge moyenne par secteur	28 / 57 mobiles voix par secteur	En faisant l'hypothèse d'une distribution de trafic homogène

Table 6.7 : Hypothèses de simulations d'ingénierie dynamique avec notre simulateur de réseau UMTS pour différents niveaux de trafic.

Sur les figures 6.20 suivantes, on représente les performances (sur la zone de la ville 1) moyennées sur 5 simulations des politiques *Select_4_0.7* et *Select_5_0.65* pour différentes densités de trafic.



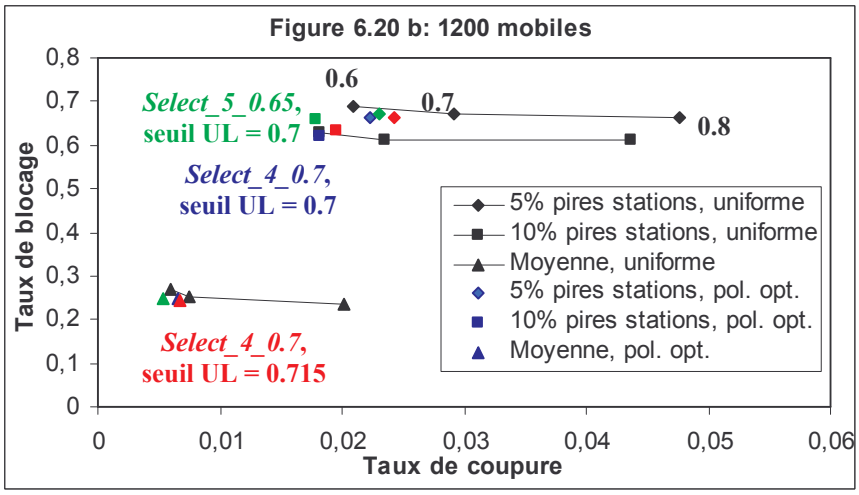


Figure 6.20 : Efficacité des politiques *Select_4_0.7* et *Select_5_0.65* pour des niveaux de trafic de 600 et 1200 mobiles (fig. a et b) par rapport aux fronts P.U. dans le plan taux de coupure / taux de blocage.

Ci-dessous nous fournissons, dans la table 6.8, les gains de performances obtenus avec les politiques *Select_4_0.7* et *Select_5_0.65* sur la zone de la ville 1 par rapport à un paramétrage uniforme par défaut (seuil de charge cible DL=0.7 et *Add_win*=3.50 dB) et pour des niveaux de trafic de 600 et 1200 mobiles.

Politique optimisée par OEP	% gains en coupure			% gains en blocage			% gains en blocage des liens de MD		
	5% pires	10% pires	Moyenne	5% pires	10% pires	Moyenne	5% pires	10% pires	Moyenne
Politique *Select_4_0.7* avec un seuil de charge cible UL = 0.7	+2%	-0,2%	-4%	+4%	+3%	-0.1%	+15%	+5%	-0,7%
	+24%	+23%	+12%	+1%	-1%	+3%	+38%	+38%	+35%
Politique *Select_4_0.7* avec un seuil de charge cible UL = 0.715	+15%	+10%	+4%	+8%	+5%	+4%	+7%	+4%	+2%
	+17%	+17%	+10%	+1%	-3%	+4%	+40%	+37%	+33%
Politique *Select_5_0.65* avec un seuil de charge cible UL = 0.7	+6%	+8%	+5%	+3%	0%	-6%	+30%	+26%	+21%
	+21%	+24%	+29%	0%	-7%	+2%	+56%	+61%	+58%

Table 6.8 : Pourcentages des gains en taux de blocage, de coupure et de blocage des liens de MD pour les politiques *Select_4_0.7* et *Select_5_0.65* pour des niveaux de trafic de 600 mobiles (1ère ligne) et 1200 mobiles (2nde ligne) par rapport à un paramétrage uniforme (seuil de charge cible DL=0.7 et *Add_win*=3.5 dB).

Globalement, la politique *Select_4_0.7* améliore les performances par rapport aux fronts P.U, pour les taux de coupure et taux de blocage des liens en MD. Cette politique avec un seuil de charge cible UL égal à 0.715 est plus adaptée aux situations de faible charge, permettant dans ce cas d'améliorer aussi le taux de blocage. En revanche, cette politique avec un seuil de charge cible UL égal à 0.7 est adaptée aux situations de forte charge, fournissant des gains importants sur les taux de coupure et taux de blocage des liens en MD. La politique *Select_5_0.65* améliore nettement les performances par rapport aux fronts P.U pour le taux de blocage des liens en MD, pour les faibles et fortes charges de trafic. Cette politique améliore aussi nettement le taux de coupure dans les cas de forte densité de mobiles. En revanche, elle ne fournit en général aucun gain sur le taux de blocage.

Les figures suivantes permettent de mettre en évidence respectivement les gains (en pourcentage) en nombre de mobiles rejetés (bloqués ou coupés) par rapport à une politique de paramétrage uniforme, les causes de blocage des appels ainsi que le nombre de mobiles coupés pour chaque politique de paramétrage dynamique et pour différents niveaux de trafic.

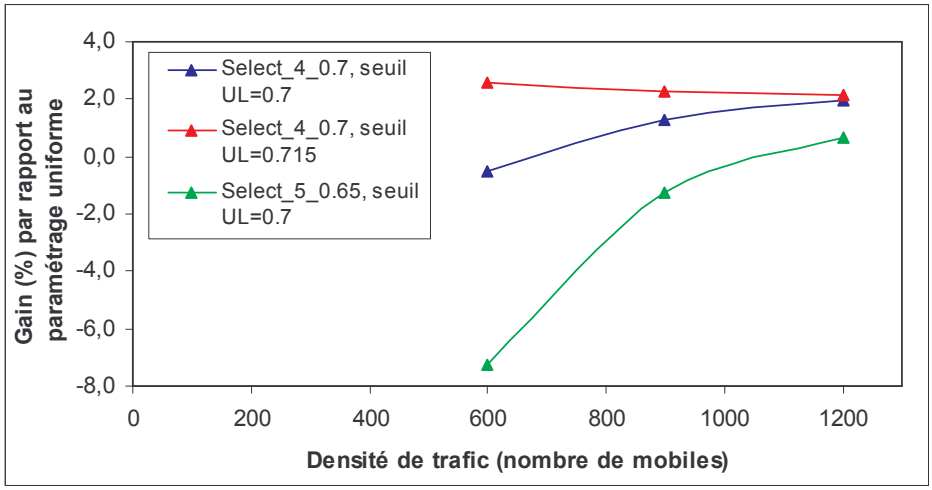


Figure 6.21 : Gains (en %) en nombre de mobiles rejetés avec les politiques *Select_4_0.7* et *Select_5_0.65* pour des niveaux de trafic de 600 / 900 / 1200 mobiles par rapport à un paramétrage uniforme (seuil de charge cible DL=0.7 et *Add_win*=3.5 dB).

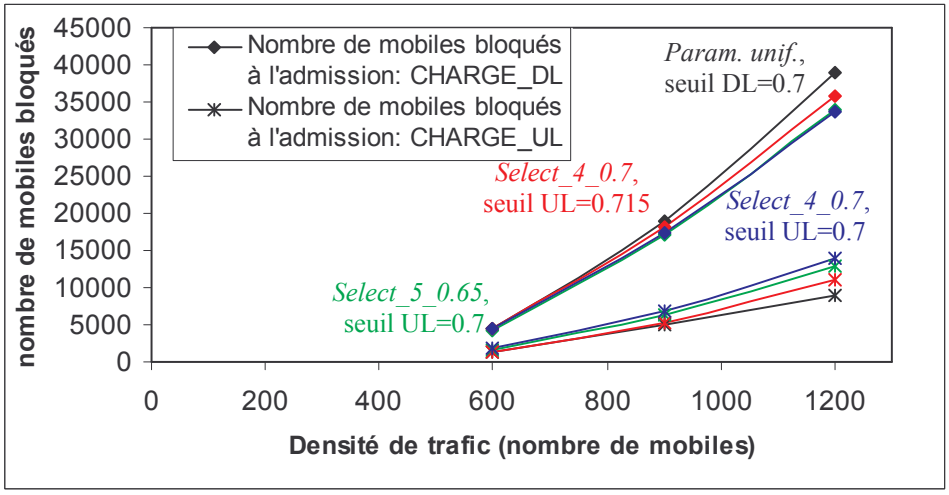


Figure 6.22 : Nombre de mobiles bloqués (à cause de la charge DL ou UL) avec les politiques *Select_4_0.7* et *Select_5_0.65* pour des niveaux de trafic de 600 / 900 / 1200 mobiles par rapport à un paramétrage uniforme (seuil de charge cible DL=0.7 et *Add_win*=3.5 dB).

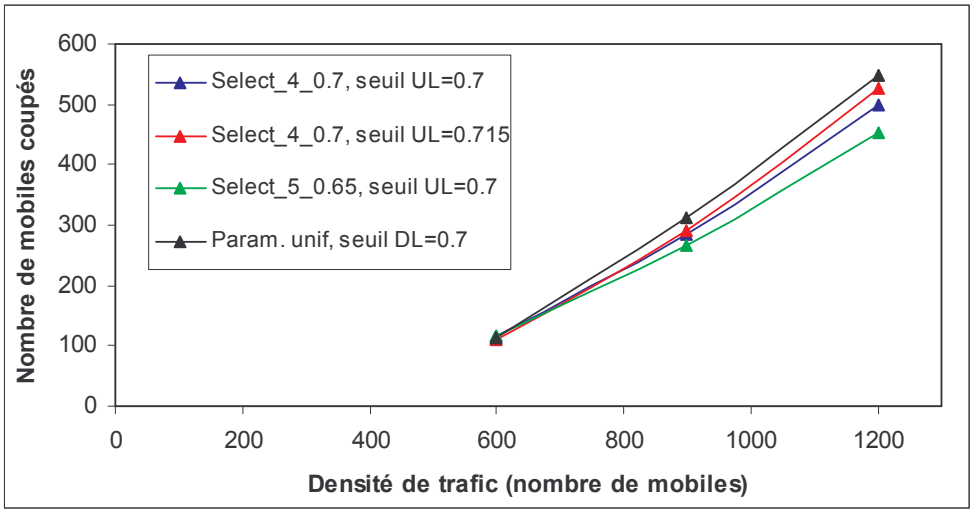


Figure 6.23 : Nombre de mobiles coupés avec les politiques *Select_4_0.7* et *Select_5_0.65* pour des niveaux de trafic de 600 / 900 / 1200 mobiles par rapport à un paramétrage uniforme (seuil de charge cible DL=0.7 et *Add_win*=3.5 dB).

La figure 6.21 confirme que la politique *Select_4_0.7* avec un seuil de charge cible UL égal à 0.715 fournit de meilleurs gains (en nombre de mobiles rejetés) dans les situations de faible charge. Inversement cette politique avec un seuil de charge cible UL égal à 0.7 est plus efficace dans les situations de forte charge. Enfin, la politique *Select_5_0.65* ne permet pas d'améliorer le nombre de mobiles rejetés par rapport à la politique de paramétrage uniforme de référence, même si son efficacité augmente avec la densité de trafic. La figure 6.22 indique que la politique initiale de paramétrage uniforme (seuil de charge cible DL=0.7 et *Add_win*=3.5 dB) donne un fort blocage à cause de la charge DL et relativement un faible blocage à cause de la charge DL. Nos politiques de gestion dynamique permettent de diminuer le blocage à cause de la charge DL en adaptant le seuil de charge cible DL, mais dégradent le blocage à cause de la charge UL. De ce point de vue, c'est la politique *Select_5_0.65* qui réalise le meilleur compromis entre ces deux causes de blocage. Enfin, la figure 6.23 montre que les trois politiques étudiées améliorent le taux de coupure dans le réseau par rapport à la politique de paramétrage uniforme de référence, la politique *Select_5_0.65* étant la plus efficace de ce point de vue.

6.4.3. Comparaison des résultats de l'OEP en mono et multi-objectifs

Dans ce paragraphe, deux solutions du front optimal de l'algorithme d'optimisation par essaim de particules en multi-objectifs pour le mode *on-line* vont être comparées à la politique *Select_4_0.7* optimisée en mono-objectif pour un coefficient de mixage égal à 4. Nous avons lancé l'optimisation en multi-objectifs pour la même fonction *Coût (10,3,1)* qu'au 6.1.2.2, privilégiant l'optimisation des performances des pires stations, sur le réseau de la ville 1 (fig. 6.1) et pour un densité initiale de 900 mobiles (mêmes hypothèses que pour l'optimisation mono-objectif au paragraphe 6.4.2.1). La figure 6.24 suivante décrit l'ensemble des solutions explorées au cours de l'optimisation en multi-objectifs dans le plan fonction coût de blocage – fonction coût de coupure (pour une charge moyenne de 42 mobiles voix par secteur).

