

Auto-tuning and Self-optimization of 3G and Beyond 3G Mobile Networks

Ridha Nasri

▶ **To cite this version:**

Ridha Nasri. Auto-tuning and Self-optimization of 3G and Beyond 3G Mobile Networks. Networking and Internet Architecture [cs.NI]. Université Pierre et Marie Curie - Paris VI, 2009. English. NNT : . tel-00494190

HAL Id: tel-00494190
https://pastel.hal.science/tel-00494190

Submitted on 22 Jun 2010

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.

THESE DE DOCTORAT DE
L'UNIVERSITE PIERRE ET MARIE CURIE

Spécialité

Informatique, Télécommunications et Électronique

Présentée par

M. Ridha Nasri

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Sujet de la thèse :

Paramétrage Dynamique et Optimisation Automatique des Réseaux Mobiles 3G et 3G+

Soutenue le 23 Janvier 2009

devant le jury composé de :

Prof. Guy Pujolle	Directeur de thèse
Dr. Zwi Altman	Encadreur
Dr. Mongi Marzoug	Rapporteur
Prof. Nazim Agoulmine	Rapporteur
Prof. Tijani Chahed	Examinateur
Dr. Salaheddine Maiza	membre Invité

À mon père Hacen,
À ma mère Ramdhana,
À tous mes frères et mes sœurs
À ma fiancée Mounira
À tous mes proches et mes amis

Remerciements

Tout d'abord, je tiens à remercier Mr. *Zwi Altman*, ingénieur de recherche et chef de projet à France Télécom R&D, pour son encadrement. Sa grande expérience, sa vision et sa connaissance approfondie m'ont acheminé toujours sur la bonne direction. J'ai apprécié les discussions techniques que nous avons eues ensemble. Il m'a permis d'élargir le spectre de mes connaissances et de mes contributions.

Mes plus vifs remerciements vont également à Mr. *Guy Pujolle*, professeur à l'université UPMC pour sa direction de ma thèse et pour ses réponses précieuses à mes demandes.

J'exprime ma gratitude à Mr. *Mongi Marzoug*, directeur du "pole modélisation des réseaux" à Orange, et Mr. *Nazim Agoulmine*, professeur à l'université d'Evry, qui consentirent à en être les rapporteurs de thèse. D'égale façon, je remercie Mr. *Tijani Chahed*, professeur à TELECOM SudParis, pour ses nombreux conseils précieux et pour sa participation en tant qu'examinateur de ma thèse. Ma précieuse gratitude va également à Mr. *Salaheddine Maiza*, Ingénieur expert en qualité réseau à SFR, d'avoir accepté mon invitation pour participer à ma soutenance.

La thèse a été effectuée au sein de l'unité de recherche REM du laboratoire NET du FTR&D. Mes gratitudes vont a toutes les personnes de l'unité qui par leurs idées et soutiens m'ont aidé à bien finir ce travail, en particulier *Hervé Dubreil*, *Abed Samhat*, *Arturo Ortega Molina*, *Zakaria Nouir* et *Salah-Eddine Elayoubi*.

Une grande partie de la thèse était faite dans le cadre d'un projet européen *Eureka-Celtic Gandalf*. Je remercie vivement tous les partenaires du projet pour leurs idées et expertises enrichissantes.

Enfin, ma reconnaissance et mon affection vont à mes chers parents qui m'ont soutenu tout au long de mes études.

Résumé

La télécommunication radio mobile connait actuellement une évolution importante en termes de diversité de technologies et de services fournis à l'utilisateur final. Il apparaît que cette diversité complexifie les réseaux cellulaires et les opérations d'optimisation manuelle du paramétrage deviennent de plus en plus compliquées et coûteuses. Par conséquent, les coûts d'exploitation du réseau 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 réseaux mobiles déployés, il sera possible de retarder les opérations de densification du réseau et l'acquisition de nouveaux sites. Le paramétrage automatique et optimal permettra donc aussi d'étaler voire même de réduire les investissements et les coûts de maintenance du réseau.

Cette thèse introduit de nouvelles méthodes de paramétrage automatique (*auto-tuning*) des algorithmes RRM (*Radio Resource Management*) dans les réseaux mobiles 3G et au delà du 3G. L'*auto-tuning* est un processus utilisant des outils de contrôle comme les contrôleurs de logique floue et d'apprentissage par renforcement. Il ajuste les paramètres des algorithmes RRM afin d'adapter le réseau aux fluctuations du trafic. Le fonctionnement de l'*auto-tuning* est basé sur une boucle de régulation optimale pilotée par un contrôleur qui est alimenté par les indicateurs de qualité du réseau. Afin de trouver le paramétrage optimal du réseau, le contrôleur maximise une fonction d'utilité, appelée aussi fonction de renforcement.

Quatre cas d'études sont décrits dans cette thèse. Dans un premier temps, l'*auto-tuning* de l'algorithme d'allocation des ressources radio est présenté. Afin de privilégier les utilisateurs du service temps réel (voix), une bande de garde est réservée pour eux. Cependant dans le cas où le trafic temps réel est faible, il est important d'exploiter cette ressource pour d'autres services. L'*auto-tuning* permet donc de faire un compromis optimal de la qualité perçue dans chaque service en adaptant les ressources réservées en fonction du trafic de chaque classe du service. Le second cas est l'optimisation automatique et dynamique des paramètres de l'algorithme du soft *handover* en UMTS. Pour l'*auto-tuning* du *soft handover*, un contrôleur est implémenté logiquement au niveau du RNC et règle automatiquement les seuils de *handover* en fonction de la charge radio de chaque cellule ainsi que de ses voisines. Cette approche permet d'équilibrer la charge radio entre les cellules et ainsi augmenter implicitement la capacité du réseau. Les simulations montrent que l'adaptation des seuils du *soft handover* en UMTS augmente la capacité de 30% par rapport au paramétrage fixe.

L'approche de l'*auto-tuning* de la mobilité en UMTS est étendue pour les systèmes LTE (3GPP *Long Term Evolution*) mais dans ce cas l'*auto-tuning* est fondé sur une fonction d'*auto-tuning* préconstruite. L'adaptation des marges de *handover* en LTE permet de lisser les interférences intercellulaires et ainsi augmenter le débit perçu pour chaque utilisateur du réseau.

Finalement, un algorithme de mobilité adaptative entre les deux technologies UMTS et WLAN est proposé. L'algorithme est orchestré par deux seuils, le premier est responsable du *handover* de l'UMTS vers le WLAN et l'autre du *handover* dans le sens inverse. L'adaptation de ces deux seuils permet une exploitation optimale et conjointe des ressources disponibles dans les deux

technologies. Les résultats de simulation d'un réseau multi-systèmes exposent également un gain important en capacité.

Abstract

With the wireless mobile communication boom, auto-tuning and self-optimization of network parameters are more than ever key issues to provide high-quality services for the end-user and to decrease the operational expenditure of the network operation. The special attention drawn to the self-optimization of radio resource management (RRM) parameters is motivated by the user need for ubiquitous communication and by the increasing complexity of networks resulting from the cooperation of radio access technologies. Formerly, RRM has been based on some algorithms (admission control, resource allocation, handover …) governed by a set of fixed thresholds. Today, RRM procedures have undergone a considerable change and the paradigm will shift towards a completely automatic network management. Optimal auto-tuning mechanisms, performed by control methods such as fuzzy logic optimized by reinforcement learning, could considerably improve network management functions with respect to traditional RRM algorithms with fixed parameters.

This thesis introduces new results in auto-tuning and self-optimization of RRM parameters in 3G and beyond 3G networks. Auto-tuning tasks are organized in the control plane where different information exchange is involved between the network nodes. Auto-tuning using fuzzy logic control is performed in a local loop, namely the controller is in continuous interaction with the network. The controller feeds the network with new parameter settings and conversely the network returns its feedback by delivering new quality indicators indicating its operating state.

Different use-cases are investigated. First, an auto-tuning of resource allocation algorithm in UMTS is studied as an alternative to the existing static resource allocation. The auto-tuning process dynamically adapts a guard band that is reserved for users using real time services. A best trade-off between real time and non real time services is achieved in the sense that the quality of service becomes comparable in the two traffic classes especially in a high load situation. The second use-case concerns the self-optimization of soft handover parameters in UMTS networks. For each cell the controller receives as inputs the filtered downlink load and that of its neighbouring cells. The controller continually learns the best parameter values in each network situation. The learning process is governed by a utility function. Simulation results reveal significant improvements in terms of network performance. The proposed auto-tuning algorithm balances the radio load between base stations and improves the system capacity by up to 30% compared to a UMTS network with fixed soft handover parameters. However, the auto-tuning increases the signalling messages load in the radio interface as well as in the core network. This negative effect is minimized by reducing the reactivity of the auto-tuning controller.

The third case deals with the auto-tuning of LTE (3GPP Long Term Evolution) mobility algorithm. The auto-tuning is carried out by adapting handover margin involving each couple of cells according to the difference between their loads. The auto-tuning alleviates cell congestion and balances the traffic and the load between cells by handing off mobiles close to the cell border from the congested cell to its neighbouring cells. Simulation results, based on simplified system and interference models, show that the auto-tuning process brings about an important gain in both call admission rate and user throughput.

Finally, an algorithm of intersystem mobility between UMTS and WLAN access technologies is proposed. The algorithm is coverage and load based, and is governed by two thresholds: the first is responsible for handover from UMTS to WLAN and the second - for the inverse direction. The self-optimization of the two thresholds is jointly performed. The results obtained using the adaptive intersystem mobility algorithm show high tracking capacity gain, and illustrate the importance of intelligent cooperation between technologies.

The results provided in this thesis are supported by theoretical analysis and extensive dynamic system level simulations with multi-cells' scenarios, including the effect of many relevant mechanisms that have an impact on the radio access. However, simulations do not exactly reflect the reality in the network operation. So auto-tuning should be tested in a real experimental network (test-bed).

Table of contents

Résumé i
Abstract iii
Table of contents v
List of figures and tables vii
Acronyms ix

Chap. 1 Introduction 1
1.1 Background and problem definition 1
1.2 Scope and objectives of the thesis 2
1.3 Original contributions 3
1.4 Thesis structure 4

Chap. 2 Auto-tuning in mobile communication: Related works 6
2.1 Introduction 6
2.2 Auto-tuning of GSM main functionalities 6
2.2.1 Overview of GSM Networks and their evolutions "GPRS and EDGE" 6
2.2.2 RRM algorithms and target auto-tuned parameters 7
2.2.3 Literature review of GSM auto-tuning 11
2.3 UMTS networks 13
2.3.1 Overview of UMTS Network and its evolutions 13
2.3.2 UMTS RRM algorithms and related parameters 15
2.3.3 Related work on the auto-tuning of UMTS parameters 20
2.4 Multi-system networks 22
2.4.1 GERAN UTRAN inter-working 22
2.4.2 3GPP-WLAN inter-working 24
2.4.3 Literature review of auto-tuning in multi-system networks 25
2.5 Conclusion 27

Chap. 3 Auto-tuning architecture and tools 28
3.1 Introduction 28
3.2 Auto-tuning architecture 28
3.2.1 Gandalf management architecture 28
3.2.2 Gandalf auto-tuning architecture 29
3.2.3 Auto-tuning information flows 31
3.2.4 Auto-tuning and optimization engine 32
3.3 Fuzzy logic controller 33
3.3.1 Mathematical framework of FLC 34
3.3.2 Example: use-case of FLC 36
3.4 Reinforcement learning 37
3.4.1 General view of machine learning 37
3.4.2 Mathematical framework of RL 38
3.5 Fuzzy Q-learning controller 43
3.5.1 Q-learning algorithm 43
3.5.2 Adaptation of Q-learning to fuzzy inference system 44
3.6 Conclusions 46

Chap. 4 Application of auto-tuning to the UMTS networks 47

4.1 Introduction 47
4.2 Correlation between quality indicators 47
4.2.1 Presentation of used quality indicators 47
4.2.2 Correlation between quality indicators 49
4.3 Auto-tuning of resource allocation in UMTS 50
4.3.1 Admission control strategy 51
4.3.2 Quality indicators, actions and reinforcement function 53
4.3.3 Performance evaluation 54
4.4 Auto-tuning of UMTS soft handover parameters 56
4.4.1 SHO algorithm 57
4.4.2 FQLC-based auto-tuning of SHO parameters 59
4.4.3 Performance evaluation 60
4.4.4 Signalling overload due to auto-tuning 64
4.5 Conclusions 66

Chap. 5 Self optimization of mobility algorithm in LTE networks 67
5.1 Introduction 67
5.2 Overview of LTE system 67
5.2.1 System requirements 68
5.2.2 System architecture 68
5.2.3 Physical layer 70
5.2.4 Self optimizing network functionalities 70
5.3 Interference in e-UTRAN system 71
5.3.1 System model and assumptions 71
5.3.2 Interference model 71
5.4 Auto-tuning of e-UTRAN handover algorithm 74
5.4.1 E-UTRAN handover algorithm 75
5.4.2 Handover adaptation and load balancing 76
5.4.3 Auto-tuning of handover margin 77
5.5 Simulations and results 78
5.6 Conclusion 82

Chap. 6 UMTS-WLAN load balancing by auto-tuning inter-system mobility 83
6.1 Introduction 83
6.2 Assumptions 83
6.2.1 UMTS-WLAN inter-working mode 83
6.2.2 Technology selection and admission control 84
6.3 UMTS-WLAN vertical handover and its auto-tuning 86
6.3.1 Vertical handover description 86
6.3.2 Auto-tuning of vertical handover parameters 87
6.4 Simulations and performance evaluations 88
6.5 Conclusion 91

Chap. 7 Conclusions and perspectives 92
7.1 Conclusions 92
7.2 Limitations and perspectives 93

References 94
Appendix A: Convergence proofs of the reinforcement learning algorithm 100
Appendix B: LTE interference model 104
Appendix C: Network system level simulator 107

List of figures and tables

Figure 1.1. Fully heterogeneous access network. 1
Figure 1.2. Targets of the thesis. 3
Figure 2.1. GSM system architecture. 7
Figure 2.2. Dynamic resource allocations between CS and PS services. 9
Figure 2.3. UMTS architecture. 14
Figure 2.4. Intra-Frequency Soft Handover 18
Figure 2.5. Centralized CRRM entity 23
Figure 2.6. Decentralized CRRM into every RNC/BSC . 23
Figure 3.1. Network management tasks with the corresponding time scale 29
Figure 3.2. Auto-tuning architecture in user, control and management planes. 30
Figure 3.3. Signalling messages between AOE and JRRM/RRM modules. 31
Figure 3.4. Auto-tuning and Optimization Engine 33
Figure 3.5. The concept of fuzzy logic controller. 34
Figure 3.6. Fuzzy logic controller based on Takagi-Sugeno approach. 35
Figure 3.7. Fuzzy sets and menbership function of the dropping and blocking rate. 36
Figure 3.8. Example of Markovian Decision Process. 39
Figure 3.9. Fuzzy Q-learning algorithm. 45
Figure 4.1. Capacity model for a UMTS base station . 51
Figure 4.2. Call success rate of each service as a function of the RT guard band for two traffic situations: (1) RT=3 and NRT=5; and (2) RT=5 and NRT=2. 52
Figure 4.3. Fuzzy Q-learning controller for auto-tuning the RT guard band in each BS. 53
Figure 4.4. Call succes rate as a function of RT call arrival rate (RT_D means RT CSR in the dynamic version, RT_F25% means RT CSR for the fixed guard band of 25%). 55
Figure 4.5. Histograms of RT CSR for all BSs with traffic arrival rate of RT=2 and NRT=6 mobiles/s. 56
Figure 4.6. Histograms of NRT CSR for all BSs with traffic arrival rate RT=2 and NRT=6 mobiles/s. 56
Figure 4.7. UMTS SHO algorithm (event 1A, 1B and 1C). 58
Figure 4.8. Evolution of the convergence criteria of the fuzzy-Q-learning controller. 61
Figure 4.9. Evolution of the quality of rule-action pairs. 62
Figure 4.10. Call succes rate versus incomming tarffic for the optimized network with autonomic management compared to a classical network. 62
Figure 4.11. Cumulative distribution function of the cell call succes rate in the optimized network compared to the classical network. 63
Figure 4.12. Distribution of cell load for the optimized network compared to a classic network with fixed configuration. 63
Figure 4.13. Percentage of mobiles in SHO situation as a function of arrival rate for the network without and with auto-tuning. 64
Figure 4.14. Cumulative distribution function of the active set update frequency. 65
Figure 5.1. LTE architecture. 69
Figure 5.2. E-UTRAN (eNB) and EPC (MME and S-GW). 69
Figure 5.3. Inter-cell interference coordination scheme. 72
Figure 5.4. Typical pattern of geographical distribution of HO procedure. 77
Figure 5.5. Example of geographical distribution of HO procedure with traffic balancing. 77
Figure 5.6. The network layout including coverage of each eNB. 79
Figure 5.7. Admission probability as a function of the traffic intensity for auto-tuned handover compared with fixed handover margin network (6dB). 79

Figure 5.8. Connection holding probability as a function of the traffic intensity for auto-tuned handover compared with fixed handover margin network (6dB). 80
Figure 5.9. Average throughput per user versus the traffic intensity for auto-tuned handover compared with fixed handover margin network (6dB). 81
Figure 5.10. Cumulative distribution function of the SINR for network with and without auto-tuning, for traffic intensity equals to 8 mobiles/s. 81
Figure 6.1. Very tightly coupled UMTS/WLAN network. 84
Figure 6.2. Selection procedures in very tightly coupled UMTS/WLAN network 85
Figure 6.3. Load based VHO algorithm between UMTS and WLAN networks. 86
Figure 6.4. Heterogeneous network layout with 59 UMTS cells (squares with arrows) and 22 WLAN APs (circles). 88
Figure 6.5. Call success rate of RT traffic as a function of NRT traffic arrival rate for the network with (squares) and without (diamonds) auto-tuning. 89
Figure 6.6. Call success rate for NRT traffic as a function of NRT traffic arrival rate for the network with (squares) and without (diamonds) auto-tuning. 90
Figure 6.7. Average throughput as a function of NRT traffic intensity. 90
Figure 6.8. Impact of auto-tuning on the execution rate of UMTS to WLAN vertical handovers. 91
Figure 6.9. Impact of auto-tuning on the success rate of UMTS to WLAN vertical handovers. 91
Figure C.1. Main blocs of the multi-system simulator architecture 107
Figure C.2. Time evolution of the multi-system simulator 108
Table 3.1. Fuzzy rules. 36
Table 4.1. Correlation between quality indicators. 50