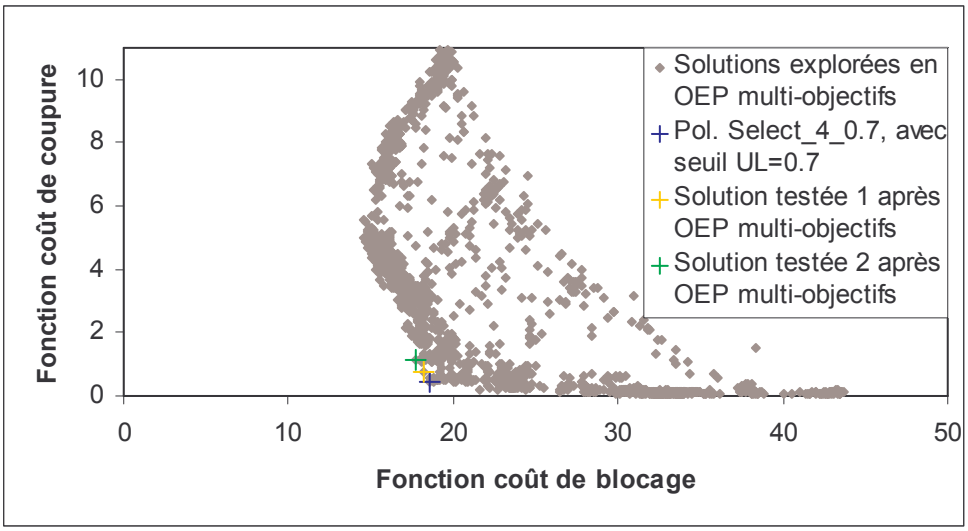


Figure 6.24 : Solutions explorées en OPE multi-objectifs (dont les solutions 1 et 2 du front optimal) comparées avec la politique *Select_4_0.7* obtenue en mono-objectif pour un coefficient de mixage égal à 4, dans le plan fonction coût de blocage – fonction coût de coupure.

Comme dans le cas de l'optimisation par algorithmes génétiques en multi-objectifs (cf. [Pic05]), un front de solutions optimales (ou de rang 1) est obtenu par l'OEP multi-objectifs. La solution *Select_4_0.7* optimisée en mono-objectif appartient d'ailleurs à ce front de rang 1. Nous sélectionnons deux autres solutions optimales, *Select_multi_1* et *Select_multi_2*, appartenant à ce front (meilleures en terme de blocage mais moins efficaces en coupure) dont nous allons comparer les performances effectives par rapport à la politique *Select_4_0.7*.

La figure 6.25 compare les performances des politiques *Select_4_0.7*, *Select_multi_1* et *Select_multi_2* par rapport aux fronts P.U. dans le plan taux de coupure - taux de blocage, moyennées sur dix simulations de réseau avec ingénierie dynamique.

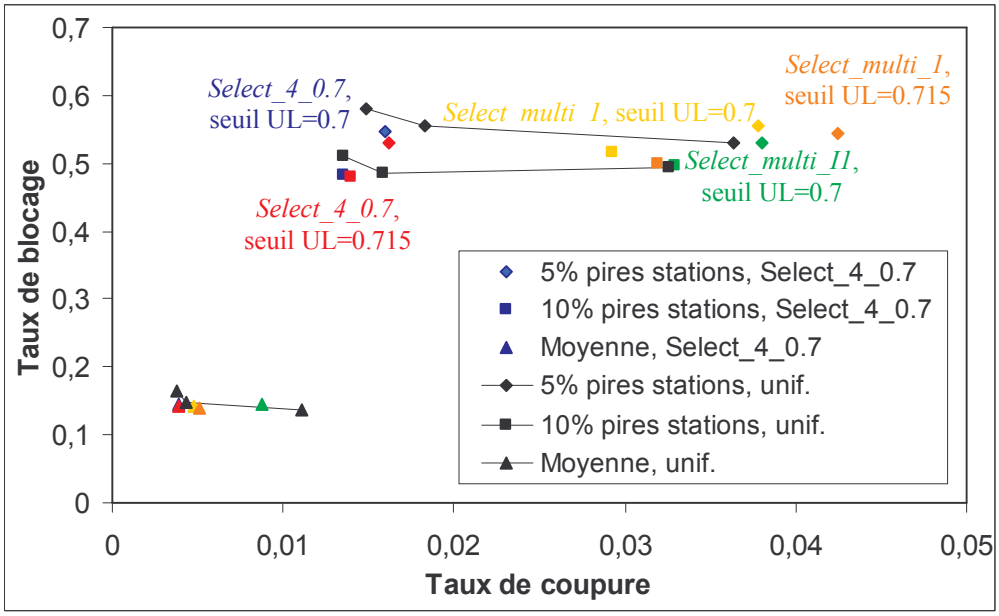


Figure 6.25 : Efficacité des politiques *Select_4_0.7*, *Select_multi_1* et *Select_multi_2* par rapport aux fronts P.U. dans le plan taux de coupure - taux de blocage.

Les politiques *Select_multi_1* et *Select_multi_2* dégradent les performances en taux de coupure, par rapport à la politique *Select_4_0.7* (comme indiqué a priori par la fonction coût

de coupure de la figure 6.24), mais sans amélioration notable des performances en taux de blocage. Du point de vue de ces indicateurs, les politiques *Select_multi_1* et *Select_multi_2* fournissent un compromis entre taux de blocage et taux de coupure moins efficace que la politique *Select_4_0.7*. La table 6.9 suivante permet de mettre en évidence les causes de blocage et coupure des appels pour chaque politique de paramétrage, moyennées sur dix simulations de réseau avec ingénierie dynamique :

Politiques	Paramétrage uniforme statique	Politique *Select_multi_1*		Politique *Select_multi_II*
Seuil de charge cible UL statique	0.7	0.7	0.715	0.7
Nombre de mobiles tirés	108529	108864	109031	109002
Nombre total de mobiles rejetés	**24127** Référence	**23280** **Gain = 3.5%**	**23093** **Gain = 4.3%**	**23853** **Gain = 1.1%**
Nombre de mobiles bloqués à l'admission: CHARGE_DL	18894	13831	14580	15590
Nombre de mobiles bloqués à l'Admission: CHARGE_UL	4883	8982	8002	7440
Nombre de mobiles bloqués à l'admission: COUV_DL	25	9	10	19
Nombre de mobiles bloqués à l'admission: COUV_UL	11	11	11	12
Nombre de mobiles bloqués à l'admission: PILOTE	1	1.5	1	2
Total de mobiles bloqués	**23814**	**22835**	**22604**	**23063**
Nombre de mobiles coupés: COUV_DL	212	112	110	173
Nombre de mobiles coupés: COUV_UL	65	73	80	65
Nombre de mobiles coupés: CHARGE_DL	0.7	0	0.1	476
Nombre de mobiles coupés: CHARGE_UL	35	261	298	76
Total de mobiles coupés	**313**	**445**	**488**	**790**

Table 6.9 : Analyse des causes de rejet des mobiles pour les politiques *Select_multi_I* et *Select_multi_II* par rapport au cas d'un paramétrage uniforme (seuil charge cible DL=0.7 et *Add_win*=3.5 dB).

Sur la table 6.9, la gestion dynamique du paramètre seuil de charge cible DL permet à la politique *Select_multi_I* de nettement diminuer le nombre de mobiles bloqués à l'admission à cause du niveau de charge DL (à comparer avec la table 6.6). Cependant cette baisse du nombre de mobiles bloqués sur le lien DL s'est accompagnée d'une hausse importante du nombre de mobiles bloqués à l'admission à cause du niveau de charge UL. Pour réduire cet effet, on a utilisé la même politique *Select_multi_I* mais avec un nouveau seuil de charge cible UL statique de 0.715. On obtient ainsi un compromis intéressant qui a permis de réduire le nombre total de mobiles bloqués et de mobiles coupés, de 4.3% par rapport au cas du paramétrage statique et uniforme (contre 2.3% pour la politique *Select_4_0.7* avec un seuil de charge cible UL de 0.715). La politique *Select_multi_II* fournit un compromis moins intéressant pour ce niveau de trafic: la diminution du nombre de mobiles bloqués à cause de la charge DL s'est accompagnée d'une nette hausse du nombre de mobiles bloqués à cause de la charge UL et surtout du nombre de mobiles coupés. Globalement le pourcentage de mobiles rejetés (bloqués ou coupés) par rapport à la politique uniforme de référence s'est améliorée de 1.1%.

Performances sur l'ensemble des stations :

Sur les figures 6.26, 6.27 et 6.28, on a classé par ordre décroissant les performances moyennes par station respectivement en termes de taux de blocage, taux de coupure et taux de blocage des liens de MD pour les politiques *Select_4_0.7*, *Select_multi_I* et *Select_multi_II*, moyennées sur dix simulations de réseau avec ingénierie dynamique :

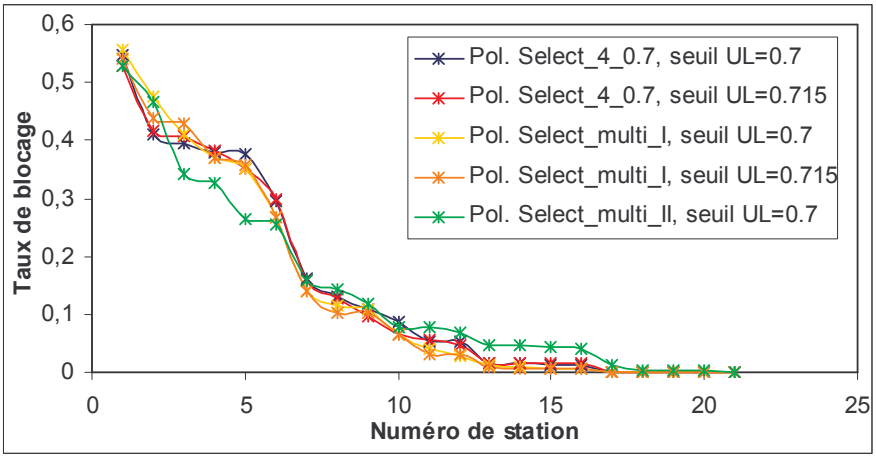


Figure 6.26 : Taux de blocage moyen par station (rangé dans l'ordre décroissant) pour les politiques *Select_4_0.7*, *Select_multi_I* et *Select_multi_II* sur la zone de ville 1.

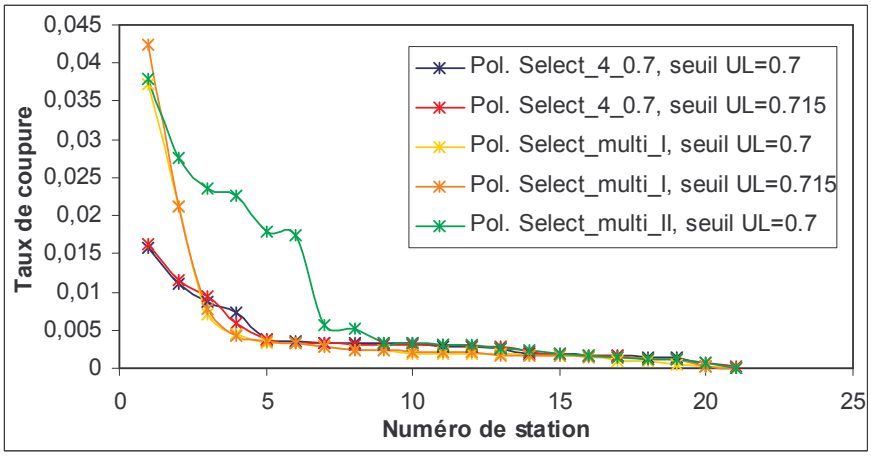


Figure 6.27 : Taux de coupure moyen par station (rangé dans l'ordre décroissant) pour les politiques *Select_4_0.7*, *Select_multi_I* et *Select_multi_II* sur la zone de la ville 1.

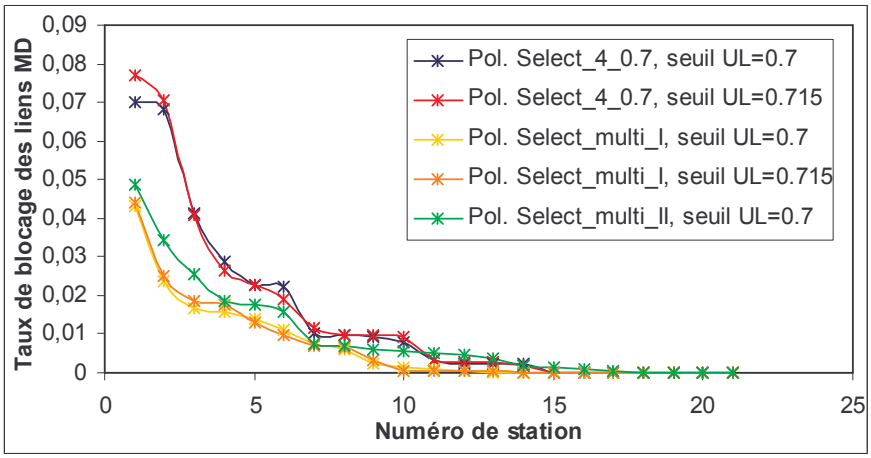


Figure 6.28 : Taux de blocage des liens de MD moyen par station (rangé dans l'ordre décroissant) pour les politiques *Select_4_0.7*, *Select_multi_I* et *Select_multi_II* sur la zone de la ville 1.

La figure 6.26 indique qu'il n'y a pas de différence majeure pour le taux de blocage moyen par station entre les différentes politiques testées, la politique *Select_multi_II* étant légèrement plus efficace, pour le tiers de stations les plus chargées. La figure 6.27 montre clairement que la politique *Select_multi_II* dégrade nettement le taux de coupure moyen par station pour le tiers de stations les plus chargées. En revanche, la politique *Select_multi_I* ne dégrade le taux de coupure moyen par station que pour les deux pires stations du réseau. Enfin, la figure 6.28 met en évidence que les politiques *Select_multi_I* et *Select_multi_II* permettent une amélioration quasiment homogène du taux de blocage des liens de MD moyen par station par rapport à la politique *Select_4_0.7*.

6.5. Premiers résultats d'optimisation de l'ingénierie dynamique avec l'algorithme du Q-learning

6.5.1. Mises en applications sur une zone réduite de la ville 2

La méthode d'optimisation des systèmes d'inférence floue par l'algorithme du *Q-learning* est testée sur une zone de la ville 2 de 32 secteurs, afin d'éviter des délais de simulations trop importants. La table 6.10 suivante résume les hypothèses de la phase d'apprentissage du *Q-learning*. Dans notre modélisation, nous avons défini un paramètre RRM à contrôler par SIF. Dans ce paragraphe, nous étudierons l'optimisation par apprentissage par renforcement d'un seul SIF simultanément, contrôlant uniquement le paramètre de seuil de charge cible DL.

Paramètres	Valeurs	Commentaires
Paramétrage 1	Coefficient de mixage égal à 1 Probabilité ε (EEP cf. 4.2.4.2) égale à 0.3 Limites des corrections: [-0.2, 0.2]	Noté: *SIF_Param_1*
Paramétrage 2	Coefficient de mixage égal à 6 Probabilité ε égale à 0.15 Limites des corrections: [-0.05, 0.05]	Noté: *SIF_Param_2*
Paramétrage 3	Coefficient de mixage égal à 7 Probabilité ε égale à 0.15 Limites des corrections: [-0.2, 0.2]	Noté: *SIF_Param_3*
Nombre d'itérations d'apprentissage	200000 itérations	Pour temps inter modifications de 4s
Paramètres contrôlés	Seuil de charge cible DL	Contrôle d'admission
Paramétrage de référence	Seuil de charge cible DL&UL=0.7 *Add_win* = 3.5 dB	Référence pour estimer les gains de performance
Front de P.U.	Seuil de charge cible DL dans {0.6, 0.7, ,0.8}	Voir sur les figures 6.29 et 6.30
Nombre d'objets de trafic	1	-
Nombre total initial de mobiles	800 mobiles	-
Service	Voix	Durée moyenne des appels = 100s
Taux d'arrivée des mobiles	8 s^{-1}	-
Charge moyenne par secteur	25 mobiles voix par secteur	La distribution de trafic étant homogène

Table 6.10 : Hypothèses de la phase d'apprentissage avec notre simulateur de réseau UMTS pour différents paramétrages de l'algorithme de *Q-learning*.

Phase d'exploitation

Suite à la phase d'apprentissage pour chacun des paramétrages décrits dans la table 6.10, nous appliquons le contrôle dynamique fourni par chacun des SIF en mode d'exploitation. La figure 6.29 décrit les performances de chacun de SIF dans le plan taux de coupure – taux de blocage.

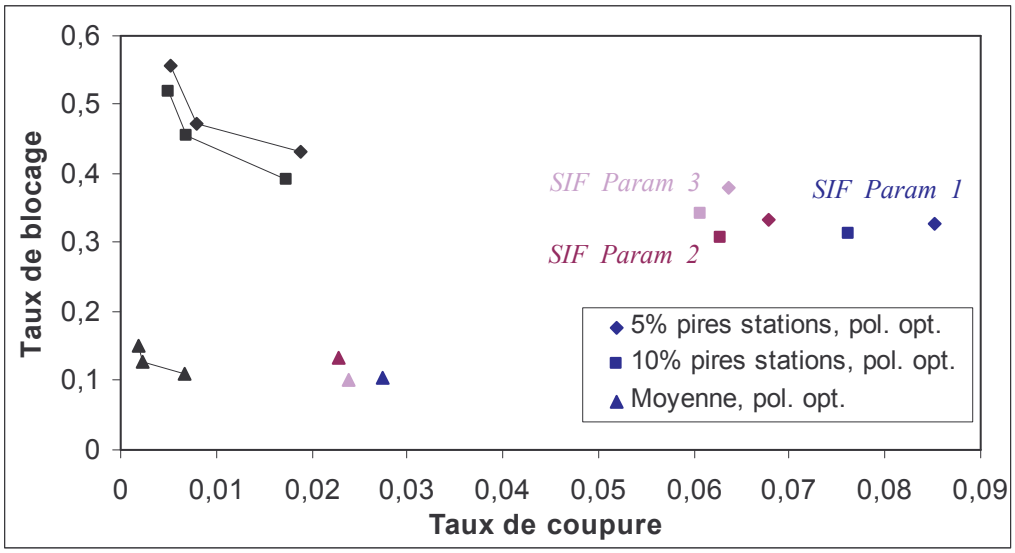


Figure 6.29 : Efficacité des politiques *SIF_Param_1*, *SIF_Param_2* et *SIF_Param_3* par rapport aux fronts P.U. dans le plan taux de coupure - taux de blocage.

Les politiques *SIF_Param_1, SIF_Param_2* et *SIF_Param_3* dégradent les performances en taux de coupure, par rapport aux fronts P.U, mais sans amélioration notable des performances en taux de blocage. Du point de vue de ces indicateurs, ces politiques fournissent un compromis entre taux de blocage et taux de coupure peu efficace. La table 6.11 décrit les causes de blocage et de coupure des appels pour chaque politique de paramétrage *on-line* :

Politiques	Paramétrage uniforme statique	*SIF_Param_1*	*SIF_Param_2*	*SIF_Param_3*
Nombre de mobiles tirés	159800	160161	160370	160297
Nombre total de mobiles rejetés	39084 Référence	34319	38369	34997
Nombre de mobiles bloqués à l'admission: CHARGE_DL	35999	24843	29781	26252
Nombre de mobiles bloqués à l'Admission: CHARGE_UL	2076	3034	3160	3283
Nombre de mobiles bloqués à l'admission: COUV_DL	668	1449	1333	1360
Nombre de mobiles bloqués à l'admission: COUV_UL	0	0	0	0
Nombre de mobiles bloqués à l'admission: PILOTE	0	0	0	0
Total de mobiles bloqués	**38743**	**29326**	**34274**	**30895**
Taux de blocage global	**24.2 %**	**18.3 %**	**21.4 %**	**19.3 %**
Nombre de mobiles coupés: COUV_DL	283	1521	1273	1411
Nombre de mobiles coupés: COUV_UL	0	0	0	0
Nombre de mobiles coupés: CHARGE_DL	0	3425	2777	2634
Nombre de mobiles coupés: CHARGE_UL	58	47	45	57
Total de mobiles coupés	**341**	**4993**	**4095**	**4102**
Taux de coupure global	**0.2 %**	**3.1 %**	**2.6 %**	**2.6 %**

Table 6.11 : Analyse des causes de rejet des mobiles pour les systèmes d'inférence floue *SIF_Param_1, SIF_Param_2* et *SIF_Param_3* par rapport au cas d'un paramétrage uniforme.

Les trois systèmes d'inférence floue ont permis une diminution des nombres de mobiles bloqués et rejetés dans le réseau. En revanche, ces trois politiques de contrôle ont nettement dégradé le nombre de mobiles coupés dans le réseau. Le meilleurs compromis est fourni par le SIF numéro 3, car le taux de coupure global est proche de la valeur 2% qui est un objectif classique des opérateurs de réseau mobile.

6.5.2. Comparaison avec l'optimisation par essaim de particules

Dans ce paragraphe, nous proposons de comparer les performances du SIF numéro 3 de la section précédente avec deux CLF optimisés par essaim de particules en multi-objectif respectivement pour l'algorithme d'admission seul (*Select_AC*) et simultanément pour le contrôle d'admission et la macro-diversité (*Select_AC+MD*) :

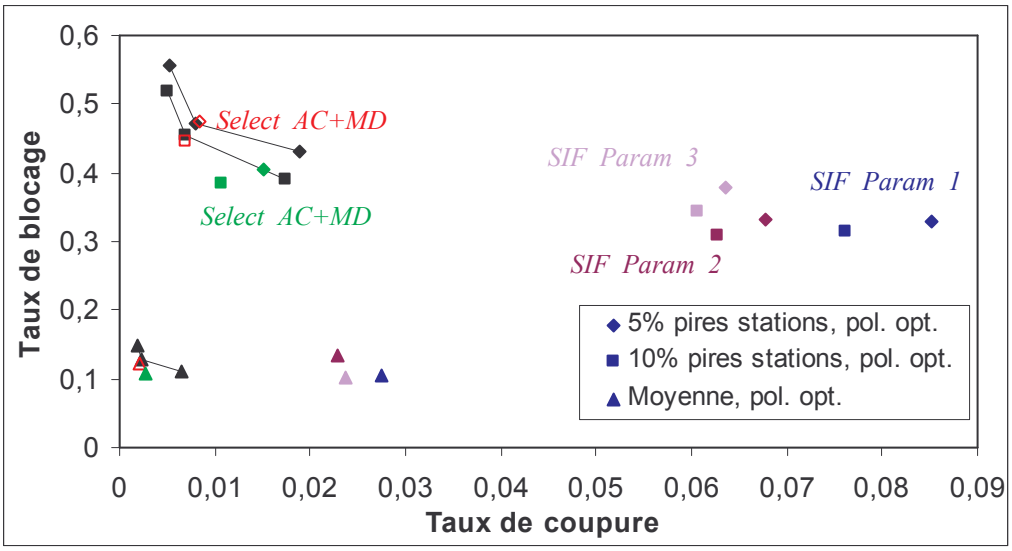


Figure 6.30 : Efficacité des politiques optimisées par OEP par rapport aux *SIF_Param_1*, *SIF_Param_2* et *SIF_Param_3* dans le plan taux de coupure - taux de blocage.

Les politiques sélectionnées par l'OEP, *Select_AC* et *Select_AC+MD*, fournissent des compromis différents en termes de taux de coupure et taux de blocage par rapport aux SIF optimisés au paragraphe précédent. La politique *Select_AC* ne fournit pas d'amélioration par rapport aux fronts de P.U. Sur la figure 6.29, le meilleur compromis entre coupure et blocage est donné par la politique *Select_AC+MD*. Cela n'est pas surprenant car c'est la seule politique ayant été optimisée simultanément pour le contrôle de l'admission et de la macro-diversité. La table 6.12 décrit les causes de blocage et de coupure des appels pour chaque politique de paramétrage *on-line* :

Politiques	Paramétrage uniforme statique	*SIF_Param_3*	*Select_AC*	*Select_AC+MD*
Seuil de charge cible UL statique	0.7	0.7	0.7	0.7
Nombre de mobiles tirés	159800	160297	160123	160300
Nombre total de mobiles rejetés	39084 Référence	34997 **Gain = +10.5%**	39126 **Perte = -0.1%**	34920 **Gain =+10.7%**
Nombre de mobiles bloqués à l'admission: CHARGE_DL	35999	26252	36372	30412
Nombre de mobiles bloqués à	2076	3283	1730	3483

l'Admission: CHARGE_UL				
Nombre de mobiles bloqués à l'admission: COUV_DL	668	1360	676	493
Total de mobiles bloqués	**38743**	**30895**	**38778**	**34388**
Taux de blocage global	**24.2 %**	**19.3 %**	**24.2 %**	**21.5 %**
Nombre de mobiles coupés: COUV_DL	283	1411	317	358
Nombre de mobiles coupés: CHARGE_DL	0	2634	0	0
Nombre de mobiles coupés: CHARGE_UL	58	57	31	174
Total de mobiles coupés	**341**	**4102**	**348**	**532**
Taux de coupure global	**0.2 %**	**2.6 %**	**0.2 %**	**0.3 %**

Table 6.12 : Analyse des causes de rejet des mobiles pour le système d'inférence floue *SIF_Param_3,* et les politiques Select_AC et Select_AC+MD par rapport au cas d'un paramétrage uniforme.

Par rapport à la politique optimisée par OEP pour le contrôle d'admission, le SIF numéro 3 (de la table 6.12) améliore nettement le nombre total de mobiles bloqués, 19.3% contre 24.2% de taux de blocage global. Mais la conséquence est une hausse importante du nombre total de mobiles coupés, 2.6 contre 0.2% de taux de coupure global. Les avantages de la politique de contrôle du SIF numéro 3 par rapport à la politique *Select_AC* n'apparaissait pas aussi clairement dans le plan taux de coupure – taux de blocage de la figure 6.30.

La politique *Select_AC* réduit le nombre de mobiles bloqués dans le réseau, par rapport au front de P.U sans dégrader le taux de coupure globale. Il n'est pas surprenant que cette politique fournisse un bon compromis, car elle a été optimisée simultanément pour le contrôle de l'admission et de la macro-diversité. Il sera nécessaire de tester une configuration équivalente où plusieurs SIF contrôlent simultanément plusieurs paramètres RRM, l'apprentissage ayant été fait de manière séquentielle.

6.6. Conclusions

Dans ce chapitre, l'amélioration des performances fournie par les contrôleurs dynamiques par rapport à un paramétrage uniforme et statique des algorithmes RRM a été étudiée. Ces méthodes de contrôle ont été testées avec les algorithmes d'admission et de macro-diversité. Nous avons tout d'abord introduit la notion de front de paramétrage uniforme dans les plans taux de coupure – taux de blocage. On a complété cette approche, par l'analyse des causes de rejet* avec et sans ingénierie dynamique. Ces méthodes d'analyse nous ont permis de comparer l'efficacité des contrôleurs optimisés par essaim de particules. Différents paramétrages initiaux du réseau et coefficients de mixage ont ainsi testés dans le mode mono-objectif. Ces contrôleurs ont ensuite été comparés avec le front optimal des solutions explorées avec l'OEP en multi-objectif. Il apparaît que le mode multi-objectif permet d'explorer une gamme de contrôleurs beaucoup plus large, fournissant des compromis plus

* Blocage ou coupure d'un appel.

variés en termes de taux de blocage et taux de coupure par station, de nombres de mobiles bloqués et coupés globalement dans le réseau.