Acronyms

2G	Second Generation
3G	Third Generation
3GPP	Third-Generation Partnership Project
AAA	Authentication, Authorization and Accounting
AOE	Auto-tuning and Optimization Engine
AP	Access Point
ARRM	Advanced Radio Resource Management
ASU	Active Set Update
B3G	Beyond 3G
BS	Base Station
BSS	Base Station Subsystem
BSC	Base Station Controller
BCCH	Broadcast Control Channel
CAC	Call Admission Control
CAPEX	CApital Expenditure
CBR	Call Blocking Rate
CDMA	Code Division Multiple Access
CDR	Call Dropping Ratio
CN	Core Network
CPICH	Common Pilot Channel
CRMS	Common Resource Management Server
CRRM	Common Radio Resource Management
CS	Circuit Switched
CSR	Call Success Rate
CSSR	Call Setup Success Ratio
DCCH	Dedicated Control CHannel
DCH	Dedicated Channel
DTCH	Dedicated Traffic CHannel
DPO	Dynamic Programming Operator
ECSD	Enhanced Circuit-Switched Data
EDGE	Enhanced Data rates for GSM Evolution
EGPRS	Enhanced General Packet Radio Service
EIR	Equipment Identity Register
eNB	Evolved Node B
EPC	Evolved Packet Core
e-UTRAN	evolved UMTS Terrestrial Radio Access Network
ETSI	European Telecommunications Standards Institute
FDD	Frequency Division Duplex
FQLC	Fuzzy Q-Learning Controller
FLC	Fuzzy Logic Controller
HARQ	Hybrid Automatic Repeat Request
HCS	Hierarchical Cell Structures

HLR	Home Location Register
HO	HandOver
HSCSD	High-Speed Circuit-Switched Data
HSDPA	High Speed Downlink Packet Access
HSPA	High Speed Packet Access
HSUPA	High Speed Uplink Packed Access
ITU-T	International Telecommunication Union –Telecommunication section
IEEE	Institute of Electrical and Electronic Engineers
IETF	Internet Engineering Task Force
JRRM	Joint Common Radio Resource Management
GSM	Global System for Mobile Communication
GPRS	General Packet Radio Service
GERAN	GSM/EDGE Radio Access Network
GGSN	Gateway GPRS Support Node
KPI	Key Performance Indicator
LTE	3GPP Long Term Evolution
MAC	Medium Access Control
MBMS	Multimedia Broadcast Multicast Service
MBSFN	Multicast-Broadcast Single-Frequency Network
MDP	Markovian Decision Process
MPM	Management and Processing Module
MS	Mobile Station
MSC	Mobile Switching Center
NMS	Network Management System
Node B	UMTS base station
NRT	Non Real Time service
NSS	Network SubSystem
OAM	Operations Administrations and Maintenance
OFDM	Orthogonal Frequency Division Multiplexing
OMC	Operation and Maintenance Center
OMS	Operation and Maintenance Subsystem
OPEX	OPerational EXpenditure
PLMN	Public Land Mobile Network
PS	Packet Switched
PSK	Phase-Shift Keying modulation
QoS	Quality of Service
RAN	Radio Access Network
RAT	Radio Access Technology
RL	Reinforcement Learning
RLC	Radio Link Control
RNC	Radio Network Controller
RNS	Radio Network Subsystem
RRM	Radio Resource Management
RSCP	Received Signal Code Power
RSSI	Received Signal Strength Indicator
RT	Real Time service

SAE	System Architecture Evolution
SDS	Semi Dynamic Simulator
SGSN	Serving GPRS Support Node
SHO	Soft HandOver
SIM	Subscriber Identity Module
SINR	Signal-to-Interference plus Noise Ratio
SMDP	Semi-Markov Decision Process
SNR	Signal-to-Noise Ratio
SON	Self Optimizing Network
SSID	Service Set ID
SSSR	Session Setup Success Ratio
TDD	Time Division Duplex
TS	Time Slot
UE	User Equipment
UMTS	Universal Mobile Telecommunication System
UTRAN	UMTS Terrestrial Radio Access Networ
VHO	Vertical HandOver
VLR	Visitor Location Register
WAG	WLAN Access Gateway
WCDMA	Wide-band Code Division Multiple Access
WLAN	Wireless Local Area Network

Chap. 1 Introduction

1.1 Background and problem definition

During the last few years, Wireless multimedia networks have known an explosion and an evolution in terms of sophisticated technologies and offered services. This is due to the exponential traffic increase related to the massive demand of diverse services. For the remedy of these intensive demand issues, mobile telecommunication actors (operators, constructors, researchers) have extended their actual networks or/and migrated to other technologies. This has resulted in evolution towards a heterogeneous wireless access network compromising a set of diverse radio technologies, but offering a single set of integrated services to the end user. The existence of new and varied radio technologies is eventually leading to greater choice and better availability of radio capacity, and ultimately services to the end user. Network core and services are evolving in parallel to the radio access mechanisms, ideally resulting in an integrated service environment offering a range of mobile, secure, quality assured services in a managed fashion over a diverse set of radio access technologies. Figure 1.1 depicts an extensive, if not intimidating completely heterogeneous environment integrating existing and envisaged network types and technologies.

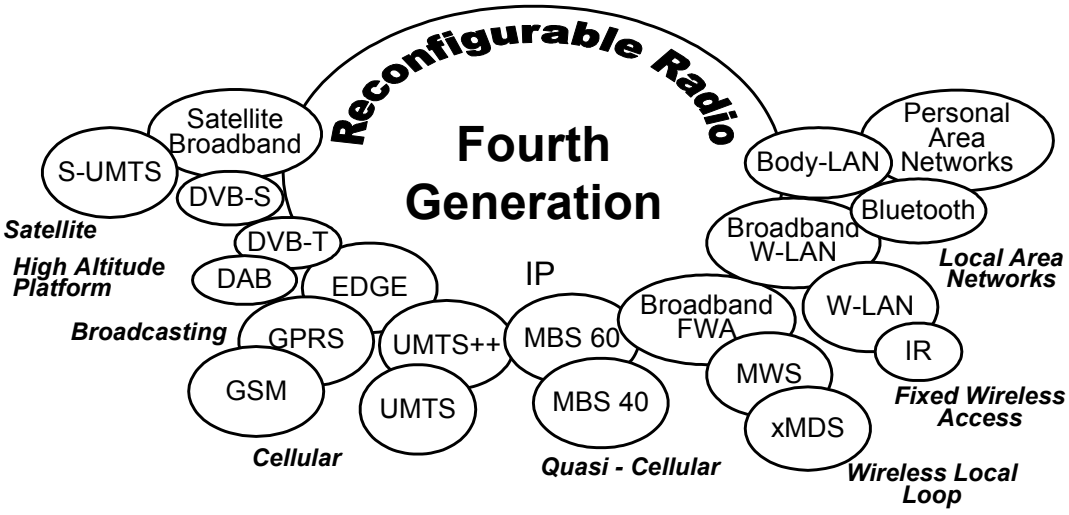


Figure 1.1. Fully heterogeneous access network.

This heterogeneous mix poses significant challenges, initially at the level of the physical interoperability/compatibility between systems and subsequently when attempting to run consistent software services over what are fundamentally different telecommunication technologies. A typical user accessing a service on a hand-held device may find their physical connection to that service switching from a WLAN link in the office, to a GPRS connection in a

car and to a 3G connection in a congested city centre. Throughout all these connections, the network infrastructure must keep the user connected, switching access technologies and maintaining a consistent user experience and service.

In addition, the fast evolution of mobile networks leads to a highly complex and heterogeneous radio access network landscape. In this context, network management becomes crucial to guarantee high quality and optimum cooperation between network subsystems. Network management and optimization tasks are today primarily manual processes. Staff carries out a series of checks and diagnosis of quality indicators to establish the causes of the problem; then analyses possible solutions and finally launches the best healing action. In this process, several applications and databases must be queried in order to analyse performance data and update the configuration of the network. Thus, operators have tried to cope with the increase of network complexity by increasing their staff and over-dimensioning resources. However, the growing size of cellular networks, together with the increasing complexity of network elements, makes this strategy no longer practical.

Likewise, due to effort and expenses, default values for parameters of RRM algorithms are set all over the network, even if non-optimum performance is achieved. Therefore, the flexibility from the large parameter set defined on a cell (or even adjacency) basis is not fully seized. These inaccurate network settings limit the network capacity, leading to premature increase of capital expenditure (CAPEX) and avoidable reductions in operator revenues. Consequently, operators currently demand automatic tools that simplify planning, rollout and operation of their networks. In addition, these automated processes will mechanize their current repetitive procedures, and also provide new optimization procedures that increase network performance at a minimum cost. In this context, auto-tuning and self optimizing network tasks are nowadays more than ever key issues to replay effectively to the need of mobile network operators and to provide high-quality service for end-users. The migration from manual to automatic cell planning has already induced a significant quality enhancement and deployment cost reduction. Similar gains are expected from the automatic and dynamic optimization of all management tasks for the operation of a mobile radio network. This expectation is especially justified in complex systems like UMTS and its long term evolution or even heterogeneous access networks.

1.2 Scope and objectives of the thesis

The main objective of this thesis is the development of dynamic and automatic optimization's algorithms to select the most appropriate values for RRM network parameters. The automatic and dynamic optimization refers in this work to auto-tuning or self-optimization techniques.

This thesis proposes new radio resource management algorithms together with methods of auto-tuning to optimise the overall systems' performance in a 3G or in the heterogeneous network. The first target is to give a simplified architecture for the implementation of the auto-tuning module in current 3G networks. The second target of the thesis is to apply the auto-tuning and the self-optimization concept to RRM algorithms in general and mobility management in particular. The auto-tuning of mobility algorithm is studied in UMTS network extended to its long term evolution and tackled again in the heterogeneous networks where only UMTS cells

and WLAN hotspots coexist. The thesis demonstrates through network simulations the feasibility of the auto-tuning concept. The targets of the thesis are presented in Figure 1.2.

The objective of the application of the auto-tuning to different RRM algorithms (left hand parts of figure 1.2) is to evaluate the gain in terms of capacity and converge that can be achieved in the network compared to the case without any auto-tuning. In the application parts, the auto-tuning concept is applied in a feedback based regulation loop. The self-tuning module calculates optimal values for the RRM parameters by dynamically processing the quality indicators delivered by the network. The control of parameters is performed based on a utility function called in the thesis reinforcement function. This function guides the auto-tuning algorithm to find the best parameter settings of the network.

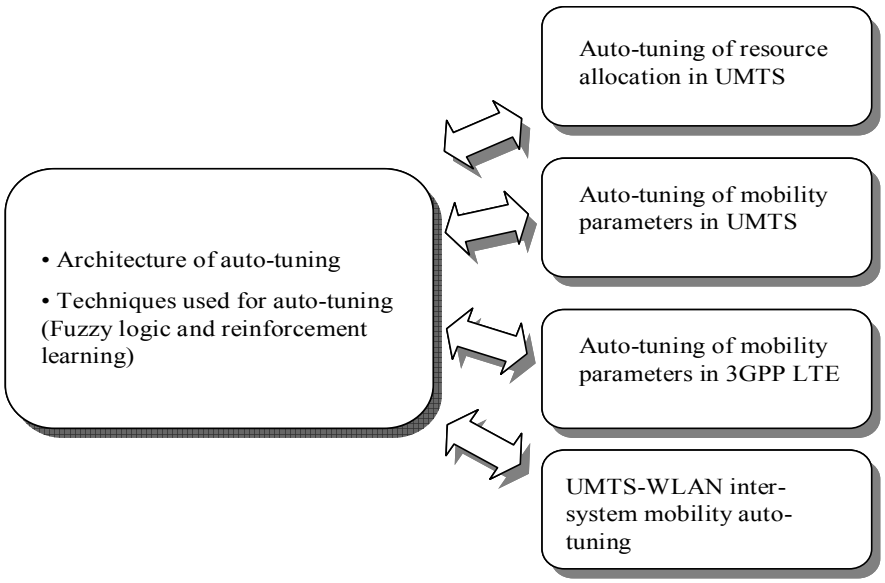


Figure 1.2. Targets of the thesis.

1.3 Original contributions

This thesis includes three major contributions: presenting a simplified architecture of auto-tuning, applying the auto-tuning to dynamic resource allocation and evaluating performances of self optimizing mobility parameters in 3 scenarios: UMTS soft handover, 3GPP LTE hard handover and intersystem mobility.

For the first contribution, major difference of the work presented here and earlier published results related to the architecture and to the way of presenting tools used for auto-tuning. In this work, an extensive explanation of fuzzy-Q-learning based automatic optimization is presented with proofs. The auto-tuning architecture is published in [1] and in [5].

With respect to the first use-case of auto-tuning, a dynamic and autonomic approach to sharing radio resources between RT and NRT services in UMTS network is studied in [10]. In this contribution, a Fuzzy Q-Learning Controller (FQLC) is used to adapt the resource reserved for

the RT traffic according to perceived quality indicators of each service. The FQLC combines both fuzzy logic theory and reinforcement learning method. Each base station is fitted out with a controller which manages its resources. Learning results from each base station are supplied to a central FQLC of the network.

In this thesis, the auto-tuning of mobility parameters is further studied and analysed in 3 different technologies. For the UMTS, we have addressed in [4] and [11] the auto-tuning of soft handover (SHO) parameters. Unlike earlier studies, the auto-tuning process uses a fuzzy Q-learning controller to dynamically adapt SHO parameters to varying network situations such as traffic fluctuation and load differences between network cells. The proposed method improves the system capacity compared to a classical network with fixed parameters, balances the load between base stations and minimizes human intervention in the UMTS mobility optimization tasks. The auto-tuning of SHO parameters and dynamic resource allocation is published later in a book chapter [6].

The mobility auto-tuning is extended to the 3G Long Term Evolution (LTE), see [7]. Due to the fact that the LTE system is a new standard, modelling interference and capacity and developing a system level simulator shall be required before studying the mobility auto-tuning. This contribution is accepted as a technical report and is extended later to a technical specification in 3GPP standard [3].

The paper [9] introduces a new WLAN-UMTS intersystem mobility algorithm which includes both coverage-based and load-based handover. The auto-tuning of the parameters governing the proposed handover is presented in [8].

The recent approaches, defining the system state in terms of load and congestion, are in their limits especially in UMTS system where different metrics are correlated such as capacity and coverage. In [2], we have proposed a new approach to determine the cell state in terms of load and congestion by jointly combining different quality indicators. This approach serves as an input for new admission and mobility management algorithms.

1.4 Thesis structure

First, the thesis presents an extensive state of the art of auto-tuning in mobile communication. The survey of the auto-tuning is preceded by a brief presentation of the related technology. The auto-tuning architecture and the presentation of fuzzy reinforcement learning algorithms are given in chapter 3. The fuzzy reinforcement learning serves as the tool used to apply auto-tuning in mobile communications. In chapter 3, we highlight the version fuzzy-Q-learning algorithm because of its use in the whole thesis except chapter 5.

Chapter 4 presents two use-cases of auto-tuning in UMTS system, the first is the dynamic resource allocation and the second is the auto-tuning of soft handover parameters. Performance evaluation of each use-case is presented based on a system level dynamic simulator.

Chapter 5 presents an in–depth interference analysis of a 3GPP LTE system. The interference analysis and system modelling serve as preliminaries to study the impact of LTE mobility auto-

tuning on the system performances defined as the average user throughput and congestion indicators of the network.

Chapter 6 is devoted to investigating the auto-tuning of UMTS-WLAN intersystem mobility. The auto-tuning is carried out assuming a very tight coupling scenario where a WLAN access point is considered as a part of a UMTS cell. The results reveal that the capacity gain of using the auto-tuning concept is further greater than the case without any auto-tuning.

The last chapter summarises and concludes the work presented, highlights the limitations and points towards potential future works.

Chap. 2 Auto-tuning in mobile communication: Related works

2.1 Introduction

Auto-tuning has been studied recently in the context of 3G networks and multi-systems environment. The main focus has been given to tune some Radio Resource Management (RRM) algorithms (such as mobility, admission control and resource allocation). The adaptation of some RRM algorithm has already been investigated for the 2nd generation networks. As the complexity of wireless communication networks increases from day to day, the need for auto-tuning becomes critical. Research activity on this topic has been conducted in both industry and academia and has been reported in the literature.

The objective of this chapter is to give a comprehensive survey on the research done in the area of auto-tuning in wireless communications. Since the work in this thesis covers the auto-tuning of different technologies, the survey is given for three technologies: GSM systems, 3G and beyond 3G systems which include the case of cooperative networks. To facilitate the reading of the dissertation, an overview of each technology is given with a special focus on RRM algorithms and related parameters that can be candidate for the auto-tuning process.

The structure of this chapter is as follows: Section 1 presents an overview of the GSM technology and related works on the auto-tuning of its RRM algorithms. Section 2 describes the UMTS system and related auto-tuning works. Section 3 treats different approaches of autonomic interconnection and adaptation of multi-system networks. As the description of these systems is involved, we limit the overview to topics directly related to auto-tuning.

2.2 Auto-tuning of GSM main functionalities

2.2.1 Overview of GSM Networks and their evolutions "GPRS and EDGE"

GSM became popular very quickly because it provided improved speech quality and, through a uniform international standard, made it possible to use a single telephone number and mobile unit around the world. The European Telecommunications Standardization Institute (ETSI) adopted the GSM standard in 1991, and GSM is now used in 135 countries. The name GERAN is used by 3GPP (3rd Generation Partnership Project) to refer to GSM radio access technology. The best way to create this successful communication system is to divide it into various subgroups that are interconnected using standardized interfaces [16]. A GSM network can be divided into three groups (see Figure. 2.1): The mobile station (MS), the base station subsystem (BSS) and the network subsystem (NSS).

At the beginning, GSM system supports only voice service. Although the first evolution of the standard, namely high-speed circuit-switched data (HSCSD) [14], enables mobile phones to support data rates up to 38.4 kbps, compared with 9.6 kbps for regular GSM networks.