L'algorithme du *Q-learning* a été testé sur le réseau de la ville 2 à 32 secteurs. Pour l'instant, les systèmes d'inférence floue optimisés avec le *Q-learning* ont produit des compromis différents des CLF optimisés par OEP en termes de taux blocage et taux de rejet par station, et nombres de mobiles bloqués et coupés globalement dans le réseau. Par rapport à la politique optimisée par OEP pour le contrôle d'admission, le SIF numéro 3 (de la table 6.12) améliore nettement le nombre total de mobiles bloqués mais au prix d'une hausse importante du nombre total de mobiles coupés. Les avantages de la politique de contrôle du SIF numéro 3 par rapport à la politique optimisée par OEP n'apparaissaient pas aussi clairement dans le plan taux de coupure – taux de blocage de la figure 6.30. Enfin, il est important de noter que le contrôleur optimisé par OEP simultanément pour l'admission et la macro-diversité, fournit un gain important sur le nombre de mobiles bloqués sans dégrader le nombre de mobiles coupés. De la même manière, il apparaît donc essentiel de développer et tester à présent l'utilisation simultanée de plusieurs SIF* permettant de contrôler dynamiquement différents paramètres RRM. L'apprentissage pour un tel ensemble de SIF serait réalisé de manière séquentielle. On commencerait avec un premier SIF qui serait optimisé en phase d'apprentissage du *Q-learning*. Dés lors, ce SIF serait uniquement utilisé en mode d'exploitation et la phase d'apprentissage serait lancée pour un deuxième SIF. En itérant ce procédé on pourrait ainsi optimiser tour à tour un ensemble de SIF contrôlant différents paramètres RRM. Suite à cette séquence d'apprentissage, on pourrait utiliser tous les SIF simultanément en mode d'exploitation pour optimiser dynamiquement pour toutes les cellules du réseau, l'ensemble des paramètres RRM sélectionnés.

* Chaque SIF ne contrôlant qu'un seul paramètre RRM.

7. Résultats obtenus avec l'ingénierie automatique en mode *off-line*

Dans ce chapitre, nous allons appliquer les concepts d'ingénierie automatique en mode *off-line* aux algorithmes de contrôle d'admission et de macro-diversité dans les environnements réalistes introduits au chapitre 6. Précisément nous allons appliquer la méthode décrite au paragraphe 5.2.2 pour optimiser, node B par node B, les paramètres de seuil de charge cible DL et/ou de *Add_win* à l'aide de simulations semi-dynamiques de réseau UMTS avec le logiciel développé en interne par France Télécom Recherche & Développement. L'optimisation par essaim de particules sera testée dans les modes mono et multi-objectif. Le front optimal du nuage des solutions explorées en multi-objectif sera ainsi comparé avec les meilleures solutions obtenues en mono-objectif.

7.1. Analyse de l'optimisation par essaim de particules en mode off-line

7.1.1. Méthode d'optimisation en mode *off-line*

Dans ce paragraphe, nous proposons de tester le schéma d'optimisation des CLF en mode *off-line* par essaim de particules, introduit au paragraphe 4.1.4. On reprend les fronts de performances dans le plan taux de coupure – taux de blocage obtenus au paragraphe 6.4.2.1. Sur ce graphique, on détermine le (ou les) paramétrage(s) initial(aux) uniformes le(s) plus efficace(s), c'est-à-dire donnant le meilleur compromis entre le taux de coupure, de blocage et de blocage des liens en MD. A partir de là, on teste différents coefficients de mixage en mono-objectif (qui dépendront de l'environnement étudié et du paramétrage initial uniforme choisi) pour l'optimisation par essaim de particules. Un ou plusieurs couples (paramétrage initial / coefficient de mixage) sont sélectionnés pour le mode *off-line* de l'algorithme d'OEP en mono-objectif. On pourra aussi choisir de lancer une optimisation en multi-objectifs (voir paragraphe 7.2). Enfin, les performances obtenues après optimisation par essaim de particules sont comparées avec les performances initiales données par les fronts de paramétrage uniformes, en moyenne sur l'ensemble des stations ou en considérant les pires stations du réseau.

7.1.2. Mises en applications sur le réseau de la ville 1

La méthode d'optimisation des contrôleurs de logique floue proposée est tout d'abord mise en applications sur la zone de la ville 1, afin d'éviter des délais de simulations trop importants.

Détermination des paramétrages initiaux donnant les meilleurs contrôleurs

Différents paramétrages initiaux du seuil de charge cible DL dans {0.6, 0.65, 0.7, 0.75} ont été testés ainsi que différents coefficients de mixage dans {2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6}. Les CLF ont été optimisés (simultanément[*] pour le contrôle d'admission et la macro-diversité) par essaim de particules en mode *off-line*, avec la fonction coût du 6.1.2.2, en mono-objectif pour différents couples de paramétrage initial et coefficient de mixage. La figure 7.1 compare les performances (moyennées sur 10 simulations) de la meilleure politique obtenue avec différents seuils de charge cible UL statique par rapport aux fronts P.U. (pour une charge moyenne de 42 mobiles voix par secteur) :

[*] Les CLF sont optimisés par une seule utilisation de l'algorithme d'OEP simultanément pour le C.A. et la M.D.

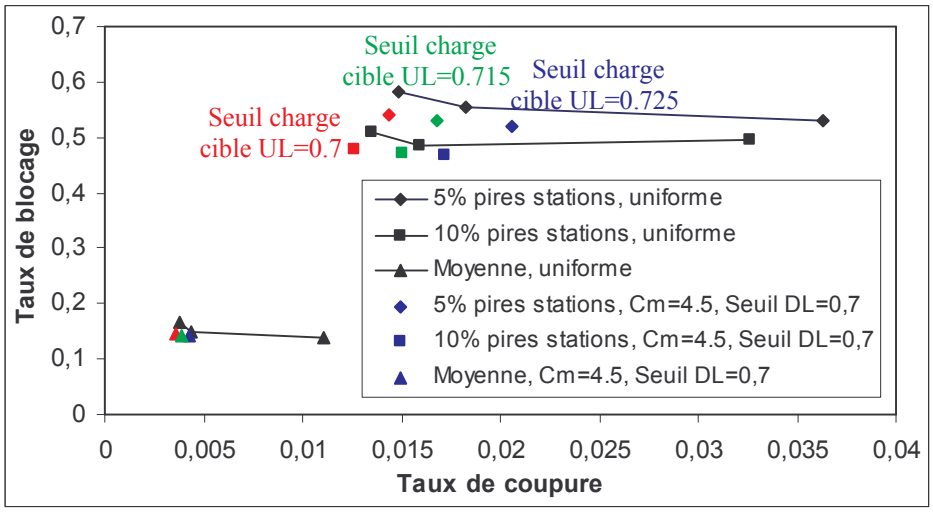


Figure 7.1 : Efficacité de la politiques optimisée en mono-objectif pour un paramétrage initial de seuil de charge cible DL égal à 0.7, un coefficient de mixage égal à 4 et pour un seuil de charge cible UL dans {0.7, 0.715, 0.725} par rapport aux fronts P.U. dans le plan taux de coupure / taux de blocage.

Sur la figure 7.1, la politique (*Cm=4.5, seuil init. DL=0.7*) en mode *off-line* a des performances moyennes en termes de taux de coupure et taux de blocage légèrement meilleures que le front de paramétrage uniforme. Pour les 5% et 10% pires stations du réseau, les améliorations de performances sont plus nettes pour les différents seuils de charge cible UL testés (0.7, 0.715 et 0.725). Nous sélectionnons donc cette politique *Select_4.5_0.7* dont nous allons analyser ci-dessous les performances d'optimisation du paramétrage *off-line*.

Analyse des résultats de la politique sélectionnée

Les hypothèses de l'optimisation *off-line* des paramètres RRM sont résumées dans la table 7.1. Nous allons ci-dessous analyser l'impact de cette optimisation *off-line* sur les causes de coupure et blocage des appels et comparer avec la situation initiale d'un paramétrage uniforme.

Paramètres	Valeurs	Commentaires
Politique testée	*Select_4.5_0.7*	Seuil de charge cible UL statique égal à 0.7, 0.715 et 0.725
Nombre de simulations	10	Résultats moyennés sur 10 simulations
Nombre d'itérations	3000 itérations	Avec un intervalle de temps de 4 s
Paramètres optimisés en *off-line*	Seuil de charge cible DL Taille de la fenêtre d'ajout	Contrôle d'admission Macro-diversité
Paramétrage de référence	Seuil de charge cible DL =0.7 *Add_win* = 3.5 dB Seuil de charge cible UL =0.7	Référence pour estimer les gains de performance
Nombre d'objets de trafic	3	-
Nombre total initial de mobiles	900 mobiles	600 sur l'objet de trafic 1, et 150 sur les objets de trafic 2 et 3
Service	Voix	Durée moyenne des communications = 100s
Taux d'arrivée des mobiles par objet de trafic	6 s^{-1}, 1.5 s^{-1}, 1.5 s^{-1}	Respectivement pour les objets de trafic 1, 2 et 3
Charge moyenne par secteur	42 mobiles voix par secteur	En faisant l'hypothèse d'une distribution de trafic homogène

Table 7.1 : Hypothèses de simulations d'optimisation *off-line* avec notre simulateur de réseau UMTS.

Ci-dessous, nous fournissons, dans la table 7.2, les gains de performances obtenus avec la politique *Select_4.5_0.7* en mode *off-line* sur la zone de la ville 1 par rapport à un paramétrage uniforme par défaut (seuil de charge cible DL=0.7 et *Add_win*=3.50 dB) :

Politique optimisée par OEP	% gains en coupure			% gains en blocage			% gains en blocage des liens de MD		
	5% pires	10% pires	Moyenne	5% pires	10% pires	Moyenne	5% pires	10% pires	Moyenne
Politique *Select_4.5_0.7* avec un seuil de charge cible UL = 0.7	**+21%**	**+21%**	**+18%**	**+3%**	**+1%**	**+2%**	**+40%**	**+36%**	**+26%**
Politique *Select_4.5_0.7* avec un seuil de charge cible UL = 0.715	**+8%**	**+5%**	**+10%**	**+4%**	**+2%**	**+4%**	**+35%**	**+27%**	**+22%**
Politique *Select_4.5_0.7* avec un seuil de charge cible UL = 0.725	**-13%**	**-8%**	**+2%**	**+6%**	**+3%**	**+4%**	**+29%**	**+22%**	**+18%**

Table 7.2 : Pourcentages des gains en taux de blocage, de coupure et de blocage des liens de MD pour la politique *Select_4.5_0.7* en mode *off-line* par rapport à un paramétrage uniforme (seuil de charge cible DL=0.7 et *Add_win*=3.5 dB).

La politique *Select_4.5_0.7* améliore nettement les performances par rapport aux fronts P.U, pour les taux de coupure et taux de blocage des liens en MD. En comparant l'efficacité de cette politique pour différents seuils de charge cible UL de 0.7, 0.715 et 0725, on s'aperçoit que l'augmentation de ce seuil dégrade nettement le taux de coupure, modérément le taux de blocage des liens de MD et améliore le taux de blocage. La table 7.3 suivante permet de mettre en évidence les causes de blocage et coupure des appels pour chaque politique de paramétrage :

Politiques	Paramétrage uniforme statique	Politique *Select_4.5_0.7*		
Seuil de charge cible UL statique	0.7	0.7	0.715	0.725
Nombre de mobiles tirés	108529	108864	108861	109029
Nombre total de mobiles rejetés	**24127** Référence	**24028** **Gain=+0.4%**	**23695** **Gain=+1.8%**	**23726** **Gain=+1.7%**
Nombre de mobiles bloqués à l'admission: CHARGE_DL	18894	17210	17966	18665
Nombre de mobiles bloqués à l'Admission: CHARGE_UL	4883	6529	5395	4712
Nombre de mobiles bloqués à l'admission: COUV_DL	25	18	18	16
Nombre de mobiles bloqués à l'admission: COUV_UL	11	10	11	12
Nombre de mobiles bloqués à l'admission: PILOTE	1	2	2	2
Total de mobiles bloqués	**23814**	**23769**	**23392**	**23406**
Taux de blocage global	**21.9%**	**21.8%**	**21.5%**	**21.5%**
Nombre de mobiles coupés: COUV_DL	212	163	176	187
Nombre de mobiles coupés: COUV_UL	65	69	74	68
Nombre de mobiles coupés: CHARGE_DL	0.7	0.5	0,1	0,2
Nombre de mobiles coupés: CHARGE_UL	35	27	54	65
Total de mobiles coupés	**313**	**259**	**304**	**320**
Taux de coupure global	**0.3%**	**0.2%**	**0.3%**	**0.3%**

Table 7.3 : Analyse des causes de rejet des mobiles pour la politique *Select_4.5_0.7* en mode *off-line* par rapport au cas d'un paramétrage uniforme (seuil charge cible DL=0.7 et *Add_win*=3.5 dB).

Sur la table 7.3, l'optimisation *off-line* du paramètre seuil de charge cible DL permet à la politique *Select_4.5_0.7* de diminuer le nombre de mobiles bloqués à l'admission à cause du niveau de charge DL. Cependant cette baisse du nombre de mobiles bloqués sur le lien DL s'est accompagnée d'une hausse du nombre de mobiles bloqués à l'admission à cause du niveau de charge UL. Pour réduire cet effet, on a utilisé la même politique *Select_4.5_0.7* mais avec un seuil de charge cible UL statique égal respectivement à 0.715 et 0.725. On obtient ainsi des compromis intéressants qui ont permis de réduire le nombre total de mobiles rejetés (bloqués ou coupés), respectivement de 1.8% et 1.7% par rapport au cas du paramétrage statique et uniforme. Les gains en mode *off-line* sur le pourcentage de mobiles rejetés sont inférieurs aux gains obtenus en mode *on-line* sur la table 6.6 dans les mêmes conditions initiales (paramétrage du seuil de charge cible DL égal à 0.7).