Transmission speeds of up to 171.2 kbps are available with mobile phones that support the GSM standard for General Packet Radio Service (GPRS) [15]. The high bandwidth is achieved by using eight timeslots, or voice channels, simultaneously to the packet switching service.

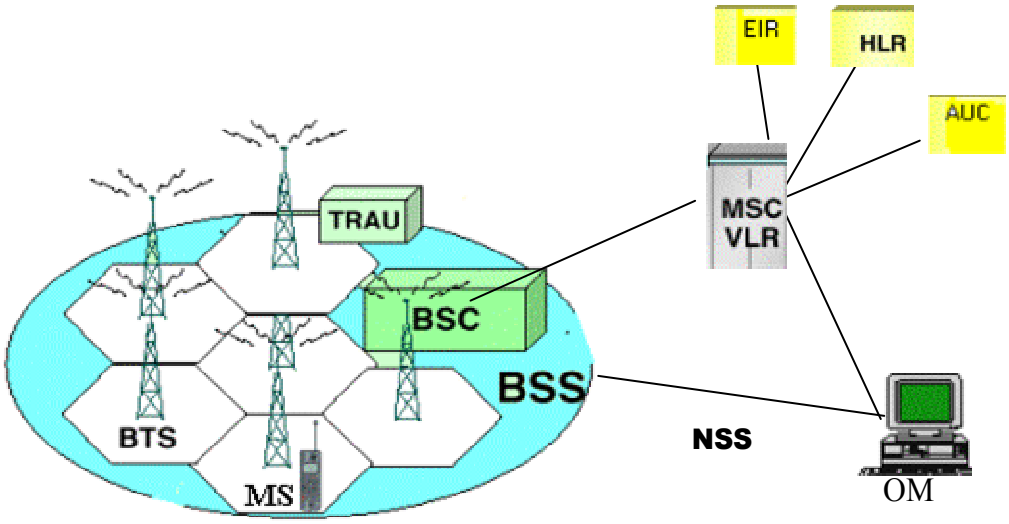


Figure 2.1. GSM system architecture.

From an operator's point of view, the introduction of GPRS facilitates the arrival of several new mobile data applications and offers advantages of managing radio resources. With GPRS, it is not only possible to use network resources in a more efficient way by treating application data flows regarding their actual needs, but also to differentiate among service users regarding their subscribed quality of service (QoS).

The evolution of GSM towards UMTS entails the exploitation of GPRS enhanced version, called EDGE (Enhanced Data Rates for GSM Evolution) [16]. Applying new modulations, new coding schemes and optimized link quality control algorithms, EDGE allows to reach higher throughputs than GPRS of up to 59.2 kbps per GSM physical channel. The additional use of 8-PSK modulation enables the introduction of new modulation and coding schemes. The advantage of the new modulation scheme is the support for higher data rates under good channel conditions, and to reuse at the same time the channel structure of the GPRS system.

The success of GSM networks is driven by RRM algorithms. RRM in such networks includes admission control and resource allocation, congestion control, packet scheduling, mobility control and cell reselection.

2.2.2 RRM algorithms and target auto-tuned parameters

In this subsection, we are going to give an overview of RRM algorithms in GSM networks. We don't hope nor do we attempt to cover all algorithms, but only a broad overview of admission,

congestion and mobility control is given. The related RRM parameters that can be auto-tuned are highlighted.

Admission control and resource allocation

The purpose of the admission control is to calculate which network resources are needed to provide the quality of service requested. According to resources' availability, the new users will be accepted or denied access. The admission control procedures aim to maximize the number of admitted users and to guarantee the QoS of calls being carried out.

The admission control functions take into account a variety of different resources: radio resources, transport resources, Radio Link Control (RLC) buffer size and base station physical resources (time slots). The availability of all these resources is requested when performing an admission control decision. According to this information, different decisions can be made, such as accept/reject the connection, queue the request or perform a directed retry. The resource control performed by the admission control also supports the service retention through the allocation/retention priority attribute. This feature is service dependent, providing users' differentiation depending on their subscription profiles. If there is no enough capacity, high priority users are accepted before lower priority users. Furthermore, once users are admitted, admission control allows high priority users to retain their connections against lower priority users in case of overload.

In circuit-switched domain, admission control decision is normally based on the availability of time slots in the target cell. However, not in all situations the availability of time slots is enough to accept a new connection. It is also possible that new calls are blocked due to high interference levels. However, in packet-switched domain, admission control manages both non-real time and real time connection request differently. In case of non-real time services such as interactive and background traffic classes, the experienced throughput decreases gradually when the number of users increases, up to the blocking situation. By limiting the non-real time load, it is possible to provide a better service to already admitted users.

Recently, dynamic resource allocation between CS and PS services has been implemented in GSM/GPRS networks [15]. For this purpose, a number of time slots is dedicated to each service and a shared band can be used dynamically by both CS and PS users (Fig. 2.2). The CS calls have priority on GPRS traffic on the shared band .This concept of "capacity on demand" is used to adapt the network to an increasing GPRS traffic. The capacity of each band is defined by the operator through specific parameters, or can be dynamically adapted as a function of traffic of each service.

To monitor the network and to check the well functionally of the admission control algorithm, some key performance indicators (KPI) are periodically monitored. We cite some of them:

System accessibility indicators: Covers the user capability to get access to the radio resources. It includes for example Call Setup Success Ratio (CSSR), call setup delay, Session Setup Success Ratio (SSSR), cell load…etc. In general, the system accessibility is measured by the Admission rate or by its complementary call blocking rate.

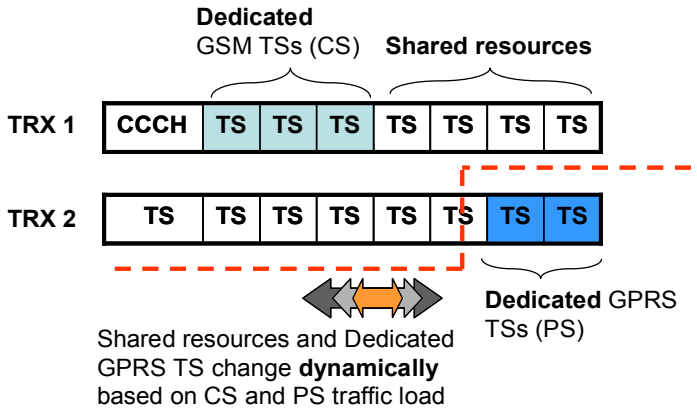


Figure 2.2. Dynamic resource allocations between CS and PS services.

System retainability: Covers the ability to keep up a voice call or a session data connection with a desired quality of service. Retainability could be defined with the dropped calls probability.

Indicators related to user situation: includes for example the received signal strength, the experienced bit rate.

For the call admission control algorithm, some parameters could be candidates for the auto-tuning processes. For example, we cite:

- Number of channels dedicated to CS service or/and the number of channels dedicated to PS service: The dynamic adaptation of these parameters could balance the perceived QoS between different service classes and leads to an efficient use of resources.
- Maximum load: used to prevent the system congestion.
- Thresholds for guard channels: New calls and handover calls are competing for the usage of the radio resource. Therefore, it's very desirable to maintain calls already in the network by reserving for them a handover guard band. This threshold can be dynamically adapted according to the call dropping rate and call blocking rate.
- Time out threshold: when the perceived qualify of service degrades and becomes below a certain threshold during a period equals to the time out threshold, the connection is dropped.

Mobility management algorithm

The mobility management comprises 2 phases dependent on the user situation. The first is when a mobile is in idle mode and the second is when it is in the connection mode. For the first mode, mobiles perform selection and reselection procedures, whereas handover mechanism is performed in the second mode.

GSM idle mode

In the GSM idle mode, the standard specifies two criteria [12]. The first one, denoted C1, is used for the cell selection and reselection procedures, and the second one is reselection criterion C2. The path loss criterion parameter C1 is defined by:

$$C1 = (A - Max(B,0))$$

Where

$$A = RLA_C - RXLEV_ACCESS_MIN$$

$$B = MS_TXPWR_MAX_CCH - P$$

RLA_C is the average received signal level. RXLEV_ACCESS_MIN is the minimum received signal level at the mobile station required for access to the system. MS_TXPWR_MAX_CCH is the maximum transmitted power level a mobile may use when accessing the system until otherwise commanded. P is the maximum output power of the mobile. All values are expressed in dBm.

The path loss criterion [13] is satisfied if $C1 > 0$.

The reselection criterion C2 is only used for cell reselection and is defined by:

$$C2 = C1 + CELL_RESELECT_OFFSET$$

CELL_RESELECT_OFFSET applies an offset to the C2 reselection criterion for that cell. This parameter may be used to give different priorities to different bands when multi-band operation is used.

Cell reselection is triggered if C1 falls below zero for a period of 5 seconds or if the C2 value of the new cell exceeds the C2 value of the serving cell by at least *Cell_Reselect_Hysteresis* dB, for a period of 5 seconds. Enhanced version of the reselection criterion is given in [12] and [13].

GSM connected mode

Handover process takes care of ensuring that any user is always connected to the most suitable cell. Handover in GSM is a hard handover, i.e. the connection is released from the old cell and established with the new serving cell. Handover can be performed for different reasons. For instance, when the user leaves the dominance area of the actual serving cell or when the call experiences bad quality.

In power-budget-based handover, the mobile measures periodically the received signal level of its serving cell and the neighbouring cells. When it detects another cell with better signal level, it carries out a power-budget-based handover to the target cell. The quality-based handover is urgently triggered due to the degraded quality or to the low received signal level. Furthermore, the mobile initiates a distance-based handover, when it becomes very far from its better cell. In GSM systems, a user can initiate a voice communication with a cell, which does not support packet switching traffic, and changes later to GPRS services. In this case, the mobile is likely to be handed over to another cell supporting GPRS services.

For the mobility control, the most appropriate parameters that can be auto-tuned are:

- Handover margin: its auto-tuning is useful for the congestion control.
- List of neighbouring cells: the optimization of the neighbouring cell list assures the connectivity of a user to the best cell. It also reduces the measurements time.
- CELL_RESELECT_OFFSET: the auto-tuning of this parameter dynamically balances the load between the GSM bands.
- Handover capture margin: used in inter-band handover.

Congestion control

Congestion control mechanisms should be designed to face situations in which the system has reached a congestion situation and therefore the QoS guarantees are at risk due to the evolution of system dynamics. The task of congestion control is to monitor, detect and handle situations when the system is reaching a near overload or an overload situation with the already connected users. This means that some part of the network has run out, or will soon run out of resources. The congestion control should then bring the system back to a stable state as seamless as possible. The congestion control process is divided into 2 phases:

Phase 1: Congestion detection phase

When users that are already admitted can not satisfy their guaranteed QoS to their services for a specific percentage of time, the network is considered to be in an overload/congestion situation. The cell load, congestion rate and other blocking indicators are the indicators of the congestion situation.

Phase 2: Congestion resolution phase

Once congestion has been detected, all new sessions and handover are rejected as they will increase more the load of the network. The control algorithm checks if the overload is caused due to users that violate their QoS restrictions in terms of bit rate, which means that it tries to find if there are users that transmit with higher bit rate than they should, according to the service agreements made between the network and the user. If such users exist, then the algorithm lowers their bit rate to the value defined in the service agreements. If there is still congestion, ongoing sessions are inserted into a table, ordered by priority.

The auto-tuning of RRM parameters of each cell is very suitable to reduce the network congestion and to dynamically balance the load between cells. For instance, we can control the size of a cell by just modifying the value of C2. With this technique a user can be forced to make a cell reselection to another under-loaded cell. Adapting handover margin is also a solution for accelerating handover of users from congested cell to neighbouring cells. Some other parameters are also considered as candidates for auto-tuning such as: maximum load and time to trigger congestion algorithm.

2.2.3 Literature review of GSM auto-tuning

The research on auto-tuning of GSM networks started in the mid-90s. *Edwards* and *Sankar* introduced in [25] a new GSM handover algorithm based on fuzzy logic control [53]. They considered the received signal strength and the mobile distance from the base station as the

monitored indicators for the auto-tuning process. The GSM handover margin was tuned according to a set of rules. Their method had been extended later to the handover in microcellular environment using more relevant quality indicators [26].

The most relevant example of auto-tuning in GSM network is given in [27] and [28]. Authors have studied an auto-tuning hierarchical network based on the dynamic adaptation of a handover threshold. The goal of the auto-tuning architecture was to improve capacity while maintaining quality in a network with GSM hierarchical cell structures (HCS). GSM HCS is a combination of micro and macro cells. Traffic is balanced between micro and macro layer by inter-layer handover. With the traffic migration between layers, a better load distribution is obtained allowing the full exploitation of the available resources. The mobile station uses the received signal strength from a cell to make handover decisions. The HCS feature defines a signal strength threshold (LEVTHR) for each cell in the micro-layer. If a mobile station connected to a macro cell measures a signal strength from a micro-cell higher than the threshold, a handover to the micro-cell is performed. Likewise, if the signal strength falls below this threshold, the mobile station abandons the micro cell.

In [28], the role of the self-tuning agent is to correctly estimate the handover threshold (LEVTHR), as a function of the capacity and the quality (interference) in all cells, to avoid congestion in the micro-cell, and to avoid degradation of QoS due to interference.

A successful field trial has been carried out in Hong Kong with the operator *SmarTone* [28]. Five GSM900 micro-cells were chosen in layer one and additional 100 neighbouring cells were also included (although it is not mentioned to which layers these cells belong). Every 10 seconds load and QoS data were sent to the self-tuning agent. The parameter *LEVTHR* is auto-tuned Every 5 minutes. Ericsson claimed that capacity and QoS are improved for the five test cells. The congestion in the tested cells was reduced from 5-10% to 1%.

With respect to the dynamic channel reservation and admission control, authors in [29] and [33] gave a comparative study between two mobility history-based schemes. The aggregate history of mobility observed in each cell is used to predict probabilistically the direction of an MS and to reserve for it a band in the target cell. It is also remarkable to see the reference [30] in which the authors introduced a method for dynamically adjusting the reserved bandwidth in each cell as a function of the monitored dropping probability or the utilization level of the guard capacity. In [33] the mobile positioning availability, as obtained by GPS (Global Positioning System) or other positioning techniques if available with the particular mobile system, serves as a basis for next cell prediction. This system is claimed to take advantage of real-time measurements instead of history-based schemes to make predictions [29]. In [34] a prediction based on a priori defined traffic model is proposed. In [35], authors use a developed mobility model and measured dropping probabilities to adjust the reserved bandwidth.

2.3 UMTS networks

This section presents a comprehensive overview of UMTS network and its evolution, HSPA (High Speed Packet Access) [17] [18]. RMM algorithms are presented as well with a special emphasis on different RRM parameters that can be auto-tuned.

2.3.1 Overview of UMTS Network and its evolutions

UMTS network has been introduced as a third generation mobile communication system. 3GPP organization [20] is in charge of its specifications. It has specified different technologies for UMTS networks: for example, Frequency Division Duplex (FDD), Time Division Duplex (TDD), and HSPA. In the majority of 3GPP specification documents, the name UTRAN is used to stand for the UMTS radio access network. The transmission rate capability of UMTS provides at least 144 kbps for full-mobility applications in all environments, 384 kbps for limited-mobility applications in the macro- and microcellular environments, and 2 Mbps for low-mobility applications particularly in the micro or pico-cellular environments. In 3GPP release'5, the transmission rate capability is enhanced for the downlink to reach 10 Mbps.

The UMTS system offers different types of quality of service (QoS) for different types of customers and their applications. A key QoS attribute includes priority access for different types of users. For example, real time priority access typically applies to voice services and reliable data transfer is applied to interactive data services. There are four different QoS classes: conversational, streaming, interactive and background class [21]. The main distinguishing factor between these QoS classes is how delay sensitive the traffic is.

Whilst the UMTS radio interface is completely new with respect to any 2G system, the core network (CN) infrastructure is based on an evolution of the current GSM/GPRS one. Figure 2.3 shows UMTS network architecture as it stands in Release'5 [19]. It consists of a set of Radio Network Subsystems (RNSs) connected to the CN via the Iu interface.

The CN primarily consists of a circuit-switched (CS) domain and a packet-switched (PS) domain. These two domains differ in how they handle user data. The CS domain offers dedicated circuit-switched paths for user traffic and is typically used for real-time and conversational services, such as voice and video conferencing. The PS domain, on the other hand, is intended for end-to-end packet data applications, such as file transfers, Internet browsing, and e-mail.

The RNS consists of a controller (the Radio Network Controller, or RNC) and one or more entities called Nodes B, which are connected to the RNC through the Iub interface. A Node B superintends a set of cells which may be FDD, TDD, or mixed. In UMTS, Different RNCs can be connected to each other through the Iur interface. RNC is the boundary between the radio domain and the rest of the network. The protocols opened in the user terminal to manage the radio link are terminated in the RNC. Above the RNC are the protocols that permit interconnection with the CN and which depend on it.

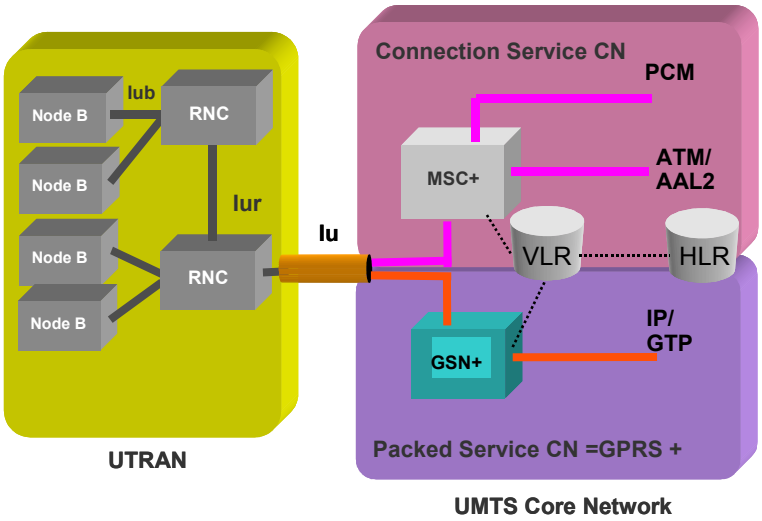


Figure 2.3. UMTS architecture.

UMTS networks uses in the air interface WCDMA (Wide-band Code Division Multiple Access) access technology. The concept of WCDMA inherits from the spread spectrum CDMA. CDMA uses a form of direct sequence which is, in essence, multiplication of a more conventional communication waveform by a pseudo-noise binary sequence in the transmitter. Spreading takes place prior to any modulation, entirely in the binary domain. The noise and interference, being uncorrelated with the pseudo-noise sequence, become noise-like and increase in bandwidth when they reach the detector. Filtering mechanism that rejects most of the interference power can enhance the Signal-to-Interference plus Noise Ratio (SINR). It is often said [19], that the SNR is enhanced by the processing gain W/R, where W is the chip rate and R is the data rate. WCDMA uses a chip rate equal to 3.84 Mcps which leads to a carrier bandwidth of approximately 5 MHz. The inherently wide carrier bandwidth of WCDMA supports high user data rates and also has certain performance benefits, such as increased multi-path diversity. WCDMA uses variable spreading factor and multi-code connections. In addition to the basic radio access capabilities, UMTS architecture provides several other advantages, including higher bandwidth over the radio interface and better handoff mechanisms, such as soft handover for circuit-switched bearer channels. Soft handover refers to the ability to maintain an ongoing connection between the mobile terminal and the network through more than one base station; this capability is particularly important in 3G systems.

UMTS system underwent a considerable evolution by the introduction of the HSDPA (High Speed Downlink Packet Access) in 3GPP release'5 [22]. HSDPA specification was released in 2002 and it was considered the most significant radio related update since release'99. HSDPA is based on a distributed architecture where the processing is closer to the air interface at the Node B for low delay link adaptation.

To achieve a high-speed downlink transmission, HSDPA implements a scheduling for the downlink packet data operation, higher order modulation, adaptive modulation and coding,

Hybrid Automatic Repeat Request (HARQ) and link adaptation according to the momentary channel conditions.

The HSDPA concept offers over 100% higher peak user bit rates than Release'99 in practical deployments. It is comparable to Digital Subscriber Line (DSL) modem bit rates in wireline communication. It extends the UMTS bit rates up to 10 Mbps. This higher bit rates are obtained with higher order modulation, 16-QAM, and with adaptive coding and modulation schemes. HSDPA is able to support not only non real time UMTS QoS classes but also real time UMTS QoS classes with guaranteed bit rates.

A new improvement in the UMTS radio interface was specified in release'6 with the introduction of HSUPA (High Speed Uplink Packed Access) [23]. HSUPA uses an uplink enhanced dedicated channel (E-DCH) with dynamic link adaptation methods as already enabled in HSDPA, i.e. shorter transmission time intervals, thereby enabling faster link adaptation, and also a hybrid HARQ with incremental redundancy, thereby making retransmissions more effective. HSUPA offers peak data rates up to 5.5 Mbps.

The last evolution of UMTS networks is observed in release'7 with the introduction of a new system which has been at the 70% of its specifications while writing this dissertation. This evolution is referred to the 3GPP Long Term Evolution (LTE) systems [97] [98]. LTE system, called sometimes "*super 3G*" or "4G", is expected to offer a spectral efficiency between 2 to 3 times bigger than 3GPP release'6. It will provide up to 100Mbit/s for 20 MHz of spectral bandwidth. Both the radio and the core network parts of the LTE technology are impacted: The system architecture is more decentralized; The RNC present in the 3G systems is removed; and RRM functionalities are moved to an "upgraded" base station called Evolved Node B (eNB). Further description of LTE system is given in chapter 5.

2.3.2 UMTS RRM algorithms and related parameters

In UMTS, RRM algorithms are needed to guarantee QoS, to maintain the planned coverage area, and to offer high capacity. RRM functionalities include mobility management, Power Control, Admission Control, Packet Scheduling, Load Control, Dynamic Channel Allocation and Code Management [24].

Since the WCDMA access technology is interference limited, power control is needed to keep the interference levels at minimum in the radio interface and to provide the desired QoS. Like other cellular networks, Handover mechanisms are needed to handle the mobility of the user across cell boundaries and to realize a continuous coverage. Admission control, load control and packet scheduling are required to guarantee the quality of service and to maximise the system throughput with a mix of different bit rates, services and quality requirements.

In this section, we survey UMTS RRM algorithms with a special focus on admission control, mobility management and load control. For each algorithm, the RRM parameters that can be auto-tuned are given.

Admission control

Admission control regulates the operation of the network in such a way that ensures uninterrupted service provision, and accommodates in an optimal way new connection requests. Admission control estimates whether a new user should have access to the system without impairing the quality requirements of existing users. If the acceptance of a user will increase the interference power and the load on the cell to a level whereby the quality of the ongoing calls is reduced and the quality of the call itself cannot be guaranteed, the user will not be admitted to the system. Due to the dependence between capacity and coverage in WCDMA technology, the coverage area of the cell is reduced below planned values and the quality of service cannot be guaranteed if the radio load increases excessively. The availability of transmission resources is also verified by the admission control process.

There is no absolute number of maximum available channels that can be allocated to potential users. This is the "Soft Capacity" property of UMTS. The number of connections can not specify the actual capacity of the cell. The limits in capacity are determined by the interference that is generated at the base station by all the signals that are transmitted by the users in the same cell and other cells, and by the propagation channel conditions in the coverage area. The user position within the cell affects the capacity of the cell. This means that the load on the cell cannot be predicted by the number of connections at any one time, and capacity values alone can not be used to admit or reject users.

The admission control algorithm is executed when a bearer is set up or modified. The effective load increase by admitting another bearer is estimated, both for the uplink and downlink. The bearer can only be admitted if the uplink and downlink admission control admit it, otherwise it is rejected due to excessive interference on the radio network.

The admission control functionality is located in the RNC where the load information from several cells can be obtained, as well as being estimated in the uplink and downlink. The term Controlling RNC (CRNC) is used to define the RNC that controls the logical resources of its UTRAN access points. In case one mobile-UTRAN connection requires resources from more than one RNC, the CRNC involved has two separate logical roles with respect to the mobile connection. When the RNC holds the Iu bearer for a certain UE it is called the Serving RNC. If another RNC is involved in the active connection, it is known as the Drift RNC.

The adaptation of parameters involving the admission control improves the system capacity and the perceived user quality. Different parameters are subjects for auto-tuning process. Among them, we cite for example:

- The percentage of power and number of codes dedicated to signalling channels: the auto-tuning of these parameters allows exploiting resource dedicated to the signalling channels for traffic channels.
- The dedicated code and power to traffic channel. It is very useful to balance the quality of service of different traffic classes.
- The load threshold: the adaptation of this parameter prevents the system from congestion

- The maximum load. The difference between the load threshold and the maximum load is generally reserved to the mobility. Its adaptation makes a best trade-off between blocking and dropping rates.
- The Downlink/Uplink SIR target. It is useful to adjust the perceived quality of each user, so an amount of power can be saved and the interference is reduced.

Mobility management

The mobility management mechanisms include handover and cell selection-reselection. The cell selection-reselection occurs when the user is in idle mode. Handover occurs when a user in connection moves from the coverage area of one cell to the coverage area of another one. It can also be performed between frequencies or to distribute load/users among neighbouring cells in the case of the overload cell situation. Efficient handover algorithms are a cost-effective way of enhancing the capacity and QoS of cellular systems.

UMTS idle-mode

The cell selection process allows the user to select a suitable cell where to camp on in order to access available services. In this process the mobile can use stored information (*Stored information cell selection*) or not (*Initial cell selection*) [37]. For the initial cell selection, the user shall scan all radio frequencies in the UTRAN bands according to its capabilities to find a suitable cell. On each carrier, the user seeks the strongest cell. Once a suitable cell is found this cell shall be selected. The procedure of *Stored Information Cell Selection* requires stored information of carrier frequencies and optionally also information on cell parameters, e.g. scrambling codes, from previously received measurement control information elements. Once the user has found a suitable cell the user shall select it. If no suitable cell is found the initial cell selection procedure shall be started.

For UMTS FDD mode, the cell selection criterion S is fulfilled when:

$$Srxlev > 0 \text{ and } Squal > 0$$

where

$$Srxlev = Qrxlevmeas - (Qrxlevmin + QrxlevminOffset) - Pcompensation$$

$$Squal = Qqualmeas - (Qqualmin + QqualminOffset)$$

The Srxlev and Squal are respectively Cell Selection Rxlev value and Cell Selection quality value. Qrxlevmeas is the measured Received Signal Code Power of the pilot channel, (CPICH_RSCP). Qqualmeas is the measured cell quality value. The quality of the received signal is expressed in CPICH Ec/N0 for FDD cells [37]. Qrxlevmin is the minimum required Rxlev in the cell. QrxlevminOffset is an Offset to the signalled Qrxlevmin taken into account in the Srxlev evaluation as a result of a periodic search for a higher priority network [40]. Qqualmin is the minimum required quality level in the cell and QqualminOffset is its Offset used as the same way of QrxlevminOffset. For Pcompensation, reader is invited to see [37]. Other specifications

of network and cell selection are found in [38] and [39]. Procedures and criterions for cell re-selection are found in [37] and [38].

UMTS connected mode

UMTS connected mode refers generally to the procedures performed by user when it is in communication state. The handover mechanism constitutes the main procedure in connected mode. In UMTS system, there are 4 types of handover: Soft handover (intra-frequency HO), softer handover (intra-frequency and intra-site HO), hard handover (inter-frequency HO) and inter radio access technology handover.

Soft Handover (SHO) occurs when a new connection is established before the old connection is released. During a soft handover it is possible for multiple cells to simultaneously support the mobile station's call. An algorithm that implements this kind of handover is the "Active Set Update" algorithm. This algorithm is found in the specification [24], however, each constructor has its own algorithm. The algorithm is detailed in chapter 4.

The soft handover parameters have to be carefully monitored, as excessive soft handovers can impair the downlink capacity. Each soft handover connection increases the transmitted interference to the network. More orthogonal codes are used in the downlink using soft handover connections than single link connections. It is the task of radio network planning and optimization to keep the soft handover overhead below 40% while still providing enough diversity in the uplink and downlink [19]. The dynamic adaptation of soft handover parameters could tackle the issue of best trade-off between soft handover overhead and cell capacity.

During a softer handover, a mobile station is in the overlapping cell coverage area of adjacent sectors of one base station. A communication link between the mobile and each sector is established by using separate codes in the downlink, so the mobile can distinguish the signals.

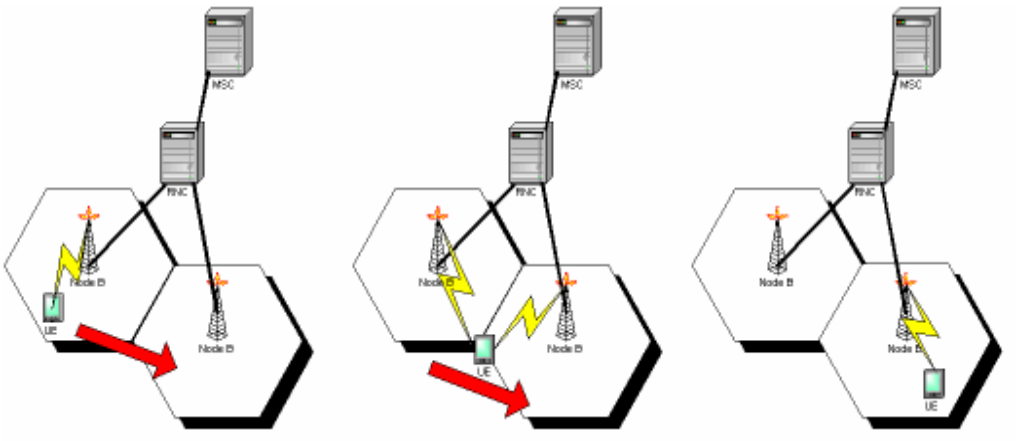


Figure 2.4. Intra-Frequency Soft Handover

In hard handover, the user has its previous radio links removed, and replaced by other radio links. Inter-frequency hard handover is used to ensure the handover path from one cell to another cell in the cell cluster where the frequencies are not the same. Generally the frequency reuse factor for

UMTS is one, meaning all the base stations transmit on the same frequency. However, this does not mean all base stations are required to transmit on a common frequency. Inter-frequency handover is used for example in Hierarchal Cell Structure (HCS) networks between separate layers and in heterogeneous networks between different Radio Access Technologies (RAT). Because hard handover involves interrupting the call while the call support is changed, it is also known as "break before make".

The soft handover parameters that can be subject of auto-tuning are:

- Maximum active set size: it is the maximum number of cells in the active set connected to the user. The auto-tuning of the maximum active set size dynamically optimizes the number of mobiles in soft handover condition and the frequency of the active set update. This latter is related to the ping-pong effect [4].
- Soft handover Hysteresis_event1A, Hysteresis_event1B and Hysteresis_event1C: they are used respectively for adding, deleting and replacement of a link in the active set of a mobile. The auto-tuning of these parameters is very useful to relief permanent localised congestion problems caused by the uneven appearance of traffic in a cellular network both in time and space. Traffic balancing through permanent adaptation of SHO parameters on an adjacency-by-adjacency basis can greatly minimise congestion without the need for any hardware upgrades, thus providing a cost-effective method to increase network capacity [4]. The hysteresis region is maintained to avoid unnecessary handovers due to the ping-pong effect.
- Neighbour cell list: its auto-tuning allows reducing the measurement time of signals coming from cells and permits the user to select or to be handed over to the best cell in the network. It is also suitable for reducing the neighbour cell list.

Congestion control

As in a GSM network, congestion control in UMTS assures the well functionality of the system and brings it back to a stable state when overload situation occurs. The congestion control procedures include two phases: detection and resolution:

Congestion detection: Some criterion must be introduced in order to decide whether the network is in congestion situation or not. A possible criterion to detect when the system has entered the congestion situation and trigger the congestion resolution algorithm is when the load factor η increases over a certain threshold η_{CD} during a certain amount of time, ΔT_{CD}, i.e. if $\eta \geq \eta_{CD}$ in 90% of the frames within $\Delta T_{CD.}$ The load factor η measures the theoretical spectral efficiency of a WCDMA cell and must be $0<\eta<1$. Usually the network is planned to operate below a certain maximum load factor $\eta_{\max}$ and the congestion detection threshold should be set in accordance to the maximum planned value.

Congestion resolution: When congestion is assumed in the network, some actions must be taken in order to maintain the network stability. The congestion resolution algorithm executes a set of rules to lead the system out of the congestion status. The system can reduce the throughput of packet data users and push mobiles to another WCDMA carrier or to a GSM network. It can also decrease bit rates of real time users and drop low priority calls in a controlled fashion.

The auto-tuning of some RRM parameters is seen too as a solution for congestion. So, by means of adapting handover parameters, traffic is pushed from overloaded cells to under-loaded cells. The adaptation of cell-selection criteria of a cell as a function of its load and the load of the neighbouring cells allows reducing the coverage area of overloaded cells and stretching under-loaded cells' ones. The service area of a cell can also be affected by adapting the power assigned to the CPICH of each cell.

2.3.3 Related work on the auto-tuning of UMTS parameters

A plethora of references addressing the issue of auto-tuning and RRM parameters' adaptation is found in the related literature. In [1], the Gandalf project proposes a physical auto-tuning functional architecture. The proposal deals with the on-line and off-line auto-tuning, namely the controller of the auto-tuning takes action quickly (every seconds or minutes) or slowly (1 hour and more). The proposed architecture was given for management plane, control plane, and user plane. The relationship between layers has been given as well.

In [41], the theoretical approach of automatic neighbour cell list optimization was introduced. An initial neighbouring cell list is used in the RNC and allows calculating network statistics. The RNC collects and sends performance statistics (such as average CPICH RSCP reported by mobile terminals, HO proportions, HO success ratio and HO effort) to an optimization tool. Based on RNC HO statistics, unnecessary cells are removed from the list and new cells are added to the list. The auto-tuned neighbour lists are sent back to the RNC. In [41], this method has also been tested in a commercial network of 95 cells to optimize neighbour cell lists in 3G system. The overall system quality (in terms of successful HO procedures) has been improved and the average length of the neighbour cell lists is significantly reduced.

In [38], 3GPP has already specified an online optimization approach of the neighbour cell list, called *Detected Set Reporting*. In the approach, the network commands mobiles to detect and report cells which are not on the neighbour list. *Detected set reporting* is event-triggered but without any active set update procedure.

In [42], authors from Nokia described a method to automate the setting of common pilot power in a WCDMA network. The CPICH power auto-tuning algorithm is based on the gradient descent method to minimize a cost function. This cost function takes into account two items: the deviation of the coverage from the target and the deviation of the load from the load in the neighbouring cells. The method showed slightly encouraging simulation results. Two methods are introduced in [43] to estimate the uplink and downlink planned E_b/N_o (meaning the estimated E_b/N_o requirements per service for proper decoding of the signal) of WCDMA services and afterwards to auto-tune the uplink and downlink E_b/N_o values for packet data. Simulation results showed that this auto-tuning method allows improving the system performance in terms of throughput. The auto-tuning of WCDMA link power per service and the cell downlink load level targets was given in [44]. The quality indicators used in [44] were call-blocking probability, packet queuing probability and downlink power outage. These indicators were compared to allowed levels and a table of heuristic rules gave the parameter adjustments depending on the deviation of the indicator from the allowed levels.

In [45], handover parameters are auto-tuned based on cost function. The cost function depends on blocked call ratio and downlink transmit power. A second order gradient method is adopted to minimize the cost function. Simulation results showed that HO window auto-tuning method allows improving the system performance in terms of blocking rate.

Homnan et al. [46] discussed the feasibility of controlling soft handover threshold in IS95 and CDMA2000 networks [19]. Authors considered the E_b/N_0, the outage probability and the remaining capacity in the serving cell as inputs to the controller. A set of rules, called fuzzy inference system, have manually been constructed based on the radio engineering experience. Likewise reference [47] introduces fuzzy inference systems applied to the issue of soft handover in CDMA mobile communication networks, presenting new algorithms for threshold adjustment aimed at reducing call blocking probability and controlling the quality of service.