7.2. Comparaison des résultats de l'OEP en mode off-line en mono et multi-objectif

Dans ce paragraphe, deux solutions du front optimal de l'algorithme d'optimisation par essaim de particules en multi-objectif pour le mode *off-line* vont être comparées à la politique *Select_4.5_0.7* optimisée en mono-objectif pour un coefficient de mixage égal à 4.5. Nous avons lancé l'optimisation en multi-objectif pour avec la même fonction coût qu'au 6.1.2.2, privilégiant l'optimisation des performances des pires stations, sur le réseau de la ville 1 (fig. 6.1) et pour un densité initiale de 900 mobiles (mêmes hypothèses que pour l'optimisation mono-objectif au paragraphe 7.1.2). La figure 7.2 suivante décrit l'ensemble des solutions explorées au cours de l'optimisation en multi-objectifs pour le mode *off-line* dans le plan fonction coût de blocage – fonction coût de coupure.

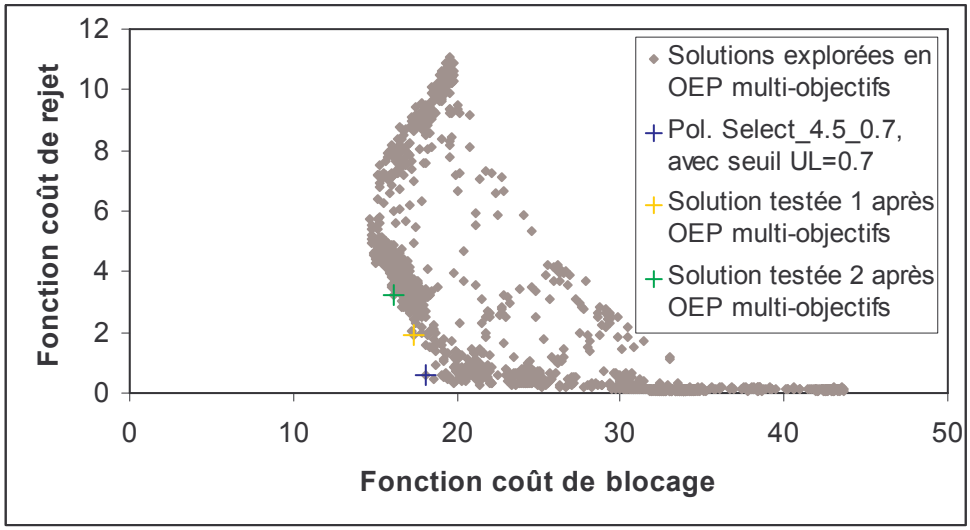


Figure 7.2 : Solutions explorées en OEP multi-objectifs (dont les solutions 1 et 2 du front optimal) comparées avec la politique *Select_4.5_0.7* obtenue en mono-objectif pour un coefficient de mixage égal à 4.5 dans le plan fonction coût de blocage – fonction coût de coupure.

La solution *Select_4.5_0.7* optimisée en mono-objectif appartient au front de solutions de rang 1. Nous sélectionnons deux autres solutions optimales pour le mode *off-line*, *Select_multi_1_Off* et *Select_multi_2_Off*, appartenant à ce front (meilleures en fonction coût de blocage mais moins efficaces en fonction coût de coupure) dont nous allons comparer les performances effectives par rapport à la politique *Select_4.5_0.7*.

La figure 7.3 compare les performances des politiques *Select_4.5_0.7*, *Select_multi_1_Off* et *Select_multi_2_Off* par rapport aux fronts P.U. dans le plan taux de coupure / taux de blocage, moyennées sur dix simulations de réseau avec ingénierie automatique en mode *off-line*.

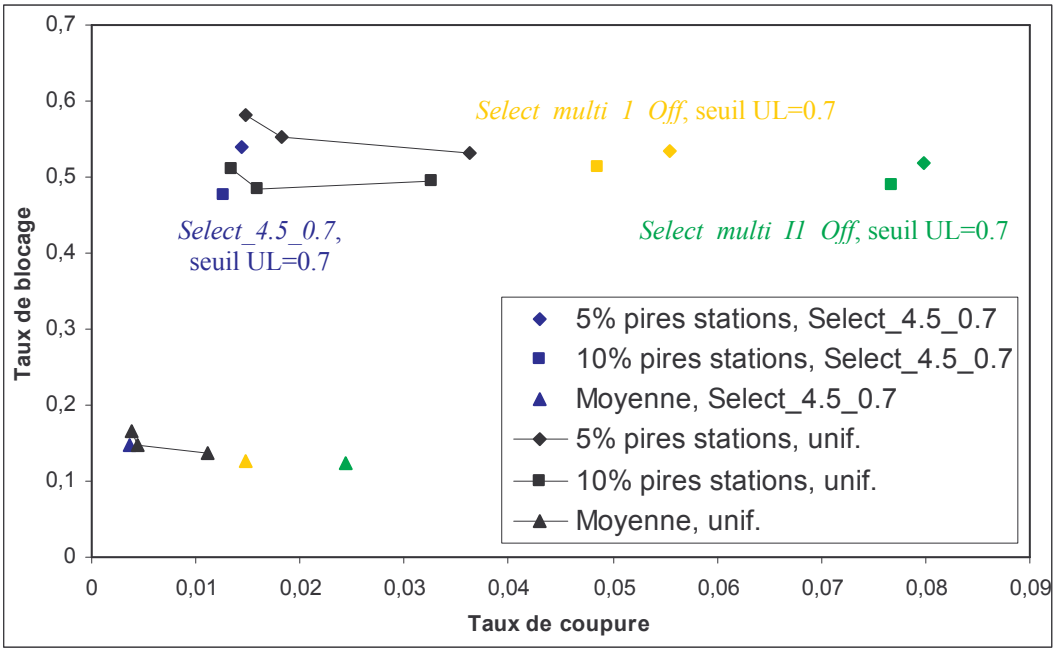


Figure 7.3 : Efficacité des politiques *Select_4.5_0.7*, *Select_multi_1_Off* et *Select_multi_2_Off* par rapport aux fronts P.U. dans le plan taux de coupure / taux de blocage.

Les politiques *Select_multi_1_Off* et *Select_multi_2_Off* dégradent les performances en taux de coupure, par rapport à la politique *Select_4.5_0.7* (comme indiqué a priori par la fonction coût de coupure de la figure 7.2), mais sans amélioration notable des performances en taux de blocage. Du point de vue de ces indicateurs, les politiques *Select_multi_1_Off* et *Select_multi_2_Off* fournissent un compromis entre taux de blocage et taux de coupure moins efficace que la politique *Select_4_0.7*. La table 7.4 suivante permet de mettre en évidence les causes de blocage et coupure des appels pour chaque politique de paramétrage, moyennées sur dix simulations de réseau ingénierie automatique en mode *off-line* :

Politiques	Paramétrage uniforme statique	Politique *Select_multi_I_Off*	Politique *Select_multi_II_Off*
Seuil de charge cible UL statique	0.7	0.7	0.7
Nombre de mobiles tirés	108529	108971	108712
Nombre total de mobiles rejetés	24127 Référence	21810 **Gain=+9.6%**	22496 **Gain=+6.8%**
Nombre de mobiles bloqués à l'admission: CHARGE_DL	18894	11363	12899
Nombre de mobiles bloqués à l'Admission: CHARGE_UL	4883	9081	7290
Nombre de mobiles bloqués à l'admission: COUV_DL	25	37	53
Nombre de mobiles bloqués à l'admission: COUV_UL	11	10	10
Nombre de mobiles bloqués à l'admission: PILOTE	1	3	3
Total de mobiles bloqués	**23814**	**20494**	**20255**
Taux de blocage global	**21.9%**	**18.8%**	**18.6%**
Nombre de mobiles coupés: COUV_DL	212	192	307

Nombre de mobiles coupés: COUV_UL	65	73	53
Nombre de mobiles coupés: CHARGE_DL	0.7	961	1822
Nombre de mobiles coupés: CHARGE_UL	35	89	60
Total de mobiles coupés	**313**	**1315**	**2241**
Taux de coupure global	**0.3%**	**1.2%**	**2.1%**

Table 7.4 : Analyse des causes de rejet des mobiles pour les politiques *Select_multi_I_Off* et *Select_multi_II_Off* par rapport au cas d'un paramétrage uniforme (seuil charge cible DL=0.7 et *Add_win*=3.5 dB).

La politique *Select_multi_I_Off* obtenue par OEP en mode multi-objectif a permis une baisse notable du taux de blocage global, de 21.9% à 18.8%, avec une hausse maîtrisée du taux de coupure global, de 0.3% à 1.2%. Un tel compromis est en général acceptable pour les opérateurs de réseau mobile, le taux de coupure moyen cible étant de 2%. La politique *Select_multi_II_Off* fournit des performances nettement moins efficaces. Le nombre total de mobiles bloqués est légèrement meilleur que pour la politique *Select_multi_I_Off*, mais le taux de coupure global dans le réseau a quasiment doublé. Sur la figure 7.3, cette observation est confirmée pour les taux de blocage et de coupure par station.

7.3. Conclusions

Dans ce chapitre, l'amélioration des performances fournie par l'optimisation *off-line* des paramètres RRM par rapport à un paramétrage uniforme des algorithmes RRM a été étudiée dans le cas de l'optimisation par essaim de particules. Ces méthodes de contrôle ont été testées avec les algorithmes d'admission et de macro-diversité. Les méthodes d'analyse, introduites au chapitre 6, nous ont permis de comparer l'efficacité des contrôleurs optimisés pour le mode *off-line*. Différents paramétrages initiaux du réseau et coefficients de mixage ont ainsi testés dans le fonctionnement mono-objectif. Ces contrôleurs optimisés ont ensuite été comparés avec le front optimal des solutions explorées avec l'OEP en multi-objectif. Il apparaît que le mode multi-objectif permet d'explorer une gamme de contrôleurs beaucoup plus large, fournissant des compromis plus variés en termes de taux de blocage et taux de coupure par station, nombres de mobiles bloqués et coupés globalement dans le réseau. Ainsi, la politique *Select_multi_I_Off* obtenue par OEP en mode multi-objectif a permis une baisse notable du taux de blocage global, avec une hausse maîtrisée du taux de coupure global à 1.2%. Un tel compromis est en général acceptable pour les opérateurs de réseau mobile, le taux de coupure moyen cible étant de 2%.

A présent, il s'agira de tester l'optimisation et l'utilisation des SIF par l'algorithme de *Q-learning*, dans le cas d'un fonctionnement *off-line*. En effet comme décrit au chapitre 5, les figures 5.10, 5.11 et 5.12 proposent l'adaptation du *Q-learning* à un tel mode de fonctionnement. Tout comme pour le mode *on-line*, il sera aussi nécessaire de développer et tester le fonctionnement simultané de plusieurs SIF, chaque SIF optimisant le paramétrage *off-line* d'un paramètre RRM différent. Enfin, si l'optimisation *off-line* des paramètres RRM par apprentissage par renforcement se révèle efficace, on pourra tester en simulations logicielles, voire sur un réseau expérimental le nouveau mode de fonctionnement introduit par la figure 5.13, le mode "semi *on-line*".

8. Conclusions et perspectives

A l'ouverture commerciale des réseaux UMTS Européens, le paramétrage des algorithmes de gestion de la ressource radio était uniforme et statique. Nous pressentions qu'une adaptation des paramètres RRM, station par station, pourrait permettre d'améliorer les performances du réseau et la qualité de service perçue par les utilisateurs. Nous avons étudié le paramétrage station par station de deux algorithmes essentiels des réseaux W-CDMA : le contrôle d'admission et la macro-diversité. Cette analyse a permis d'extraire d'une part les indicateurs de qualité du réseau significatifs de l'état des cellules, et d'autre part les paramètres les plus importants de ces algorithmes RRM.

Deux stratégies peuvent alors être définies pour le contrôle automatique des stations du réseau mobile. La première consiste en une adaptation *off-line* des valeurs des paramètres RRM. Une telle solution présente l'avantage d'être utilisable à moyen terme pour optimiser les réseaux UMTS existants, par une modification du paramétrage des stations au niveau d'un nœud central de management comme l'OMC. La seconde stratégie vise à paramétrer dynamiquement les algorithmes RRM du réseau UMTS. Cette approche est plus contraignante, car elle nécessite la prise en compte des contraintes temporelles liées à la normalisation actuelle de l'UMTS (*Releases* R'99, R4 et R5) et le développement d'une architecture appropriée de contrôle dynamique. Apres avoir ainsi défini les modes de fonctionnement possibles pour l'ingénierie automatique, nous avons introduit deux cadres théoriques pour le développement de nos contrôleurs.

La première modélisation consiste à définir nos modules de contrôle comme des processus de type FKBC (*Fuzzy Knowledge Based Controller*). En se fondant sur notre expérience du paramétrage RRM des réseaux mobiles, nous avons construit deux systèmes de matrices de décisions contrôlant respectivement un paramètre du contrôle d'admission et un paramètre de l'algorithme de macro-diversité. Ces systèmes de matrice ont été décrits de manière synthétique par des représentations paramétriques. Afin d'améliorer l'efficacité des contrôleurs de logique floue ainsi définis, une méthode d'optimisation itérative stochastique a été adaptée : l'optimisation par essaim de particules (OEP). Deux utilisations possibles ont été développées pour cet algorithme d'OEP : les versions mono-objectif et multi-objectif. La première application permet de converger rapidement vers des solutions correspondant à un compromis particulier entre les indicateurs de blocage et de coupure dans le réseau. En revanche, la seconde version permet d'explorer l'ensemble des solutions possibles pour nos contrôleurs de logique floue. L'optimisation par essaim de particules a permis d'améliorer nettement l'efficacité des contrôleurs automatique, comme décrit dans les chapitres 6 et 7. Mais plusieurs caractéristiques de cette technique d'optimisation nous ont contraint à tester et étudier d'autres méthodes d'optimisation dans le contexte du contrôle automatique des paramètres du réseau UMTS : perte d'informations dans les fonctions coûts globales, utilisation de l'OEP exclusivement avec des simulations logicielles[*], manque de souplesse et de généricité dans la définition des CLF.

Du fait des limites de l'approche par OEP, nous avons donc étudié la possibilité d'utiliser des méthodes d'apprentissage par renforcement pour contrôler le paramétrage des réseaux mobiles. Pour être le plus générique possible, nous avons choisi de modéliser ce problème d'optimisation par un système d'inférence floue (SIF). Parmi les différents algorithmes d'apprentissage par renforcement, nous avons choisi d'adapter la méthode du *Q-learning* à

[*] Pas d'application possible dans les réseaux réels.

notre SIF. L'utilisation du SIF permet de ramener notre problème d'optimisation sur des ensembles continus à un processus de décisions Markovien. Plusieurs propriétés du *Q-learning* nous permettent d'espérer que cette méthode devrait être plus appropriée à notre problème : possibilité d'apprendre simultanément et parallèlement de toutes les stations du réseau, utilisation des contrôleurs (pré-calibrés en simulations) dans les réseaux réels, généricité de la méthode de contrôle par *Q-learning*[*]. Au chapitre 6 en mode *on-line*, les systèmes d'inférence floue optimisés avec le *Q-learning* ont produit des politiques de contrôle différentes des CLF optimisés par essaim de particules, avec des gains plus importants notamment sur les taux de blocage dans le réseau. Pour l'instant, les contrôleurs basés sur les SIF ne permettent de contrôler simultanément qu'un seul paramètre RRM[**]. Il apparaît donc essentiel de développer et tester à présent l'utilisation simultanée de plusieurs SIF[***] permettant de contrôler dynamiquement différents paramètres RRM. La phase d'apprentissage pour un tel ensemble de SIF serait réalisée de manière séquentielle.

Lors de prochains travaux, nous chercherons à tester l'efficacité des contrôleurs basés sur l'algorithme de *Q-learning* dans les modes *off-line* et "semi *on-line*" décrits au chapitre 5. En effet, ces deux modes de fonctionnement présentent l'avantage d'être utilisables à court terme dans les réseaux UMTS réels, contrairement au contrôle dynamique des paramètres RRM. Par ailleurs, comme indiqué au chapitre 2, on pourra aussi étendre et adapter l'utilisation de l'ingénierie automatique aux autres algorithmes et paramètres de configuration des réseaux UMTS : algorithmes de contrôle de charge, *packet scheduling*, contrôle de puissance maximale, voire les paramètres antennaires. L'intégration de nouveaux indicateurs de performance ou de paramètres à contrôler est d'ailleurs facilitée dans le cadre de contrôleurs fondés sur le *Q-learning*. Cette méthode de contrôle automatique nous apparaît suffisamment générique pour s'adapter aussi à différents scénarii de gestion coordonnée multi-système.

Une partie des suites des travaux de contrôle automatique ont pour cadre notre participation au projet coopératif EUREKA - CELTIC Gandalf. Les objectifs de Gandalf sont de développer des méthodes de surveillance (*monitoring*), d'analyse et de gestion des pannes (*troubleshooting*), et de contrôle automatique des performances et de la qualité de service dans le contexte de réseaux d'accès radio multi-systèmes coopératifs [Stu05]. Dans ce cadre, nous allons donc étudier l'adaptation des méthodes d'ingénierie automatique à un environnement hétérogène de réseaux d'accès radio : UMTS, GSM / GPRS, WLAN, EDGE. Pour cela, il sera nécessaire de prendre en compte l'ensemble des algorithmes en développement favorisant la gestion commune des ressources radio :

- Prise en compte des algorithmes *Advanced-RRM*, intra système et inter fréquences.
- Etude des algorithmes *Common-RRM*[****] inter systèmes en développement au 3GPP.
- Etude des algorithmes intégrant la gestion des technologies WLAN.
- Analyse et intégration des nouvelles stratégies de *Joint-RRM* (intégrant l'ensemble des technologies d'accès radio), fondées sur des nouvelles méthodologies de contrôle.

L'ingénierie automatique multi-système de la ressource radio pourrait permettre d'optimiser l'exploitation des différents réseaux d'accès radio, en sélectionnant la technologie d'accès radio et le niveau de qualité de service les plus adaptés aux demandes des utilisateurs.

---

[*] Introduction simple de nouveaux indicateurs de performances dans un contrôleur basé sur des SIF (avec la possibilité d'intégrer la connaissance acquise auparavant), de nouveaux algorithmes à contrôler, de nouveaux systèmes à prendre à compte…

[**] Le seuil de charge cible DL pour les tests du chapitre 6.

[***] Chaque SIF ne contrôlant qu'un seul paramètre RRM.

[****] Au 3GPP, les algorithmes *Common-RRM* concernent les systèmes GSM/GPRS et UMTS.

9. Références

[3GPP1] TR 25.881 "Improvement of RRM across RNS and RNS/BSS," Release5, 2001.

[3GPP2] TS 23.234 "3GPP system to WLAN interworking: system description", Release 6.

[3GPP3] TS 25.215 "Physical layer – Measurements (FDD)," v 3.5.0, Dec. 2000.

[3GPP4] TS 25.331 "Radio Resource Control (RRC) Protocol Specification," Release5, 2002-2003.

[Agu04] "A fuzzy-Neural Based Approach for Joint Radio Ressource Management in Beyond 3G Framework," R. Agusti, O. Sallent, J. Pérez-Romero, L. Giupponi, 1st Int. Conf. on Quality of Service in Heterogeneous Wired/Wireless Networks (QSHINE), Dallas, Oct. 2004.

[Ben04] "UMTS design strategies based on indicator matrix approach," S. Ben Jamaa, Z. Altman, A. Ortega and B. Fourestié, *IEEE Int. Conf. on Communications ICC 2004*, Paris, France, June 2004.

[Ben05a]"Quality indicator matrices and their contribution to WCDMA network design," S. Ben Jamaa, H. Dubreil, Z. Altman, A. Ortega, *IEEE Transactions on Vehicular Technology*, vol. 54, May 2005.

[Ben05b] "Steered optimization strategy for automatic cell planning of UMTS networks," S. Ben Jamaa, Z. Altman, J.M. Picard and A. Ortega, *Vehicular Technology Conference Spring 2005*, Stockholm, June 2005.

[Brev04] Brevet EP 1 555 842 A1 intitulé "Procédé de détermination de la capacité d'un réseau de télécommunication cellulaire," H. Dubreil et A. Daviaud (FTR&D/DMR/IIM), publié le 20 juillet 2005.

[Cle02] "The Particle Swarm-Explosion, Stability, and Convergence in a Multidimensional Complex Space," M. Clerc, J. Kennedy, IEEE Transactions on Evolutionary Computation, vol. 6, pp. 58-73, 2002.

[Cap04] "Call Admission Control technique for UMTS," A. Capone, S. Redana, Vehicular Technology Conf. 2001, VTC 2001 Spring IEEE, vol.4, pp. 2828-2832, 2001.

[Cha01a] "Calculating the Outage Probability in a CDMA Network with Spatial Poisson Traffic," C.C. Chan and S.V. Hanly, IEEE Trans. on Vehicular Technology, vol. 50, no. 1, Jan. 2001.

[Cha01b] "Mobility Management Incorporating Fuzzy Logic for a Heterogeneous IP Environment," P.M.L. Chan, R.E. Sheriff, Y.F. Hu, P. Conforto, and C. Tocci, IEEE Communications Magazine, pp. 42-51, Dec. 2001.

[Cha04] "A Call Admission Control in Mobile Cellular CDMA Systems using Fuzzy Associative Memory," R.N.S. Chandra, D. Sarkar, IEEE Communications Society, 2004.

[Cho01] "Forward-link capacity of a DS/CDMA system with mixed multirate sources," W. Choi, J.Y. Kim, IEEE Trans. Veh. Technol., vol. 50, no. 3, May 2001.

[Cor96] "A hybrid genetic algorithm-evolution strategy process for learning fuzzy logic controller knowledge bases," O. Cordon and F. Herrera, in Genetic Algorithms and Soft Computing, Berlin Germany, pp. 251-278, 1996.

[Dia96] "An Introduction to Fuzzy Control," D. Diankov, H. Hellendoorn and M. Reinfrank, Springer-Verlag, 2nd Ed., 1996.

[Din04] "A call admission control strategy based on fuzzy logic for W-CDMA Systems," P. Dini, S. Guglielmucci, IEEE Communications Society, 2004.

[Dub03] "Radio capacity estimation of a CDMA system with multi-service traffic," H. Dubreil and A. Daviaud, *WWC 2003*, San Francisco, May 2003.

[Dub05] "Particle swarm optimization of fuzzy logic controller for high quality RRM auto-tuning of UMTS networks," H. Dubreil, Z. Altman, V. Diascorn, J.M. Picard, and M. Clerc, *Vehicular Technology Conference Spring 2005*, Stockholm, June 2005.

[El-A01] "A model-based Q-learning scheme for wireless channel allocation with prioritized handoff," E.S. El-Alfy, Yu-Dong Yao, H. Heffes, Global Telecommunications Conference, vol. 6, pp. 3668 – 3672, Nov. 2001.

[Ela03a] "On the capacity of multi-cell UMTS," S-E. Elayoubi, T. Chahed and G. Hébuterne, IEEE Globecom 2003, San Francisco, pp. 487-491, Dec. 2003.

[Ela03b] "Resource Management in UMTS : from Admission Control to Packet Scheduling," S-E. Elayoubi, T. Chahed and G. Hébuterne", ICON2003, Sydney, Oct. 2003.

[Gar04] "Apprentissage par renforcement – tutorial," P. Garcia, IRISA, n° 1655, 2004.

[Gaï05] "A new architecture for autonomous network equipment," D. Gaïti, UTT, G. Pujolle, LIP6, M. Salaun, France Telecom, and H. Zimmermann, Ginkgo-Networks, MAN'2005 IFIP Open Conference on Metropolitan Area Networks Architecture, protocols, control, and management HCMC, Viet Nam, April 11-13, 2005

[Glo97] "Fuzzy Q-Learning," P.Y. Glorennec and L. Jouffe, in Proc. Fuzz-IEEE, Sixth Int. Conf. on Fuzzy Systems, Barcelona Spain, pp. 659-662, 1997.

[Gus02] "Real-time performance monitoring and optimization of cellular systems," P. Gustas, P. Magnusson, J. Oom, and N. Storm, Ericsson Review, n. 1, pp. 4-13, 2002.

[Hil00] WCDMA Downlink capacity estimation," K. Hiltunen, R. De Bernardi, Conference Proceeding, VTC 2000, p. 992-996.

[Hol02] "WCDMA for UMTS, Radio Access for Third Generation Mobile Communications," H. Holma and A. Toskala, Wiley, 2nd Edition, 2002.

[Hom00] "Adaptation of CDMA soft handoff thresholds using fuzzy inference system," B. Homnan, V. Kunsriruksakul, W.Benjapolakul, Proc. IEEE ICPWC, pp. 259-263, Nov. 2000,

[Hom04] "QoS controlling soft hand-over based on simple step control and a fuzzy inference system with the gradient descent method," B. Homnan, W.Benjapolakul, IEEE Transactions on vehicular technology, vol. 53, no. 3, may 2004.

[Jor92] "Forward models: Supervised learning with a distal teacher," M.I. Jordan, Cogn. Sci., vol.16, pp. 307-354, 1992.

[Jou98] "Fuzzy Inference System Learning by reinforcement Methods," L. Jouffe, IEEE Trans. Syst., Man, Cybern., vol. 28, pp. 338-355, Aug. 1998.

[Kim00] "Capacity estimation for an SIR-based power controlled CDMA system supporting ON-OFF traffic," D.K. Kim, D.K. Sung, IEEE Trans. Veh. Technol., vol. 49,no. 4, July 2000.

[Lai00] "Verification of 3G radio network dimensioning rules with static network simulations," J. Laiho-Steffens, A. Wacker, K. Sipilä, VTC 2000, p. 478-482.

[Lai01] "Radio network planning process and methods for WCDMA," J. Laiho, A. Wacker, Ann. Télécommun., 56, no. 5-6, 2001, p. 317-331.

[Lin91] "Neural-network-based fuzzy logic control and decision system," C. Lin and C.S.G. Lee, IEEE Trans. Comput., vol. 40, pp. 1320-1336, Dec.1991.

[Lo03] "A Neural Ressource Manager for Hierarchical Cellular Systems Supporting Multimedia Services," K.R. Lo, C.J. Chang, and C.B. Shung, IEEE Trans. on Vehicular Technology, vol. 52, no. 5, Sept. 2003.

[Mag01] "An Architecture for Self-Tuning Cellular Systems," P. Magnusson, J. Oom, *Proc. of the 2001 IEEE/IFIP International symposium on Integrated Network Management*, pp. 231-245, 2001.

[Mea02] "Common Radio Resource Management (CRRM)," F. Meago, COST273, May 2002.

[Nas05] "Dynamic Radio Resource Management in Wireless Networks: Towards autonomic mobile networking," R. Nasri, Z. Altman and H. Dubreil, IEEE Communication Magazine, 2005, submitted.