Ye and his colleagues have proposed in [48] a new CDMA fuzzy call admission control algorithm based on the estimation of some indicators to decide whether to accept or to block a new or handover call. In their admission control algorithm, fuzzy logic is used to estimate the user mobility and the effective bandwidth that would be used by the new user requesting access. Likewise references [49] and [50] present call admission control schemes utilizing the fuzzy logic approach in CDMA systems, highlighting its effectiveness in handling mobility and traffic model uncertainty.

In [51], an optimized fuzzy logic controller has been proposed for simultaneously auto-tuning admission control and macro-diversity parameters. Fuzzy rules are designed and optimized using a combinatorial method called *particle swarm*. Since the optimization is very time-consuming, it is carried out off-line. In the thesis dissertation of Herve Dubreil [52], both combinatorial method and Reinforcement Learning (RL) [53] have been used to optimize the auto-tuning process of handover and admission control parameters in UMTS systems. The auto-tuning tool was a fuzzy logic controller.

In [54] the authors have proposed a RL-based call admission control algorithm. Instead of controlling RRM parameters, their proposed scheme assists the system in a decision process. In each network situation, the controller has only two actions: reject or accept the call. A small number of states (system input indicators) is used, namely the number of mobile users in each traffic class.

In [55] and [56], the QoS provisioning problem by means of Q-learning algorithm has been presented. Two algorithms of resource management are considered: call admission control and bandwidth adaptation. Two QoS constraints are taken into account in [56]: the probability of hand-off dropping and the normalized average allocated bandwidth of each traffic class. In their papers, authors have chosen the Q-learning algorithm, because it does not require closed form formulae of the system. Chen et al. [57] have used as well Q-learning method to adapt the transmission rate in the WCDMA radio resource control and to accurately estimate the cost for the multi-rate transmission problem. All the above references using reinforcement learning have modelled the system as a Semi-Markov Decision Process (SMDP).

2.4 Multi-system networks

Beyond 3G systems (B3G) consist of a number of existing systems with different radio access technologies such as GERAN, UTRAN, WLAN, etc. providing different QoS and coverage. The interoperation between them is an essential requirement to achieve a seamless service as well as efficient mobile environment with end-to-end QoS.

When high degree of integration between different technologies exists, the provisioning of service is more efficient and the selection mode of the best radio access is faster as well as the handover procedure. On the other hand, a higher level of integration requires undoubtedly bigger effort in the definition of interfaces and mechanisms able to support the necessary exchange of data and signalling between the different radio access networks. Therefore, more sophisticated RRM mechanisms (e.g. aware of the different resources of each available RAN, of the different offered services, etc.) are needed to allow a fruitful system inter-working. This sophisticated RRM mechanisms, called equally Common RRM (CRRM) [64], Advanced RRM (ARRM), Joint RRM (JRRM) [1] or Multi-RRM (MRRM), provide a solution to transform overlapping radio access networks into a jointly, efficiently co-ordinated, pool of radio resources. These CRRM algorithms include inter-system mobility mechanisms, admission control, load control and packet scheduling. CRRM entity shall have information on the state of the radio resources in each RAN. For example, the measurements such as load information of the neighbouring cells are an important input of the CRRM algorithm and must be collected from the different RRM in each RAN.

2.4.1 GERAN UTRAN inter-working

Interconnection between UTRAN and GERAN has recently been presented by 3GPP in technical reports [58] [59] and the handover from an UMTS to a GSM/GPRS network can be triggered even if this handover is left to vendors' implementations.

In release'99, procedures for inter-system handover have been defined, but they could result in a failure due to high load in the target cell. This situation has motivated 3GPP standardization to introduce cell load information exchange between systems in Release'5. The notion of CRRM was introduced in 3GPP in 2001 and a technical report has been edited [58], collecting proposals and simulations from different companies. The Release'5 work resulted in the introduction in GSM and UMTS of the possibility to exchange cell load information between RNC and BSC. In [58], two architectures for mapping CRRM logical functionality into physical entities have been proposed. The first is the centralized approach (CRRM server approach) and the second is the decentralized scheme (integrated CRRM approach), where CRRM can be implemented in each RNC and BSC equipment.

CRRM server approach: this approach implements RRM and CRRM entities into separate nodes, CRRM is a stand-alone server. The CRRM server (Fig 2. 5) pilots local RRM entities implemented in the RNC and BSC. Consequently, Each RRM entity in the functional model may be requested by its responsible CRRM entity to report certain information (e.g. measurements) with respect to its radio resources. Then, CRMS (Common Resource Management Server) gathers measurements from cells under its coverage. For each specific operation (handover, cell

change order, etc.), the RNC/BSC sends to the CRMS the list of candidate cells, including the mobile measurements for these cells and information about the QoS required by the mobile. This function shall allow for requesting immediate replies to a measurement request as well as event- or timer-triggered measurement reporting. It is assumed that this reporting is controlled by the responsible CRRM server. The CRMS, after applying some algorithms, returns the prioritised list of candidate cells.

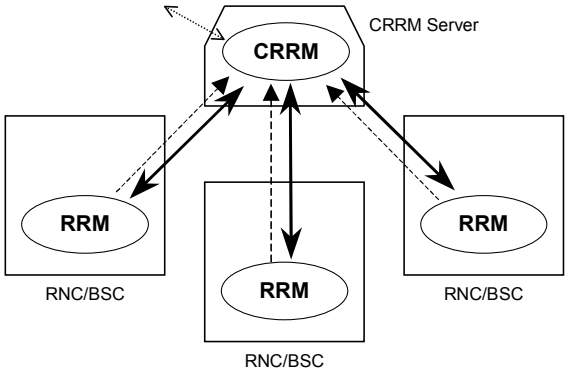


Figure 2.5. Centralized CRRM entity [58].

Integrated CRRM approach: this approach integrates the CRRM functionality into the existing UTRAN/GERAN nodes: each UTRAN RNC and GERAN BSC is equipped with a CRRM and local RRM entities (Fig 2. 6). The functional interface between RRM and CRRM is not realised as an open interface in this solution. Only "*Reporting Information*" is exchanged over open interfaces. The Iur and the proposed Iur-g (between BSC and RNC) already include almost all the required ingredients to support the CRRM functionality. The main benefit of this integrated CRRM solution is that with limited changes and already existing functionality it is possible to achieve optimal system performance. The co-location of CRRM and RRM in only one equipment does not influence the decisions of the local RRM. CRRM is not supposed to be consulted for channel switching or for soft handover for example, since the RRM in RNC shall handle these cases. CRRM will be used only for inter-system mobility including selection, re-selection and handover.

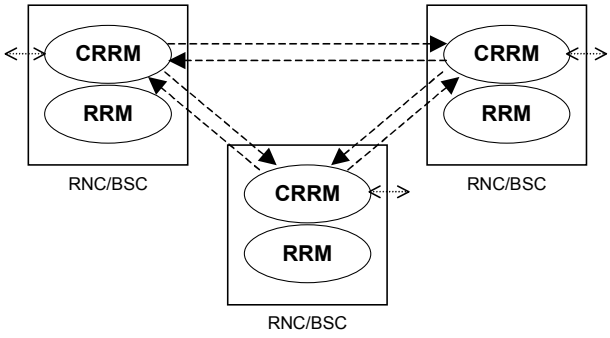


Figure 2.6. Decentralized CRRM into every RNC/BSC [58].

2.4.2 3GPP-WLAN inter-working

Interoperability between 3GPP systems (UTRAN and GERAN) and WLAN networks is an essential requirement to achieve a seamless as well as efficient mobile environment across these different access technologies. The WLAN may be an integral part of the UMTS network where one operator controls both WLAN and UMTS networks. This refers to the Tight Coupling and Very Tight Coupling presented in [64]. In the other hand systems can be separate, i.e., an independent WLAN is interconnected to the network of one UMTS operator or shared by multiples operators. Reference [63] investigates different interconnection level between 3G cellular networks and WLAN. The interaction levels or scenarios, as defined by 3GPP in the technical report [60], span from the simple common billing to the seamless service continuity when moving from the 3G access network to the WLAN access network and vice-versa. WLAN-3GPP system interconnection is defined as a wireless IP connectivity service where the user can obtain access via a Wireless LAN technology. The interworking must be independent of the underlying WLAN radio technologies. Note that in comparison to the 3GPP work developed to normalize the interoperability between RRMs of GERAN and UTRAN, there is still no standard entity devoted to manage RRMs when connecting UTRAN and WLAN.

The first interaction level, presented in [60], aims just at unifying billing and customer care procedures for users in 3G and WLAN networks, i.e., a single customer relationship. It does not require any particular interoperability requirements at radio resource management. The second level focuses on the 3GPP system based access control and charging. It includes authentication, authorization and accounting (AAA) procedures provided by the 3GPP system for WLAN users in 3GPP networks. This allows a direct access to internet from the WLAN access network with an authentication based on the mobile operator's infrastructure. The third level of 3GPP proposal offers to WLAN users the access to all the packet-switching based services provided by 3G systems. These levels are investigated by 3GPP in [61] where a special attention to basic UMTS AAA and charging issues is addressed. The network selection procedures for WLAN network and for 3GPP network are detailed in [61] and [62]. The name of the WLAN network is provided in WLAN beacon signal in so-called SSID (Service Set ID) information element. There are two modes in WLAN network selection: Manual mode and automatic mode. In the manual mode, the terminal shall try to find all available SSIDs through scanning. Once a list of all available SSIDs has been obtained, it shall be possible for the terminal to obtain a list of all available WLAN names from each SSID and shall present them to the user to select one. In the automatic mode, after the scanning procedure, the selection of the WLAN is done automatically according to predefined criteria, for example based on preference lists. As of UMTS selection, the terminal can use procedures similar to existing 3GPP procedures for network selection.

Regarding the last two levels, i.e. level 4 and 5, they are not considered in [61] and refer respectively to the tight and very tight coupling proposed by 3GPP. Level 4 allows service continuity when user changes the access between 3G and WLAN networks. In level 5, seamless service continuity is supported when user moves from 3G network to WLAN and vice-versa. Many open issues to be considered in levels 4 and 5 such as the criteria and the decision

mechanisms for the change of the access network, the change of the offered QoS that can occur, the seamless mobility management, etc.

2.4.3 Literature review of auto-tuning in multi-system networks

Several studies in the literature have proposed an efficient and intelligent JRRM in multi-systems' context. Since the JRRM algorithm must decide which RAT every user in the network is attached to at a given time, most of the proposed studies formulate the JRRM as a fuzzy decision, where engineer experience is taken into account, or/and multi-objective decision problem which can be solved with techniques from the Multiple Criteria Decision Making (MCDM) field [65] [66] [67]. On the other hand, the proposed methods mainly differ in which pertinent system indicators are used to make a good decision.

In [68] and [69], fuzzy layer selection algorithms are presented to decide to which network a user can be handed over. In a vertical handover, the JRRM assigns users to layers in a system consisting of macro-cells and micro-cells based on network indicators. For this purpose, network measurements (i.e. mobile speed, layer occupancy) are fed to a fuzzifier, where a value between 0 and 1 is assigned, corresponding to the degree of membership to given fuzzy set [53]. A fuzzy set is a linguistic representation of the input variable (e.g. high, low). Subsequently, an inference engine makes use of predefined fuzzy rules to indicate, for each RAT, the suitability of selecting it. Thus, the decision process is based on heuristic decision rules extracted from previous experience in this type of networks. The main advantage of fuzzy logic over classical rule-based expert systems is that, during the inference process, several rules can be fired in parallel, which is the key for approximate reasoning. A similar approach based on heuristic decision rules is applied in [70] to solve the inter-working between WLAN and 3GPP networks.

In [71] and [72], the previous approach is extended with MCDM techniques in order to combine network measurements, operator policies and user preferences in the JRRM decision process. A vertical handover algorithm is presented, where both handover initiation and handover decision processes rely on fuzzy logic. During the initiation stage, the need for a handover decision is evaluated based on the heuristic rules residing in a fuzzy inference system. Although this triggering mechanism has been traditionally based on threshold crossing, more refined mechanisms can take advantage of experts' knowledge by means of fuzzy inference. Thus, network performance indicators from each connection (e.g. RSSI, BER/BLER, network coverage and perceived QoS) are monitored to detect abnormal behaviour that requires urgent actions. Once the handover decision process is triggered, the different RATs are compared in terms of different criteria. For this purpose, additional information from the candidate networks (e.g. RSSI and available bandwidth) and the user (e.g. user speed, battery status) must be retrieved. The actual values of these performance indicators are fed to a fuzzifier, whose output is a membership value proportional to the degree of fulfilment of the corresponding criterion for each RAT. A decision matrix is thus obtained, where all the candidate solutions are ranked based on the score for every criterion. To build a unique overall score for each RAT, the MaxMin aggregation method [76] is used. In this method, the membership value assigned to each RAT represents the minimum degree of fulfilment among criteria. The RAT with the maximum membership value is finally selected.

In [73], a similar approach is adopted, where only three criteria related to the network, the user and the operator preferences are considered for the handover decision. To extract the network score for each RAT, several network performance indicators (i.e. RSSI, resource availability and mobile speed) are aggregated by fuzzy heuristic rules. Thus, a unique indicator of adequacy from network perspective is raised for every RAT. To build a final score for each RAT, the resulting values from this network criterion are later combined with operator and user criteria. An additional difference is the initiation of the handover decision process as a periodic event. In the timer-based case, handover is triggered when a timer expires. The latency of this mechanism has traditionally favoured its use for alternative (i.e. non-urgent) handovers, which do not have stringent delay constraints. Hence, this mechanism is suitable to achieve better overall network performance (e.g. through load balancing between layers) or satisfy user preferences (e.g. QoS-cost trade-off). Nonetheless, timers might be used indistinctively to trigger imperative or alternative handover actions, as it is the case, as long as the timer is initialized to a small value (e.g. 100ms). Simulation results in a scenario with three concentric cells prove the robustness of the method. However, the computational load might well be an issue in a real environment, since the handover decision process should be performed every time for every user in the network.

In [74], the previous model is converted into a neuro-fuzzy decision model to take advantage of learning capability of neural networks. The parameters of the fuzzy model (e.g. means, deviations and shapes of the membership functions) are adjusted by the back propagation method [77]. The objective of this adaptation process is to minimize some overall network performance indicator, such as the ratio of non-satisfied users (i.e. ratio of users that receive a bandwidth below a certain desired value). An error measurement is defined as the difference with the target performance value. The gradient of this error respect to the network's modifiable weights is calculated. Finally, this gradient is used in a simple gradient descent algorithm to find weights that minimize the error.

In [75], it is underlined that the use of fuzzy logic in previous work is not to deal with imprecise information, but to ease the use of classical (i.e. not fuzzy) MCDM methods. Fuzzy sets have been traditionally used only to model the flexibility of soft constraints. Thus, the membership functions in the fuzzification stage are employed to raise a score from crisp measurement values to represent the degree of satisfaction of the different criteria. The resulting crisp values are subsequently used by classical MCDM to determine the ranking order of the alternatives, as suggested in [76]. Following this approach, the performance of the three popular MCDM methods is evaluated in the context of vertical handover. The methods considered are the simple additive weighting method [65], the MaxMin method [76] and the technique for order preference by similarity to ideal solution [65]. Results show that the simple additive weighting method provides a relative conservative ranking, less sensitive to user preferences and attribute values. Likewise, it is argued that MaxMin method gives disputable results, since it only uses a small part of the information of the decision matrix.

2.5 Conclusion

This chapter provides an overview of current systems with special emphasis on RRM algorithms and auto-tuning of related parameters. At the beginning of the chapter, GSM infrastructure and RRM algorithms are given. For each RRM algorithm, the related parameters that can be auto-tuned are highlighted. A literature survey on GSM auto-tuning is given for admission and congestion control and mobility management. Likewise, the UMTS architecture and its evolution are given in the second section based on 3GPP specifications. RRM algorithms and related parameters are presented as well. Achievements and current studies of UMTS RRM auto-tuning are presented illustrating the benefits of self-tuning for improving the network capacity and QoS. In the last section, inter-system cooperation and JRRM algorithms are presented. A special focus is given to inter-connexion mechanisms between GSM and UMTS on the one hand and between 3GPP systems and WLAN on the other hand. In the latter case, preliminary studies have been conducted in 3GPP with different interconnection scenarios. New research avenues of intersystem mobility and admission control are summarized. Most of these researches formulate JRRM algorithms as multi-objective decision problems that can be solved by simple (not fuzzy) or fuzzy MCDM methods. Many aspects in auto-tuning techniques and applications remain unexplored and have motivated the present thesis.

Chap. 3 Auto-tuning architecture and tools

3.1 Introduction

In auto-tuning and self configuration context, the core challenge we face is how to design an efficient architecture easy to be implemented in real networks and what tools can be chosen to optimize the auto-tuning tasks. The auto-tuning functionalities should be merged with the network elements according to whether a long term or short tem auto-tuning is required. The simplest way for auto-tuning implementation is to add a new layer in the Operation and Maintenance Centre (OMC) of the network; however this solution can only support long term or off-line auto-tuning [27]. Besides, tools used in the self-configuration should be based on heuristic and optimal methods characterized by high level of intelligence. We use in this thesis, Fuzzy Logic Controllers (FLC) as open loop regulators of RRM parameters. The fuzzy logic theory translates experiences from radio engineers into a simple set of rules which are not always optimally developed. To optimize FLC, we use Reinforcement Learning (RL) algorithms as a dynamic FLC-optimizing engine.

In the present chapter, we start by introducing the self-tuning architecture concepts as proposed by the Eureka Celtic Gandalf project [1]. Both on-line and off-line auto-tuning approaches are discussed. The next sections discuss the fuzzy logic controller and the reinforcement learning algorithms and more specifically the fuzzy Q-learning algorithm [4].

3.2 Auto-tuning architecture

3.2.1 Gandalf management architecture

The architecture proposed by the Gandalf project concerns the use of auto-tuning and troubleshooting [1] [83] in advanced- (for a distinct Radio Access Network (RAN)) and joint- (for two or more RANs) RRM [1]. Figure 3.1 shows the different management tasks ranging from the off-line to the online auto-tuning according to the corresponding time scale. In the long term optimization, we have the off-line auto-tuning and the troubleshooting whereas for the short time scale, dynamic auto-tuning is merged with the RRM and JRRM functionalities. Network off-line auto-tuning are long term management tasks, with time scale varying from a day to months, in which optimal fixed network parameter values and thresholds are derived and manually injected in the network. Diagnosis and troubleshooting are also long term management tasks and used to detect faults and sub-optimal parameters via alarms and monitoring of counters and Key Performance Indicators (KPIs), often centralized in the Network Management System (NMS). This function can be performed on daily basis or longer time period.

The online auto-tuning task should be instantaneous, typically of the order of minutes, seconds or less. To facilitate the information exchange between RRM functionalities and the on-line auto-

tuning management, the auto-tuning should be close to the RRM functions. In other words, it must link the management and the control plane.

3.2.2 Gandalf auto-tuning architecture

In [1], the Gandalf project proposes an auto-tuning functional architecture. Figure 3.2 illustrates a modified and simplified version of the architecture proposed in [1], where only auto-tuning part is taken into account. The proposed architecture, in essence, comprises several blocs: RRM Data Collector, Management and Processing Module (MPM), the JRRM and RRM functions, and the Auto-tuning and Optimization Engine (AOE).

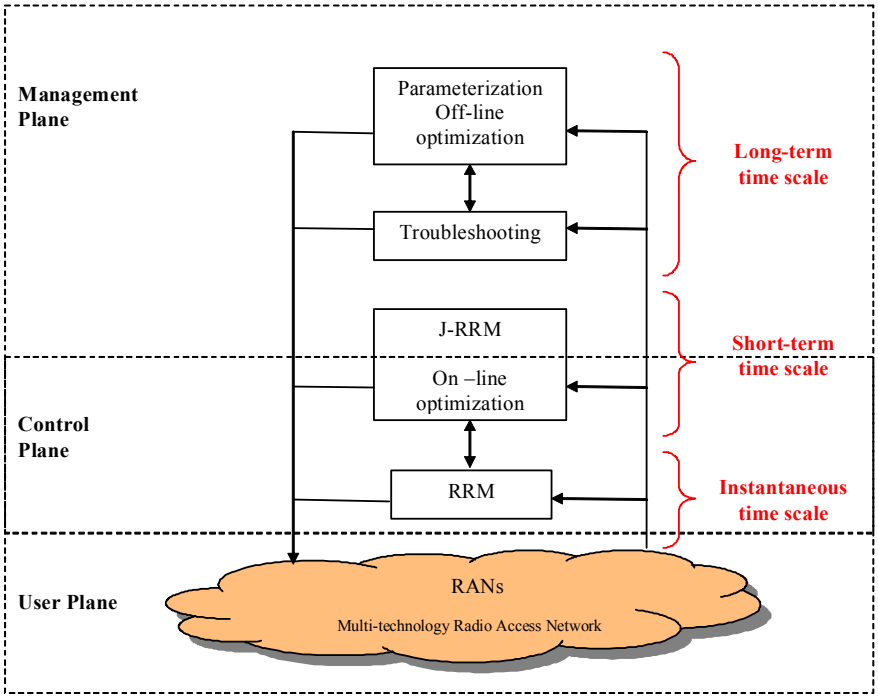


Figure 3.1. Network management tasks with the corresponding time scale [1].

The RRM Data Collector collects measurements concerning the performance of each RAN. Data is the KPIs delivered by the RNC or another equivalent entity. Examples of KPIs could be: traffic blocking, resource availability and access, handovers success and failure, receiver level quality, voice call quality, packet call quality, dropped call rate... The crude data is injected in the MPM.

The MPM is expected to handle both multi-technology and multi-vendor environments. MPM harmonizes and renders transparency the KPIs obtained from different technologies and different vendors.

As mentioned in the previous chapter, the JRRM element is responsible for the global, multi-RAN radio resource management, and aims at improving the quality of service in the global network. The JRRM differs from the RRM, since it exploits the total available heterogeneous infrastructure to provide efficient RRM. Whenever the RRM entity is unable to reconfigure the

system in such a way to solve a problem, the KPIs of each RAN can be directed to a higher level, and the problem may be treated and solved in JRRM or in the AOE. Indeed, the JRRM can reconfigure RRM parameters that will better suite the network needs, or take a decision such as inter-RAN vertical handover, selection and reselection as well as roaming of users from one service provider to another. JRRM enables to reconfigure RRM parameters whenever an auto-tuning engine is directly implemented in the JRRM entity or the latter receives auto-tuning commands from the AOE.

The AOE consists in defining parameter targets to be achieved based on the available KPIs. The main parameters that are fine-tuned are signalling parameters, RRM or JRRM parameters. The auto-tuned parameters are periodically injected in the JRRM or RRM elements. The online auto-tuning process can be considered as a part of the RRM or JRRM functionalities. In this case, RRM is called advanced RRM [1] [85]. Further details of the AOE are found in section 3.2.4.

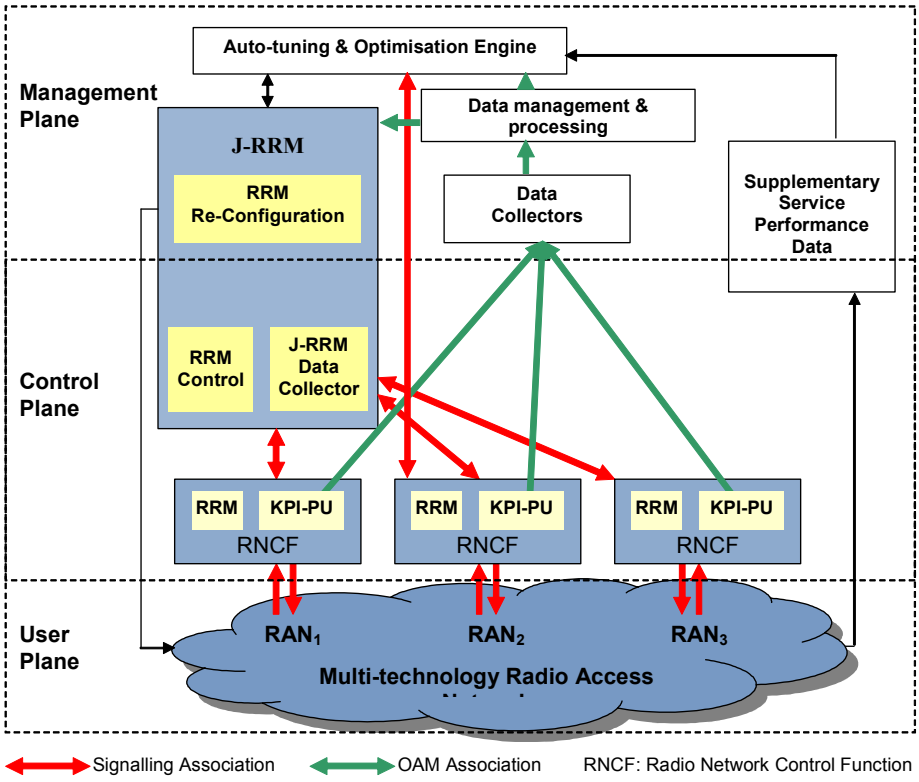


Figure 3.2. Auto-tuning architecture in user, control and management planes.

Data is sent from the upper level to the AOE through the OAM (Operations Administrations and Maintenance) links if off-line auto-tuning is used because it requires a high time to be exchanged. To collect more network data, supplementary service performance data elements can be added. Theses allow complementing counters and KPIs implemented by vendors, and can be used to test, validate or monitor part of the network. There exist many capture tools which are used to give more information on the network performances and to define specific KPIs needed by the operators. Such tools are *Vallent*, *Sunrise*, *Cigale* and *Terms investigation* [84].

In the case of online auto-tuning, signalling messages are sent between entities. In [85], examples of signalling messages and auto-tuning information flows are presented.

3.2.3 Auto-tuning information flows

An example of auto-tuning message sequence chart (MSC) is presented here based on [85]. The MSC expresses how signalling messages could be sent between entities in case of auto-tuning is implemented. The MSC is depicted in figure 3.3 and described by 7 steps:

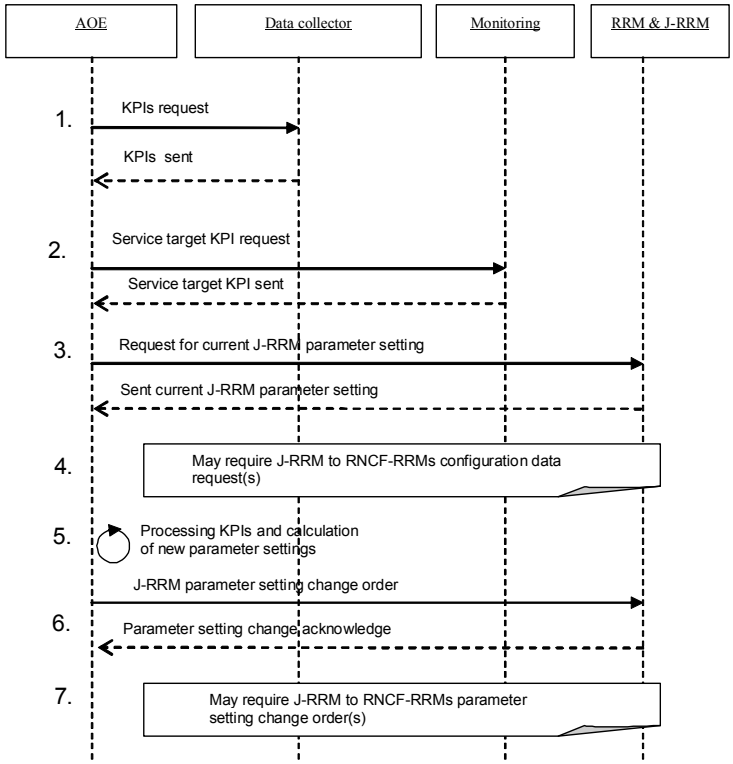


Figure 3.3. Signalling messages between AOE and JRRM/RRM modules.

Step (1): The AOE requests and receives KPIs from the data collector module.

Step (2): The AOE requests and receives inter system service target indicators from the monitoring module, such as target blocking and dropping rates in each system. These indicators together with those from the data collector module are used to guide the optimization process.

Step (3): The AOE checks the current parameter settings of the J-RRM to guarantee coherence of the proposed modifications (in steps (4-5)).

Step (4) (optional): It may be necessary for the J-RRM to request current RRM parameter values from the network elements.

Step (5): KPIs are processed and the new RRM parameters are set by the fuzzy Q-learning controller inside the AOE. To allow learning, new functions depending on the controller may be built and stored in a look-up table (see next sections).

Step (6): Modification order of RRM or JRRM parameter settings is transmitted to RRM module, and an acknowledge message is sent back.

Step (7): The JRRM module maps the parameter changes requested by the AOE into parameter changes of RRM for each RAN.

3.2.4 Auto-tuning and optimization engine

The AOE is presented by the simplified figure 3.4. It chiefly contains 4 blocks: data processing, Fuzzy Logic Controller (FLC), Q-learning algorithm and a look-up table. The first part serves as the intermediary between the network and the AOE. In order to stabilize the auto-tuning process, the data processing module should filter KPIs, received from the network (from any RRM or JRRM module), over a filtering period. In fact, averaging KPIs allows smoothing out its random fluctuations. It makes actions or decisions of the controller to be based on underlying trends and not instantaneous changes. Mathematically, the filtering function of any received quality indicator x is made by averaging the instantaneous value of the indicator x during a discrete time window T_f:

$$Fil(x) = \frac{1}{T_f} \sum_{i=0}^{T_f - 1} x(t - i) \tag{3.1}$$

The data processing module allows also defining an objective function from the received KPIs. The objective function, called also reward signal, should be maximized by the Q-learning algorithm. The objective function is obtained by making operation between received KPIs.

The FLC is responsible for dynamically adapting RRM or JRRM parameters according to received quality indicators. The adaptation is performed optimally thanks to the Q-learning algorithm. The combination of both FLC and Q-learning algorithm justifies the name Fuzzy-Q-Learning Controller (FQLC). FLC and Q-learning algorithm are further given in next sections.

The fourth part is the look-up table, used as the memory of the FQLC. A predefined system state set and its corresponding best parameterization of the network are stored in the look-up table. This latter is periodically updated during the learning process of the controller.

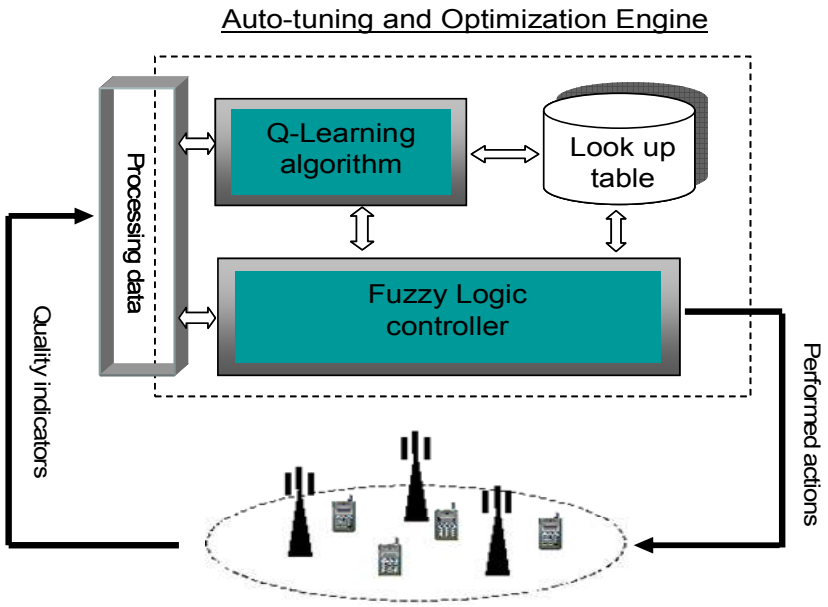


Figure 3.4. Auto-tuning and Optimization Engine

3.3 Fuzzy logic controller

Fuzzy Logic Controller (FLC) is similar to the conventional controller in the sense that it must address the same issues common to any control problem, such as system stability and performances [86]. However there is a fundamental difference between FLC and conventional controller in term of system modelling. Conventional controller starts with a mathematical model of the system and controllers are designed based on the model. FLC, on the other hand, starts with heuristics and human expertise (in terms of fuzzy *if-then* rules) and controllers are designed by synthesizing these rules. This is extremely important in mobile wireless communications because it is almost impossible to obtain an accurate yet simple mathematical model of the system.

FLC has been found particularly well suited for parameter auto-tuning in radio access networks [46] [47]. This is due to its simplicity and ability to convert knowledge from experience into a set of '*if-then*' rules, which mimics the reasoning of the operator (network expert). It is easily understood that the network operator has a critical role in the first step of the FLC design. As a result, an expert system is obtained, which is able to map performance indicator values into control actions, as an operator would do. To solve the problems associated to the separate triggering of rules, a fuzzy inference engine is incorporated into the system. FLC stands then for the process of formulating the mapping from a given input to an output using fuzzy logic theory [86] [87].

Fuzzy logic theory inventors split the fuzzy control process into three phases as illustrated in figure 3.5. The first phase is the fuzzification step and consists of converting the crisp input data to fuzzy data sets. This mapping process involves finding the degree of membership of the crisp input in predefined fuzzy sets. The second phase is the inference process and consists of making

decision from the "*if then*" rules by combining all fuzzy input sets. The last phase is the defuzzification which maps the fuzzy outputs to the final controlled crisp parameters [87].

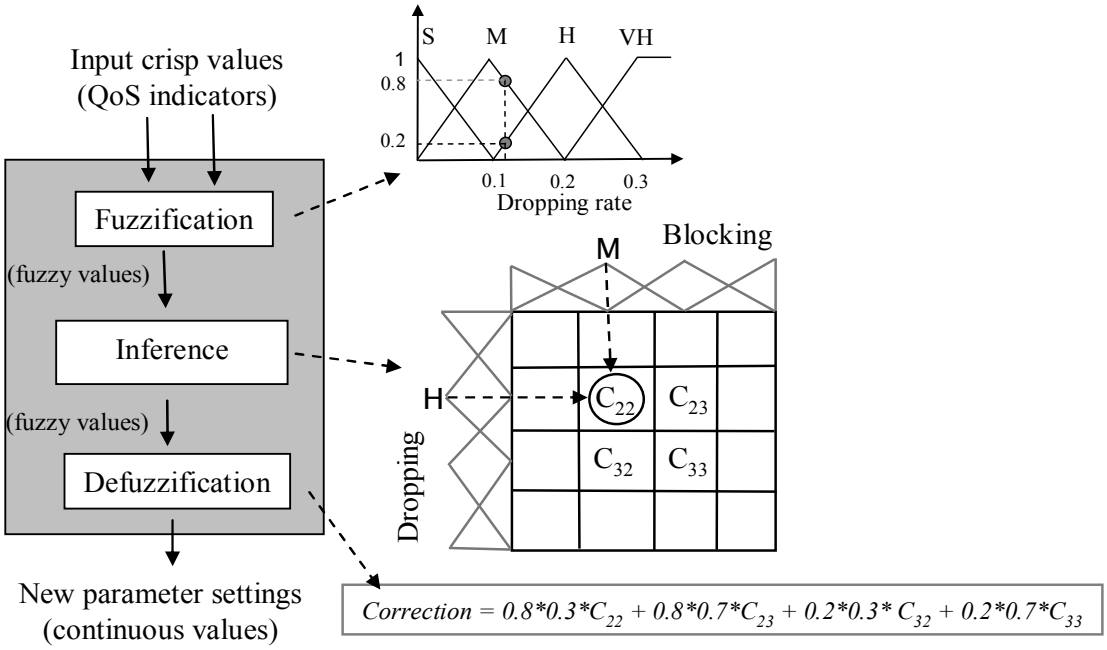


Figure 3.5. The concept of fuzzy logic controller.

In the literature, there are basically two FLC types: *Mamdani* and *Takagi-Sugeno* approach [78] [79]. The main difference between *Mamdani* and *Sugeno* is the output membership function. *Mamdani's* fuzzy inference approach [78] uses a singleton output membership function in the defuzzification process which simplifies the computation. By contrast, *Takagi-Sugeno* type [79] uses a more complex output membership function for extended flexibility at the expense of computation load. In this dissertation, we use FLC based on *Takagi-Sugeno* approach as depicted in figure 3.6.