[Per02] "Admission Control for Different UE-MAC Algorithms in UTRA-FDD", Proc. 3rd Int. Conf. on 3G Mobile Telecom. Technologies, May 2002.

[Pic04] "Dynamic Control of UMTS Networks by Load Target Tuning," J.M. Picard, H. Dubreil, F. Garabedian, Z. Altman, *IEEE International Symposium VTC 2004*, Genoa, Italy, May. 11-14, 2004.

[Pic05] "Automatic Cell Planning Strategies for UMTS Networks," J-M. Picard, Z. Altman, S. Ben Jamaa, M. Demars, H. Dubreil, B. Fourestié and A. Ortega, *International Journal of Mobile Network Design and Innovation*, accepted.

[Sal03] "Provisioning multimedia wireless networks for better QoS: RRM strategies for 3G W-CDMA," O.Sallent, J. Perez-Romero, R. Agusti and F. Casadevall, IEEE Communications Magazine, Vol. 41, Iss. 2, pp. 100 – 106, Feb. 2003.

[Sal04] "Autonomic networking pour un opérateur," M. Salaun, Ecole DNAC, Egypte, Dec. 2004.

[Sal05] "On autonomic communications," M. Salaun, MAN'2005 IFIP Open Conference on Metropolitan Area Networks Architecture, protocols, control, and management HCMC, Viet Nam, April 11-13, 2005.

[Sen04] "Call admission control in cellular networks: a reinforcement learning solution," S.M. Senouci, A.L. Beylot and G. Pujolle, Int. J. Network Mgmt, No. 14, pp 89-103, 2004.

[Sha94] "Microcell engineering in CDMA cellular networks," J. Shapira, IEEE Trans. on Vehicular Technology, vol. 43, Iss. 4, pp. 817 – 825, Nov. 1994.

[Spi00] "Estimation of capacity and required transmision power of WCDMA downlink based on a downlink pole equation," K. Spiliä, Z.-C. Honkasalo, J. Laiho-Steffens, A.Wacker, VTC 2000, pp. 1002-1005.

[Stu05] "The EUREKA GANDALF Project: Monitoring and Self-Tuning Techiques for Heterogeneous Radio Access Networks," P. Stuckmann, Z. Altman, H. Dubreil, A.

Ortega, R. Barco, M. Toril, M. Fernandez, M. Barry, S. McGrath, L. Moltsen Nielsen, *Vehicular Technology Conference Spring 2005*, Stockholm, June 2005.

[Tol02a] "Adaptative load balancing between multiple cell layers," A. Tölli, P. Hakalin, VTC 2002 Fall, Sept. 2002.

[Tol02b] "Performance Evaluation of Common Radio Resource Management," A. Tölli, P. Hakalin, and H. Holma, in Proc. of IEEE Int. Conf. on Communications, vol.5 pp. 3429-3433, 2002.

[Vee99] "The Coverage – Capacity Tradeoff in Cellular CDMA systems," V.V. Veeravalli and A. Sendonaris, IEEE Trans. on Vehicular Technology, vol. 48, no. 5, Sep. 1999.

[Vee03] "Optimization Using Particle Swarms with Near Neighbor Interactions," K. Veeramachaneni, T. Peram, C. Mohan, L.A. Osadciw, presented at Genetic and Evolutionary Computation Conference (GECCO), 2003.

[Yan01] "Competitive Takagi-Sugeno fuzzy reinforcement learning," X.W. Yan, Z.D. Deng, Z.Q. Sun, Proc. of the IEEE Int. Conf. on Control Applications (CCA '01), pp. 878 – 883, 5-7 Sept. 2001.

[Ye05] "Call Admission Control in Wideband CDMA Cellular Networks by using Fuzzy Logic," J. Ye, X. Shen, J.W. Mark, IEEE Trans. on Mobile Computing, Vol 4, No2, pp. 129- 141, March-April 2005.

[Yu04] "Efficient QoS Provisioning for Adaptative Multimedia in Mobile Communication Networks by reinforcement Learning," F. Yu, V.W.S. Wong and V.C.M. Leung, Proc. of the first Int. Conf. on Broadband Networks, BROADNETS'04 IEEE, 2004.

[Zha01a] "UMTS Air interface Voice/Data capacity – Part 2: forward link analysis," Q. Zhang, proceedings of VTC2001 spring, 2001.

[Zha01b]"UMTS air interface voice/data capacity – Part 1: Reverse link analysis," Q. Zhang, proceedings of VTC2001 spring, 2001.

[Zha04] "Handover Decision using Fuzzy MADM in Heterogeneous Networks," W. Zhang, IEEE Wireless Communications and Networking Conference (WCNC), Atlanta, March 2004.

Stages

[Dia04] "Conception d'un outil d'ingénierie dynamique pour la macro-divesité", V. Diascorn, rapport de stage, Juillet 2004.

Thèses

"Contribution a l'amélioration du trafic multimédia dans un cœur de réseau IP," M. Salaün, Université de Bretagne Occidentale et ENST Bretagne, 2004.

"Contribution to optimization of UMTS radio planning," S. Ben Jamaa, ENST Paris, 2004.

Liste de publications – Hervé Dubreil

[Alt05] Soumission d'un chapitre Wiley "Auto-tuning of RRM parameters in UMTS networks" Zwi Altman, Hervé Dubreil, Ridha Nasri, Ouassim Ben Amor, Jean-Marc Picard and Vincent Diascorn, September 2005.

[Ben05a]"Quality indicator matrices and their contribution to WCDMA network design", S. Ben Jamaa, H. Dubreil, Z. Altman, A. Ortega, *IEEE Transactions on Vehicular Technology*, vol. 54, May 2005.

[Ben05b] "Off-line Optimization Strategies of Fuzzy Logic Controllers for RRM Auto-Tuning of UMTS Networks", Ouassim Ben Amor, Zwi Altman, Hervé Dubreil, Ridha Nasri, and Maurice Clerc, Conference ICC 2006, submitted September 2005.

[Dub03] "Radio capacity estimation of a CDMA system with multi-service traffic", H. Dubreil and A. Daviaud, *WWC 2003*, San Francisco, May 2003.

[Dub05a] "Particle swarm optimization of fuzzy logic controller for high quality RRM auto-tuning of UMTS networks", H. Dubreil, Z. Altman, V. Diascorn, J.M. Picard, and M. Clerc, *Vehicular Technology Conference Spring 2005*, Stockholm, June 2005.

[Dub05b] Soumission d'un chapitre "Ingénierie Automatique des Réseaux mobiles" du traité Hermès sur l'Autonomic Networking, Hervé Dubreil, Ridha Nasri et Zwi Altman, August 2005.

[Nas05a] "Dynamic Radio Resource Management in Wireless Networks: Towards autonomic mobile networking," R. Nasri, Z. Altman and H. Dubreil, IEEE Communication Magazine, submitted April 2005.

[Nas05b] "Fuzzy-Q-Learning-Based Autonomic Management of Macro-diversity Algorithm in UMTS Networks," R. Nasri, Z. Altman and H. Dubreil, Annals of Telecommunications, Special Issue on Autonomic Communication, submitted July 2005.

[Nas05c] "WCDMA Downlink Load Sharing with Dynamic Control of Soft Handover Parameters", Ridha Nasri, Zwi Altman, Hervé Dubreil, Conference VTC 2006, submitted September 2005.

[Pic04] "Dynamic Control of UMTS Networks by Load Target Tuning", J.M. Picard, H. Dubreil, F. Garabedian, Z. Altman, *IEEE International Symposium VTC 2004*, Genoa, Italy, May. 11-14, 2004.

[Pic05] "Automatic Cell Planning Strategies for UMTS Networks", J-M. Picard, Z. Altman, S. Ben Jamaa, M. Demars, H. Dubreil, B. Fourestié and A. Ortega, *International Journal of Mobile Network Design and Innovation*, accepted.

[Stu05] "The EUREKA GANDALF Project: Monitoring and Self-Tuning Techiques for Heterogeneous Radio Access Networks", P. Stuckmann, Z. Altman, H. Dubreil, A. Ortega, R. Barco, M. Toril, M. Fernandez, M. Barry, S. McGrath, L. Moltsen Nielsen, *Vehicular Technology Conference Spring 2005*, Stockholm, June 2005.

2 Brevets

10. Annexes

10.1. Architecture des réseaux mobiles UMTS

Les réseaux mobiles UMTS sont fondés sur la technologie d'accès radio *wideband-CDMA*, qui est une technologie d'accès multiples en division par code en large bande [Hol02]. Cela signifie que les bits d'informations sont étalés en large bande (de 5 MHz en UMTS) en multipliant les données des utilisateurs par des séquences de *bits* quasi aléatoires (les *chips*) dérivées des séquences de codes CDMA. Les séquences allouées aux canaux de trafic dédiés (DCH) au sein d'une même cellule sont orthogonales entres elles sur le lien descendant.

Fonctionnellement, le réseau UMTS se compose du réseau d'accès radio ou UTRAN (*UMTS Terrestrial RAN*) et du réseau cœur qui est responsable du routage des appels et des transferts de données vers les réseaux extérieurs [Hol02]. Du point de vue des spécifications et de la normalisation, l'UTRAN se fonde sur de nouveaux protocoles développés pour la technologie W-CDMA, alors que la définition du réseau cœur est héritée du système GSM.

Comme décrit sur la figure 10.1, l'UTRAN est composé de deux éléments de réseau distincts :

- La station de base (ou *node B*) : ses principales fonctions sont d'assurer les traitements nécessaires à la couche physique radio: codage de canal, entrelacement, adaptation de débit, étalement… Cet élément remplit aussi quelques fonctions de gestion de la ressource radio (RRM) comme la boucle de contrôle de puissance interne.
- Le *Radio Network Controller* (RNC) est responsable du contrôle des ressources radio de l'UTRAN. Il s'interface avec le réseau cœur (par l'intermédiaire d'un MSC et d'un SGSN).

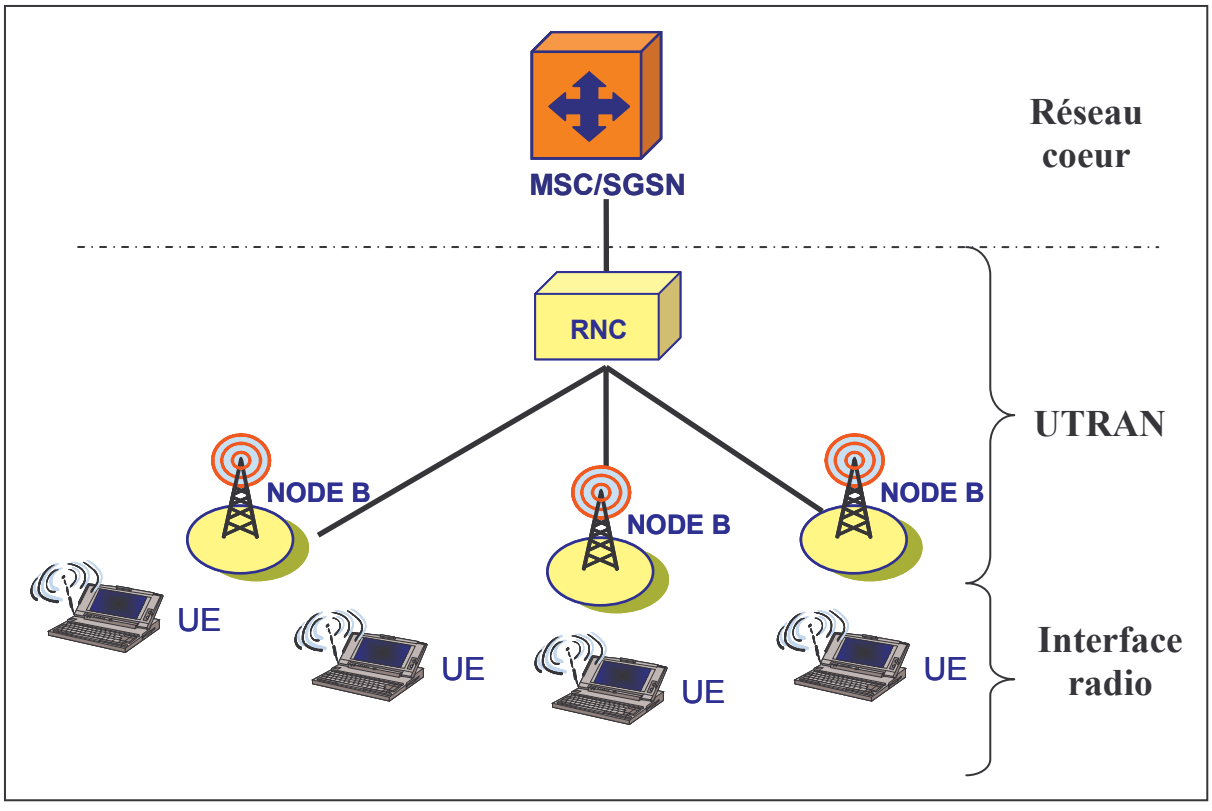


Figure 10.1 : Architecture du réseau d'accès radio UMTS.

Les MSC (*Mobiles Services Switching*) et SGSN (*Serving GPRS Support Node*) assurent respectivement le routage vers le réseau cœur des communications respectivement en mode circuit et en mode paquet.