3.3.1 Mathematical framework of FLC

In the first phase of *Takagi-Sugeno based* FLC, each crisp input variable x_i *(i=1, 2,...n)* is mapped into m_i fuzzy variables (or sets), denoted E_{ij_i} *(j_i=1, 2,...m_i)*. This procedure allows mapping a continuous state space into a discrete space. Note that we can divide each variable to exactly *m* fuzzy sets; in this case $m_i = m, \forall i$. The index *i* can also be omitted from *j* in E_{ij_i}. The mapping between the crisp and the fuzzy variables is made by a membership function $\mu_{ij}(x_i)$ which defines the membership degree of the crisp input x_i with the fuzzy variable E_{ij}. In our work we use the triangular membership function, defined as

$$\mu_{ij}(x_i)=\begin{cases}1-\dfrac{E_{ij}-x_i}{E_{ij}-E_{ij-1}}\ ,\ if\ E_{ij-1}\le x_i\le E_{ij}\\ 1+\dfrac{E_{ij}-x_i}{E_{ij+1}-E_{ij}}\ ,\ if\ E_{ij}\le x_i\le E_{ij+1}\quad ;i\in\{1,2,...n\}\ and\ j\in\{1,2,...m\}\\ 0\ ,\qquad\qquad else\end{cases}\tag{3.2}$$

However, other functions sometimes are employed when higher accuracy or nonlinearity is required such as trapezoid or Gaussian functions.

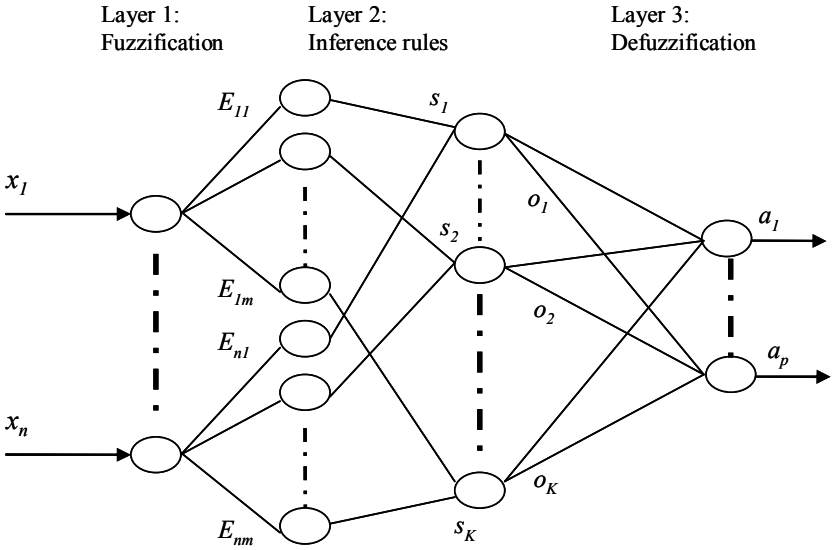


Figure 3.6. Fuzzy logic controller based on Takagi-Sugeno approach.

Recall that in the inference process, "*if-then*" rules are constructed based on network expert. The rules are the realizations of the input variables and the corresponding fuzzy consequences. As shown in figure 3.6, the inference process maps the fuzzy sets into a set of rules.

Definition 3.1

A predicate is called an inference rule if it has the form:

$$If\ x_1\ is\ E_{1j_1}\ and\ x_2\ is\ E_{2j_2}\ ...and\ x_i\ is\ E_{ij_i}\ ...and\ x_n\ is\ E_{nj_n}\ then\ a=o_s;\ j_i\in\{1,2,...m_i\}$$

The statement $x_1\ is\ E_{1j_1}\ and\ x_2\ is\ E_{2j_2}\ ...and\ x_i\ is\ E_{ij_i}\ ...and\ x_n\ is\ E_{nj_n}$ stands for the rule $R_{j_1j_2\cdots j_n}$. Rules form a space set denoted *S* and each rule is denoted by $s\in S$. The cardinal of the space *S* is equal to $K=\prod_{1\le i\le n}m_i$ which is reduced to m^n when $m_i=m,\forall i$.

a is the crisp FLC output variable and o_s is its fuzzy realization in the rule $s=R_{j_1j_2\cdots j_n}$.

The rule $R_{j_1 j_2 \cdots j_n}$ has a membership function deduced from those of E_{ij} [79]:

$$\alpha_s(\mathbf{x}) = \mu_{j_1 j_2 \cdots j_n}(\mathbf{x}) = \prod_{i=1}^{n} \mu_{ij_i}(x_i) \tag{3.3}$$

In the defuzzification phase, the FLC output a corresponding to the input $\mathbf{x} = (x_1 \ldots x_i \ldots x_n)$ is given by the gravity centre of conclusions o_s in each rule weighted by the membership function $\alpha_s(\mathbf{x})$. If the member functions α_s are chosen to satisfy the normalized condition. i.e. $\sum_{s \in S} \alpha_s(\mathbf{x}) = 1$,
the output action a is:

$$a = \sum_{s \in S} \alpha_s(\mathbf{x}) \cdot o_s \tag{3.4}$$

3.3.2 Example: use-case of FLC

Now that we have taken a broad look at how FLC is generally designed, let us see how we can simply use it in a simple example of auto-tuning UMTS admission control parameter. We want to adjust the admission control threshold in an UMTS network according to the observed blocking and dropping rates. What we now aim to do is use blocking and dropping rates as inputs to the FLC to determine the appropriate load target threshold. This is a very simple example, as there are only two input variables and only one output. In the fuzzification phase, we can set up, according to equation (3.2), the membership functions and fuzzy sets as illustrated in figure 3.7. For each input variable, we have three fuzzy sets: low, medium and high.

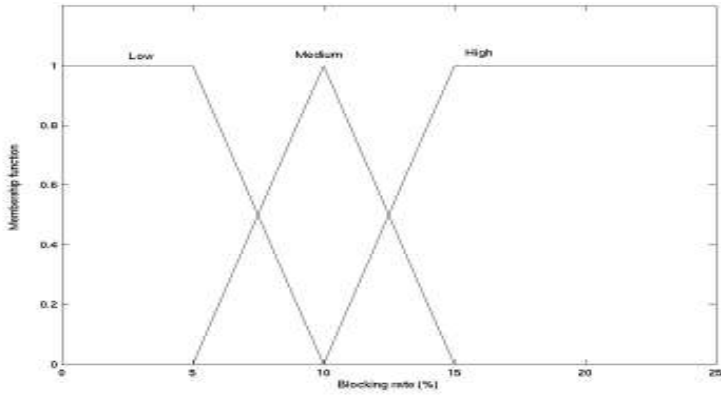


Figure 3.7. Fuzzy sets and menbership function of the dropping and blocking rate.

Blocking / **Dropping**	Low	Medium	High
Low	Do nothing	increase	increase
Medium	Decrease	Do nothing	increase
High	Decrease	decrease	decrease

Table 3.1. Fuzzy rules.

Once the fuzzification step is completed, we move to the construction of the *if-then* rules. Table 3.1 illustrates such rules. Each element of the table represents the degree of modification of the controlled input (load target threshold). As we have *n=2* variables, each of which has *m=3* fuzzy sets, the number of rules is then *K=9*.

The output fuzzy actions in this example can be increase, decrease the admission control threshold, or do nothing. One can set up several more rules to handle more possibilities by adding for example the actions 'more increase', 'more decrease'...

To show the defuzzification process, assumes that at some given time, the FLC receives from the UMTS network a blocking rate equal to 7% and a dropping rate equal to 2.7%. According to the predefined fuzzy sets and membership functions, the FLC interprets the blocking rate as low with 0.6 of truth and medium with 0.4 degree. On the other hand, the dropping rate is medium with 0.3 degree and high with 0.7 degree. If we assign value to the actions such as -0.1 for increase, 0 for do nothing, and +0.1 for decrease, we might have the following output as a correction for the admission control threshold:

Crisp action = (-0.1)*(0.6)*(0.3) + (0)*(0.4)*(0.3) + (-0.1)*(0.6)*(0.7) + (-0.1)*(0.4)*(0.7)

= -0.088

Assume that the recent admission control threshold is 0.75, in the next time the base station will be configured automatically by adjusting its admission control threshold to 0.662.

What we have taken now is a simple example with only nine rules. Of course, many other parameters could be tuned such as the reserved bandwidths mentioned in the previous chapter, maximum allowed bit rate for some admitted mobiles (the case of elastic data traffic), and handover parameters…etc. Therefore, several inputs and outputs parameters might certainly generate so many rules and then the FLC design becomes very complex. In addition, systems with varying conditions can greatly benefit from the adaptation of the controller parameters to deal with such variations. A cellular network is one example of such dynamic systems, since the traffic demand greatly varies in the spatial and temporal domain. Thus, the controller can be adaptive for the sake of efficacy, i.e. rules and parameters in the controller can be automatically and optimally modified when the network experiences substantial changes in order to keep the performance metrics as high as possible [80] [81]. If there are available data about network behaviour and performance, modification of such rules and parameters can be accomplished by self-learning algorithms. This way, auto-tuning can combine two powerful tools: the ability to express simple optimization rules using experience and know-how of operators (FLC) and the ability to apply computationally-intensive self-learning methods to optimise controller parameters based on existing and future network performance data (reinforcement learning).

3.4 Reinforcement learning

3.4.1 General view of machine learning

Machine learning is a field of artificial intelligence (or automatic learning) that deals with the design and development of algorithms and techniques that allow computer or any other agent to

learn and improve its performance based on previous results. The machine can induce patterns, regularities, or rules from past experiences. There are three types of learning: supervised, unsupervised and semi-supervised learning. From a theoretical point of view, supervised and unsupervised learning differ only in the causal structure of the model. In supervised learning, the learning algorithm generates a function that maps inputs to desired outputs. One standard formulation of the supervised learning task is the approximation problem: the learner is required to learn (to approximate) the behaviour of a function which maps inputs into one or several outputs by looking at several input-output examples of the function. For a system using supervised learning, a teacher must help the system in its model construction by defining inputs and providing their labels. In contrast, in unsupervised learning, no teacher helps the learner. Thus, the learner itself must understand and discover relationships between data components. The semi-supervised learning is between supervised and unsupervised learning. The learner in this case is assisted indirectly by a teacher via the reward received for each couple of input-output.

The Reinforcement Learning (RL) as a kind of semi-supervised learning is slightly comparable to the human learning [88]. During the human life, many problems could be met and the resolution of some ones creates human reflex and knowledge about some dangerous situations which require more attentions. RL consists to teach an agent that some decisions are good and some others are bad. Giving a reward when the agent does something good and a punishment when it does something bad allows it to segregate the good from the bad and recognize its harmful decisions. Therefore, the agent could develop better ways to make good decisions.

3.4.2 Mathematical framework of RL

The fundamental purpose of RL is to improve a current agent policy after each interaction with the environment. In this case the reinforcement is local and therefore does not give a complete evaluation of the agent policy. In fact, RL algorithms do not use directly the policy but evaluate the performances of the strategy via a set of value functions resulting from the theory of *Markovian Decision Process* (MDP) [89].

A MDP is a controlled stochastic process similar to Markov chain, except that the transition probability depends on the action taken by the decision maker (agent or controller) at each time step. The MDP is formulated by the quintuple (S, A, T, p, r); where:

- S is a state space that contains a finite number of states,
- A is a finite set of actions. We denote by $A(s) \subset A$ those actions that are available at state s,
- T is the time. T is a sub-set of positive real number,
- p are the transition probabilities between states,
- r is the reinforcement function or reward depending on states and actions.

As shown in figure 3.8, at each time t of T, the agent observes the current state $s \in S$ and performs an action $a \in A$ that shifts the system to another state $s' \in S$ with a probability $p_t(s'/s,a)$. One step later, the agent receives a reward $r_t(s,a) \in IR$.

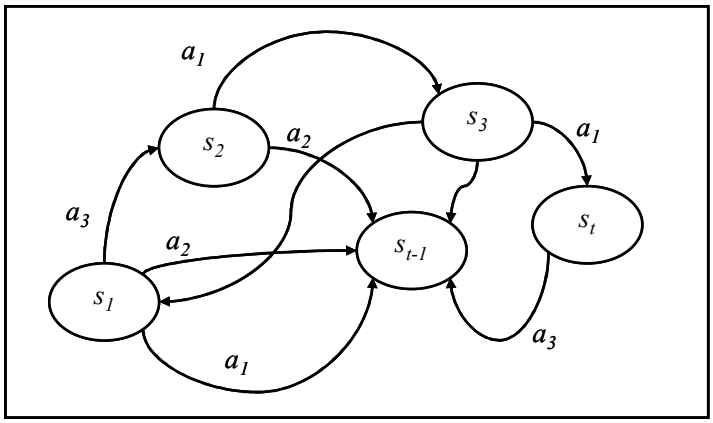


Figure 3.8. Example of Markovian Decision Process.

Definition 3.2: MDP process

Let $h_t = (s_0, a_0, ..., s_{t-1}, a_{t-1}, s_t)$ be the process history observed at time *t*.

The process (*S*, *A*, *T*, *p*, *r*) is called an *MDP* process if the transition probability between s_t and s_{t+1}, given the performed action a_t, depends only on s_t.

$$\forall h_t, a_t, s_{t+1} \quad P(s_{t+1}/h_t, a_t) = P(s_{t+1}/s_t, a_t) = p_t(s_{t+1}/s_t, a_t) \tag{3.5}$$

This does not necessarily mean that the stochastic process (s_t) is itself a *Markovian* process. It depends on the agent policy.

Definition 3.3: Agent policies

The agent policy implements a mapping from state space and action space. RL methods specify how the agent changes its policy as a result of its experience. The set of all policies forms a space, denoted $\Pi = \{\pi : s \in S \rightarrow a = \pi(s) \in A\}$.

For each policy π, we denote $q_\pi(a, h_t)$ the probability that in a given history h_t the action *a* is triggered. According to whether all the history is involved in the agent policy or not, two types of policies are defined:

History-dependent policies: the probability $q_\pi(a, h_t)$ depends on the whole history h_t.

Markov policies: the probability $q_\pi(a, h_t)$ is only a function of s_t and not of the whole history.

Definition 3.4: Goals and rewards

In RL, the purpose of the agent is formalized in terms of a special signal, called the reward and denoted *r* that passes from the environment to the agent. The reward is just a single number whose value varies from step to step. The agent's goal is to maximize the total amount of reward it receives, called goal or return function. This means maximizing not just immediate reward, but cumulative reward in the long run. The use of a reward signal to formalize the idea of a goal is one of the most distinctive features of RL [88].

The most commonly used return function, and the one that will be used throughout this thesis, is the discounted cumulative future reward, expressed as:

$$R = \sum_{t=0}^{+\infty} \gamma^t r_t \tag{3.6}$$

where γ is a parameter between 0 et 1, called discount factor. The infinite sum R has a finite value as long as the reward sequence (r_t) is bounded. The discount factor is used as a measure that indicates the relative importance of future rewards.

Definition 3.5: Value functions

Almost all RL algorithms are based on estimating some *value functions*, functions of states (or of state-action pairs), that estimate how good it is for the agent to be in a given state (or how good it is to perform a given action in a given state). The notion of "*how good*" here is defined in terms of future rewards that can be expected, or, to be precise, in terms of expected return. Of course, the rewards the agent can expect to receive in the future depend on what actions it will take. Accordingly, value functions are defined with respect to particular policies. So for each policy π, the value function V^π is expressed by the expected return function:

$$\forall \pi, \quad V^\pi : s \in S \rightarrow V^\pi(s) \in IR$$

For the discounted cumulative reward of the equation (3.6), the value function is:

$$\forall s \in S, \quad V^\pi(s) = E^\pi[R/s_0 = s] = E^\pi[\sum_{t=0}^{+\infty} \gamma^t r_t / s_0 = s] \tag{3.7}$$

The operator E^π stands for the mathematical expectation given the policy π.

Let Ω be the space of all the functions going from the state space S to the real number IR. We define in this space the norm $\max \forall V \in \Omega, \quad \|V\| = \max_{s \in S} |V(s)|$.

The value function space Ω is also a partially ordered set:

$$\forall U, V \in \Omega \quad U \leq V \Leftrightarrow \forall s \in S \quad U(s) \leq V(s)$$

Proposition 3.1

Let π be a history-dependent policy. For any initial state x, there exists a Markov policy π', such that: $V^{\pi'}(x) = V^\pi(x)$.

The proof of this proposition is given in appendix A.

Remarks

i) From the previous proposition, we deduce that every history-dependent policy can be replaced by a *Markovian* policy having the same value function if the initial state is given. From now on, we use only the markovian policy, unless contrary mentioned.

ii) If the policy is *markovian*, the process (s_t) is itself a *markovian* process with a transition matrix P_π, defined by:

$$\forall s, s' \in S \quad P_{\pi s,s'} = \sum_{a \in A} q_\pi(a,s) p(s'/s,a) \tag{3.8}$$

iii) According to the previous notation, the value function can be expressed as:

$$V^\pi(x) = \sum_{t=0}^{+\infty} E^\pi\left[\gamma^t r_t(s_t, a_t)/s_0 = x\right]$$
$$= \sum_{t=0}^{+\infty} \sum_{s \in S} \sum_{a \in A} \gamma^t r_t(s,a) P^\pi(s_t = s, a_t = a/s_0 = x)$$

Theorem 3.1

Let r_π be the reward vector whose elements are $\sum_{a \in A} q_\pi(a,s) r(s,a)$ and V^π (the same notation of the value function) the value vector whose elements are $V^\pi(s)$. The size of r_π and V^π is equal to the number of states. The matrix expression of the value function V^π is then:

$$V^\pi = (I - \gamma P_\pi)^{-1} r_\pi \tag{3.9}$$

The proof of theorem 3.1 is given as well in appendix A.

Definition 3.6: Optimal value function

A policy π is defined to be better than or equal to a policy π' if its expected return is greater than or equal to that of π' for all states. In other words, $\pi \geq \pi'$ if and only if $V^\pi(s) \geq V^{\pi'}(s)$ for all $s \in S$. There is always at least one policy that is better than or equal to all other policies. This is an optimal policy. Although there may be more than one, we denote all the optimal policies by π^*. They share the same value function, called the optimal value function, denoted V^*, and defined as $V^*(s) = \max_{\pi \in \Pi} V^\pi(s) \quad \forall s \in S$.