10.2. Différentes représentations paramétriques des matrices de décisions

Nous introduisons ci-dessous deux représentations paramétriques possibles pour les matrices de décisions introduites au chapitre 3 dans les contrôleurs de logique floue :

- **Approche par combinaison linéaire de deux matrices élémentaires**

A partir de deux matrices élémentaires (mat1 et mat2) et de deux facteurs multiplicatifs (α et β), la matrice est construite par combinaison linéaire :

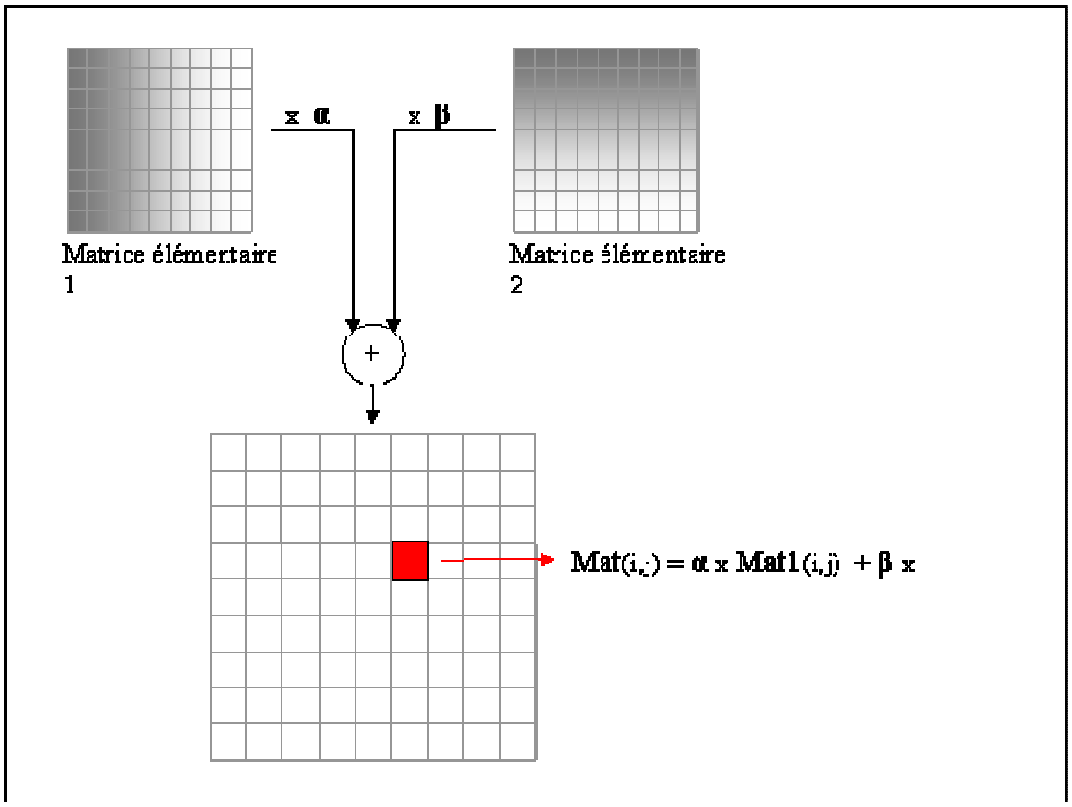


Figure 10.2 : Matrice de décisions obtenue à partir d'une combinaison linéaire de deux matrices élémentaires.

- **Approche par calcul de distance à un point**

Dans cette approche, on paramètre la distance et l'angle ainsi que les valeurs minimales et maximales de la matrice de décisions :

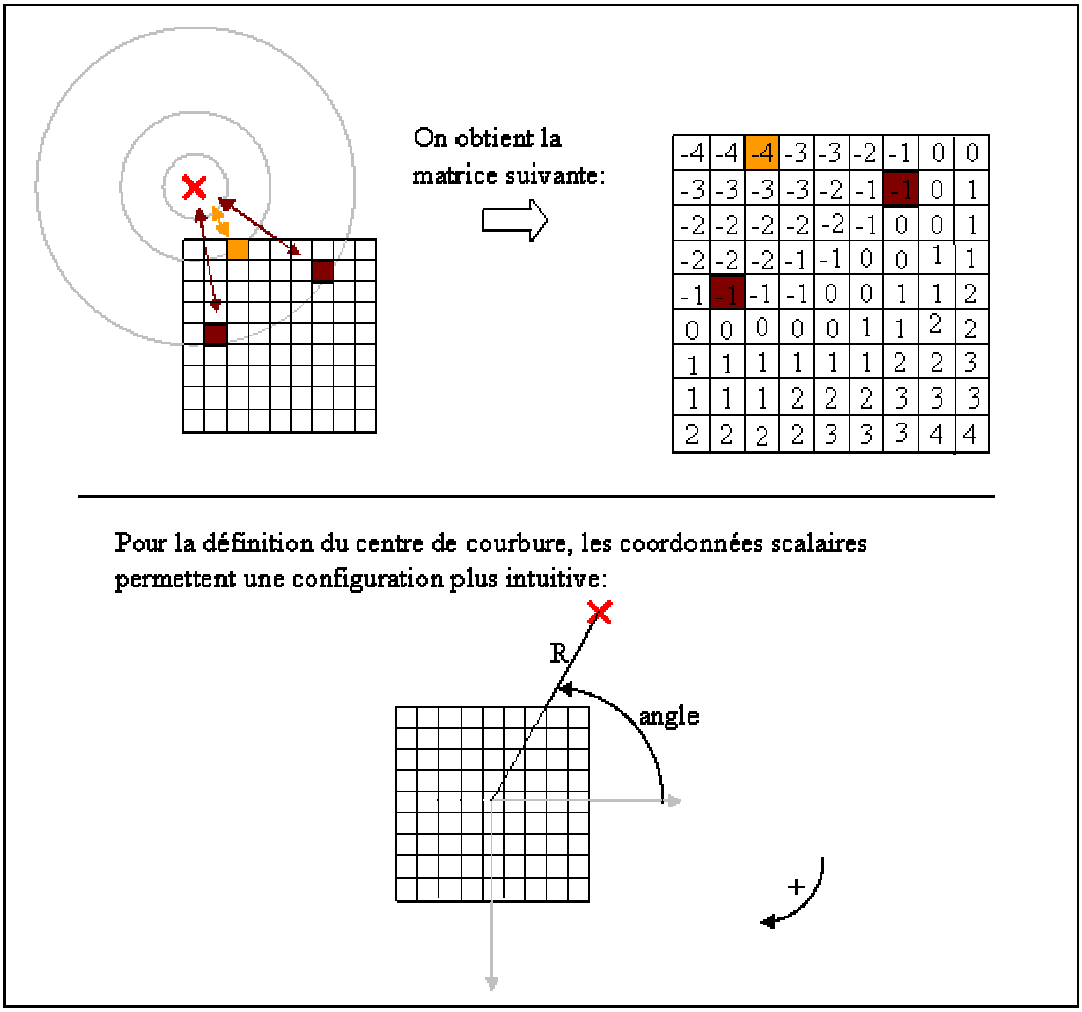


Figure 10.3 : Matrice de décisions obtenue à partir d'un calcul de distances.

11. Glossaire

3GPP	3rd Generation Partnership Project.
ACP	Automatic Cell Planner : outil de planification cellulaire.
BER	Bit Error Rate : taux d'erreur binaire.
CAC	Call Admission Control : contrôle d'admission des appels.
CAPEX	Capital Expenditures : budget d'investissement.
CDMA	Code Division Multiple Access : technologie d'accès multiples en division par code.
CIO	Cell Individual Offset : offset individuel par cellule.
CIR	Carrier to Interference Ratio : rapport porteuse sur interférence.
CLF	Contrôleur de logique floue.
CM	Contrôle du maintien.
CPICH	Common Pilot Channel : canal commun pilote.
DCH	Dedicated Channel : canal dédié.
DL	Downlink : sens descendant, c'est à dire du node B vers le terminal.
DPCH	Dedicated physical channel : canal physique dédié.
Ec/Io	Energie reçue par chip divisée par la densité de puissance dans la bande.
EDGE	Enhanced Data rates for Global Evolution.
FDD	Frequency Division Duplex : séparation des voies montantes et descendantes par transmission sur des fréquences différentes.
FKBC	Fuzzy Knowledge Based Controller : processus de contrôle basé sur la logique floue.
GPRS	General Packet Radio Service : service de transmission de données par paquets y compris sur la voie radio.
GSM	Global System for Mobile communications.
HCS	Hierarchical Cellular Systems:
IRP	Integration Reference Points.
IS	Influence set : ensemble d'influence qui rassemble dans un même goupe les stations qui interférent le plus les unes avec les autres.
KPI	Key Performance Indicator : indicateur de performances.
MD	Macro-diversité.
MDP	Markovian Decision Processus : processus de décisions Markovien.
MRC	Maximum Ratio Combining.
NMC	Network Management Center : Système d'administration générale d'un PLMN.
OAR	Optimisation par Apprentissage par Renforcement.
OEP	Optimisation par Essaim de Particules.
OMC	Operations & Management Center : centre d'administration, de surveillance et de gestion du RAN, en général associé à un sous-système particulier.
OPEX	Operating Expenditures : charges opérationnelles.
PLMN	Public Land Mobile Network.
RAN	Radio Access Network.
RNC	Radio Network Controller.
RNO	Radio Network Optimisation : module d'optimisation radio à l'OMC.
RRM	Radio Resource Management : gestion de la ressource radio.

RSCP	CPICH *Received Signal Code Power* : puissance reçue sur le code d'un canal.
QoS	Quality of Service.
SIB	System Information Broadcast.
SIR	Signal Interference Ratio : rapport signal sur interférence.
SHO	Soft Hand-Over.
SIF	Système d'Inférence Floue.
SMS	Short Message Service : service bidirectionnel de messages courts.
TCH	Traffic Channel : canal de traffic.
TDD	Time Division Duplex : séparation des voies montantes et descendantes par transmission à des instants différents sur la même fréquence.
UL	Uplink : sens montant, c'est à dire du terminal vers le node B.
UMTS	Universal Mobile Telecommunications System.
UTRAN	UMTS Terrestrial RAN.
W-CDMA	Wideband CDMA.
WiFI^{TM}	Wireless Fidelity : un réseau WiFI^{TM} est un réseau répondant à la norme 802.11.
WLAN	Wireless Local Area Network : réseau local sans fil.

RESUME en français

Les opérateurs pourraient améliorer la qualité de service et les performances des réseaux UMTS en adaptant judicieusement les paramètres de gestion de ressource radio (RRM). L'objectif de cette thèse est de développer des méthodes de paramétrage automatique des réseaux UMTS. Deux modes de fonctionnement ont été envisagés : le paramétrage *off-line* et l'adaptation *on-line*. Le premier mode permet à l'aide d'un simulateur de réseau mobile de déterminer le réglage des paramètres RRM le plus adapté pour une demande en trafic donnée. Le second mode ou ingénierie dynamique est un processus qui permettra d'ajuster dynamiquement certains paramètres contrôlant des algorithmes de RRM pour adapter le réseau mobile aux fluctuations de trafic. Le processus d'ingénierie automatique est basé sur une boucle de régulation qui est orchestrée par un contrôleur de logique floue. Ce contrôleur reçoit en entrées des indicateurs de qualité filtrés et produit en sortie un jeu de paramètres RRM modifiés. Pour améliorer les performances de ces contrôleurs, deux méthodes d'optimisation automatique ont été développées et testées numériquement à l'aide d'un simulateur de réseaux mobiles. La première est une méthode d'optimisation itérative stochastique : l'optimisation par essaim de particules. La seconde approche est fondée sur l'algorithme de *Q-learning* qui appartient à la famille des méthodes d'apprentissage par renforcement. Pour cela, il a été nécessaire de modéliser nos contrôleurs comme des systèmes d'inférence floue (SIF). Ces méthodes de contrôle ont été testées sur deux algorithmes RRM importants des réseaux UMTS : le contrôle d'admission et la macro-diversité. La méthode de contrôle automatique par SIF nous apparaît suffisamment générique pour pouvoir non seulement s'adapter aux autres algorithmes RRM, mais aussi à différents scénarii de gestion coordonnée des réseaux d'accès radio multi-systèmes.

TITRE en anglais

Optimization of fuzzy logic controller for auto-tuning of UMTS networks

---

RESUME en anglais

The operator can enhance quality of service (QoS) by judiciously tuning basic UMTS functions, which are governed by Radio Resource Management (RRM) parameters. Auto-tuning aims at automatically controlling network performance by adjusting certain RRM parameters to adapt the network to traffic fluctuations. Two control methods have been investigated and developed: the off-line optimization of control parameters and the on-line auto-tuning. The first approach aims at modifying the RRM parameters based on network simulations and traffic assumptions. The second method allows to dynamically adjust the RRM parameters of nodes B to adapt the network to traffic variations. The RRM auto-tuning controllers are based on fuzzy logic control theory. The control process utilizes quality indicators that are sampled, filtered and processed to choose new corrective parameter settings for sectors with insufficient QoS in an auto-regulation loop process. To guarantee high performance of the auto-tuning process, two methods have been adapted to perform the optimization task of Fuzzy Logic Controller: the particle swarm which is a stochastic optimization method and the Q-learning algorithm that belongs to the reinforcement learning strategies. With the second method, the controllers have been modelled as fuzzy inference systems (FIS). The optimization methods have been tested to improve performance of admission control and macro-diversity algorithms. Furthermore, the FIS auto-tuning approach appears to be sufficiently generic to be adapted to auto-tuning algorithms in a multi-system radio access network context.

DISCIPLINE

Communications & Electronique

MOTS-CLES

Auto-tuning, RRM parameters, fuzzy logic controller, particle swarm, fuzzy inference system, reinforcement learning, UMTS network

---

INTITULE ET ADRESSE DU LABORATOIRE :

Département Communications & Electronique de l'Ecole Nationale Supérieure des Télécommunications, 46 rue Barrault F-75634 Paris Cedex 13.