The objective of RL is then to find a policy π^* that corresponds to the optimal value function. To prove the existence of the optimal value function, one can use the *Dynamic Programming Operator* (DPO).

Definition 3.7: DPO

The operator DPO, denoted here *L* (to refer to the learning), is a mapping over the space Ω such that

$$\forall V \in \Omega \quad \forall s \in S \quad LV(s) = \max_{a \in A} \left\{ r(s,a) + \gamma \sum_{s' \in S} p(s'/s,a) V(s') \right\} \tag{3.10}$$

With matrix notation, the previous expression is

$$\forall V \in \Omega \quad LV = \max_{\pi \in \Pi} \{ r_\pi + \gamma P_\pi V \} \tag{3.11}$$

The existence and the uniqueness of the optimal value function are given by the following theorem.

Theorem 3.2: Bellman equation

If S and A are finite sets, then V^* is the unique solution of the equation

$$V = LV \tag{3.12}$$

The theorem is proved in appendix A.

To solve the *Bellman* equation, there are two different approaches: the first is the policy iteration and the second is the value iteration.

In the policy iteration, there are two steps: policy evaluation and policy improvement. Each iteration preserves monotonicity in terms of the policy performance. The policy evaluation step obtains V^π for a given policy π by solving the corresponding fixed-point functional equation over all $s \in S$:

$$V^\pi(s) = r(s, \pi(s)) + \gamma \sum_{s' \in S} p(s' / s, \pi(s)) V^\pi(s') \tag{3.13}$$

The policy improvement step takes a given policy π and obtains a new policy π^* that satisfies the condition $\pi^*(s) = \arg\max_{a \in A} \left\{ r(s,a) + \gamma \sum_{s' \in S} p(s' / s, a) V^\pi(s') \right\}$, $\forall s \in S$.

The policy improvement step ensures that the value function of π^* is no worse than that of π.

With respect to the value iteration approach, it is often the most efficient computational technique for finding the optimal value function and its corresponding policy. The principle is to update a given value function by applying the operator L successively. Since L is a contraction, the solution of *Bellman* equation is the limit of the sequence $V_{n+1} = LV_n$, for every starting value function V_0.

The running-time complexity of value iteration is polynomial in $|S|$, $|A|$, $1/(1-\gamma)$; in particular, one iteration is $O(|A||S|^2)$ in the size of the state and action spaces [89].

As observed in the expression of the operator L, the resolution of *Bellman* equation requires the knowledge of the system (i.e. the transition probabilities between states given a performed action, random rewards). However in complex systems, such wireless communication networks, it is impossible to know explicitly the system model which makes hard to solve the *Bellman* equation. This challenge has given rise to a number of approaches intended to result in more tractable computations for estimating the optimal value function and finding optimal or good suboptimal policies. The Q-learning algorithm, introduced by Watkins [92], tackles the inexplicitly system model. It is considered the most practical approach thanks to its simplicity. As the controller,

used in this thesis, is based on fuzzy Q-learning, we explain in the next section the Q-learning algorithm and its fuzzy version.

3.5 Fuzzy Q-learning controller

3.5.1 Q-learning algorithm

Q-learning, perhaps the most well-known example of reinforcement learning, is a stochastic approximation-based solution approach to solving Bellman equation. It is a model-free approach that works for the case in which the transition probabilities and one-stage reward function are unknown. Instead of using only one value function, Q-learning algorithm employs another value function depending on both state and action, called quality function or Q-function. It equals the expression appearing inside the "max" operator of equation (3.10).

$$Q^{\pi}(s,a) = r(s,a) + \gamma \sum_{s' \in S} p(s'/s,a) V(s') \tag{3.14}$$

Using the Q-function, the operator L becomes $LV(s) = \max_{a \in A} \{Q^{\pi}(s,a)\}$, $\forall V \in \Omega$, $s \in S$. Moreover, the value iteration approach can be expressed in terms of two sequences $\{V_t\}$ and $\{Q_t\}$ constructed from the Q-function as

$$V_{t+1}(s) = LV_t(s) = \max_{a \in A} \{Q_t(s,a)\} \tag{3.15}$$

and

$$Q_{t+1}(s,a) = r(s,a) + \gamma \sum_{s' \in S} p(s'/s,a) \max_{a' \in A} \{Q_t(s,a')\} \tag{3.16}$$

The Q-learning algorithm is a stochastic form of the value iteration. From equation (3.16), it is understood that performing a step of value iteration requires knowing the expected reward and the transition probabilities. Although such a step cannot be performed without a model, it is nonetheless possible to estimate the appropriate update. The term

$$r(s,a) + \gamma \sum_{s' \in S} p(s'/s,a) \max_{a' \in A} \{Q_t(s,a')\}$$

is replaced by its simple unbiased estimate $r_t + \gamma \max_{a' \in A} \{Q_t(s',a')\}$ [91]. So, the successor state is an unbiased estimate of the sum and r_t is an unbiased estimate of $r(s,a)$. This reasoning leads to the following relaxation algorithm, where we use $Q_t(s,a)$ to denote the learner's estimate of the Q-function at time *t*.

$$Q_{t+1}(s_t,a_t) = (1-\kappa_t) \times Q_t(s_t,a_t) + \kappa_t \times \left(r_t + \gamma \max_{a' \in A} Q_t(s_{t+1},a')\right) \tag{3.17}$$

The variable κ_t is the learning rate and it equals zero except for the state that is being updated at time *t*.

The proof of the convergence of the Q-learning algorithm is widely studied. It can be found in [91]or [92]. The same way as in theorem 2, the proof of Q-learning convergence is based on the contraction mapping theorem (theorem 3). As shown in [91], the convergence is achieved under the assumption that each state is visited in an infinite number of times and the sequence of the learning rate κ_t satisfies the conditions

$$\sum_{t=0}^{\infty} \kappa_t = \infty \text{ and } \sum_{t=0}^{\infty} \kappa_t^2 < \infty .$$

3.5.2 Adaptation of Q-learning to fuzzy inference system

The convergence of the Q-learning algorithm requires that the state space *S* should be finite. However, in mobile communication, we face continuous states (continuous quality indicators), such cell load or blocking probability, that can not be simple inputs to the MDP Q-learning algorithm. To handle continuous input indicators, a simple interpolation procedure is introduced in the Q-learning algorithm using fuzzy inference system. As explained in section 3.3, the set of continuous input indicators are mapped into a set of rules. Now, instead of applying the learning directly to the input indictor, the agent learns on the rules and fuzzy actions. To do this, a fuzzy quality-value *q* is assigned to each rule and each action.

Unlike simple fuzzy inference system, where only fuzzy action corresponds to a fuzzy rule, in fuzzy Q-learning, each rule $s = R_{j_1 j_2 \cdots j_n}$ has $A(s)$possible competing discrete actions $\{o_k\}$. The predicate, given in definition 3.1, becomes for each rule

If **x** *is* $s = R_{j_1 j_2 \cdots j_n}$ *then* $a = o_1$ with quality $q(s,o_1)$
or o_2 with quality $q(s,o_2)$
……
or $o_{A(s)}$ with quality $q(s, o_{A(s)})$

The agent stores the parameter vector $q(s,o_k)$ associated with each of these state-action couples in the look-up table. These q-values are updated whenever the agent performs an action and the system visits a new crisp state. The value functions of crisp input **x** and crisp action *a*, at time *t*, is calculated as a linear interpolation of the *q*-values:

$$Q_t(\mathbf{x}, a) = \sum_{s \in S} \alpha_s(\mathbf{x}) q_t(s, o_s) \tag{3.18}$$

$$V_t(\mathbf{x}) = \sum_{s \in S} \alpha_s(\mathbf{x}) \max_{o \in A(s)} q_t(s, o) \tag{3.19}$$

The sum is performed over all the rules of the FIS. Here, we use the notation *s* to refer rules because the set of rules forms a discrete state space *S* defined as the same way in the previous

section. Recall that the number of fuzzy states |*S*| is given in section 3.3 by $K = \prod_{1 \le i \le n} m_i$. o_s is the selected fuzzy action in rule (fuzzy state) *s*. Recall that $\alpha_s(\mathbf{x})$ is the membership function (given in equation 3.3) of rule *s* applied to the crisp input **x**.
The update of the q-value is similar to the update of Q-function in simple Q-learning algorithm. So, for fuzzy Q-learning, the iterative equation (3.17) becomes

$$q_{t+1}(s,o) = q_t(s,o) + \alpha_s(\mathbf{x}_t)\kappa_t\left(r_t + \gamma V_t(\mathbf{x}_{t+1}) - Q_t(\mathbf{x}_t, a)\right) \qquad (3.20)$$

Finally, the fuzzy Q-learning algorithm for a stochastic environment (such mobile communication system) is given below.

1. *Initialize Q-look-up table:* $\forall s \in S, o \in A \;\; q(s,o) = 0$;
   Time t=0;
   Repeat:
2. *Receive the crisp system input* $\mathbf{x}_t = (x_1, x_2, ..., x_n)$ *from the system;*
3. *Fuzzification: mapping from* $\mathbf{x}_t$ to fuzzy states $s \in S$ (or rules $R_{j_1 j_2 \cdots j_n}$);
4. *For each rule* $s \in S$ *select an action* o_s *with the EEP policy*
   $o_s = \arg\max_{o \in A(s)} q(s,o)$ *with a probability ε*
   or $o_s = random\{o, o \in A(s)\}$ *with a probability 1-ε*
5. *Calculate the inferred action* a_t *(equation (3.4))*
6. *Calculate its corresponding quality (equation (3.18))*
7. *Execute the action* a_t *that leads the system to the crisp state* $\mathbf{x}_{t+1}$. *The controller receives the reinforcement* r_t.
8. *Calculate the membership functions* $\alpha_s(\mathbf{x}_{t+1})$ *for* $s \in S$ *(equation (3.3))*
9. *Calculate the value of the new state (equation (3.19)):*
10. *Update the elementary quality* $q(s,o)$ *of each rule* s *and action* $o \in A(s)$ *(equation (3.20)):*
11. *Save the elementary quality* $q(s,o)$ *in the Q-look-up table.*
12. *if convergence is obtained then stop the(n) learning process*
13. *t=t+1.*

Figure 3.9. Fuzzy Q-learning algorithm.

In the algorithm, actions are selected during the learning process using an Exploration/ Exploitation Policy (EEP) [53]. This policy, called also ε-greedy [82], allows the controller to exploit its knowledge throughout its learning process. For each rule, the controller chooses the best action with a probability *ε* and a random action with a probability *1-ε.*

In this dissertation, ε is set to 0.80, the discount factor γ is fixed to 0.95. With respect to the leaning rate κ_t, it is equal to $1/(1+v_t(s,o))$ [54], where $v_t(s,o)$ is the total number of times that this fuzzy state-action pair has been visited before the time *t*.

The agent stops completely the learning process if the convergence is reached. In exploitation phase, the agent (or controller) chooses only the fuzzy action that maximizes the q-value in each rule, i.e. $o_s = \arg\max_{o \in A(s)} q(s,o), \forall s \in S$. The used convergence criterion is:

$$\theta = \max_{s \in S, o \in A} \left| q_{t+1}(s,o) - q_t(s,o) \right| \tag{3.21}$$

The convergence is reached when θ becomes very low (lower than $\theta_c = 10^{-4}$ for example).

3.6 Conclusions

In this chapter, the auto-tuning architecture has been given for both online and off-line auto-tuning. The requirements for efficient auto-tuning architecture are highlighted by pointing out the relation between network entities and the auto-tuning engine. An example of signalling messages between the network and the AOE has been described. The AOE has been investigated with a particular focus on fuzzy reinforcement learning mechanisms for optimizing the auto-tuning process. A complete and detailed proof of the Q-Learning algorithm has been presented, and to our knowledge, does not exist in the present form. The Reinforcement Learning with the fuzzy Q-learning implementation has been described in depth for the design task of optimal fuzzy logic controllers.

In chapter 4, we will show how this controller can be implemented in UMTS networks to dynamically adapt resource allocation algorithm and handover parameters. The fuzzy Q-leaning will be also used for adapting inter-system mobility in Chapter 6.

Chap. 4 Application of auto-tuning to the UMTS networks

4.1 Introduction

The complexity of UMTS networks and the permanent traffic variations make dynamic engineering (or self-tuning) of RRM parameters a promising avenue for improving network performance. The main objective of this chapter is to show how techniques presented in the previous chapter can dynamically perform parameters' setting of UMTS network as a function of traffic fluctuation. The optimal setting uses fuzzy Q-learning controller which combines both fuzzy logic controller and Q-learning algorithm. The combination of these two techniques is implemented, as described in chapter 3, to tune exactly two RRM algorithms. The first is the dynamic resource allocation between Real Time (RT) and Non-Real Time (NRT) services. The second case is the automatic setting of soft handover algorithm parameters namely *Hysteresis_event1A* and *Hysteresis_event1B,* according to 3GPP specifications [19] [24]. As presented previously, the Fuzzy Q-Learning Controller (FQLC) needs a vector of quality indicators as controller inputs and delivers corrections for RRM parameters. So, for each case study, we are going to use the set of quality indicators that are well related to the parameters set for auto-tuning. To do that, we measure the correlation between quality indicators and we use only pertinent indicators for the parameters' adaptation.

The structure of this chapter is as follows: Section 2 describes quality indicators candidate for the control process and the correlation between them. The choice of indicators and the correlation analysis are done in the context of UMTS auto-tuning but the same methodology follows in the next chapters. Section 3 treats the case of dynamic resource allocation between RT and NRT service. For this case study, a guard band for RT calls is reserved and dynamically adapted to optimize resource utilization and to achieve optimal tradeoffs between QoS of RT and NRT users. In section 4 soft handover parameters are auto-tuned to improve network performance in terms of call success rate (CSR). We show using simulation results that the auto-tuning of mobility parameters balances the traffic between the network Base Stations (BS) which is the origin of the capacity gains. This parameters' adaptation brings about up to 30 percent of capacity gain. The network performance without auto-tuning concept is utilized as a benchmark to evaluate the added value brought by the auto-tuning process. Concluding remarks end this chapter in section 5.

4.2 Correlation between quality indicators

4.2.1 Presentation of used quality indicators

From operators' point of view, the degree of customers' satisfaction and the well operating of the network are measured by a set of metrics and quality indicators, called also key performance indicators. Metrics are low level measurements carried out on the network interfaces whereas

quality indicators are high level measurements and quantities derived from processing low level measurements. For example, the number of blocked calls is a metric but the blocking rate is a network quality indicator. So, quality indicators are obtained by smoothing or filtering metrics counters.

In this chapter, we focus on the following indicators:

i) Load metrics

In UMTS technology, the load can be evaluated through the total transmitted power in downlink compared to the maximum power. However this metrics is not enough to judge the availability of resources in a cell. We have to take into account also the limitation due to the channel elements, orthogonal codes and the availability of signalling channels. We can also define the load generated by each service class by summing powers allocated to it. NRT load can also be calculated using average throughput or average delays.

In release 4, the exchange of load information on the Iur interface is already standardized between two RNCs and in release 5 this is extended to the Iur-g between RNC and BSC. Besides, in release 5, the RNC has the capability to send and receive cell load information from target/source system through the Iu interface. In release 5, a distinction is made between RT load (conversational and streaming classes) and NRT load (interactive and background classes). However the load is defined as a generic measure. The cell load is assumed to be evaluated in a vendor specific manner by one RNC and only a value varying between 0 and 100% is sent to other RNCs.

ii) Call Setup Success Ratio (CSSR) versus Call Blocking Rate (CBR)

This indicator is defined as the ratio of the number of successful call setups divided by the number of call setup attempts during a certain period of time. Unsuccessful call setup attempts can be caused by the non-availability of radio or network resources, the failure of network elements or by the lack of coverage. This indicator can be defined for one cell (i.e. defined based on the number of calls passed and accepted in the cell) or can be defined as seen by the network (i.e. defined based on a large number of calls passed by any user under the network). Usually, connection success ratios in the network should be higher than 98%. The percentage of users blocked while requesting access to the network is called Call Blocking Rate (CBR). The sum of CSSR and CBR is 1.

iii) Call Dropping Ratio (CDR) versus Call Success Rate (CSR)

CDR is the ratio of dropped calls during ongoing conversations divided by the number of successfully started calls during a defined period.

CSR is defined as the percentage of calls that are admitted to the network and normally end their communications without any undesirable interruption.

$$CSR = CSSR \cdot (1 - CDR) \tag{4.1}$$

Both indicators are influenced by changing network conditions (e.g. radio or load conditions), equipment malfunction or the mobility of the user.

iv) Average throughput

This indicator is the average of user bit rates. It is calculated based on session calls (NRT calls) connected to a cell or to the whole network depending on the need for global indictor or local indicator per a cell. This indicator is influenced by the available bit rate and the round-trip time (RTT), while these again depend on the bearer parameters, on the load change and radio conditions as well as delays in sub-networks such as the core network, corporate intranets and/or the public internet.

v) Average user satisfaction

The satisfaction of a NRT mobile is the ratio between the allocated bit rate and the requested bit rate. For each cell, the average satisfaction, denoted S_{NRT}, is calculated as the average of all NRT user (connected to the cell) satisfactions [10]:

$$S_{NRT} = \frac{1}{M_{NRT}} \sum_{m \in BS} \frac{R_m}{R_m^{\max}} \tag{4.2}$$

where, M_{NRT} is the number of NRT mobiles connected to the cell. R_m is the perceived bit rate of the mobile m and $R_m^{\max}$ is its requested bit-rate.

In this study, for each mentioned indicator, we store the instantaneous values (metrics) and the filtered values (quality indicators). However, to avoid using highly oscillating values in the input of the auto-tuning controller, we use only filtered indicators. Recall that, the filtering operator of a quality indicator $x(t)$ over a filtering period T_f is defined by:

$$Fil(x) = \frac{1}{T_f} \sum_{i=0}^{T_f - 1} x(t-i) \tag{4.3}$$

By using smooth indicators, we want to make sure that the parameters' setting performed by the FQLC is based on underlying trends and not instantaneous changes.

4.2.2 Correlation between quality indicators

The objective of calculating correlation between quality indicators is to avoid using, at the controller inputs, quality indicators that are auto-correlated. In fact, using correlated indicators distorts the results and increases the complexity of getting dynamic optimal solutions at the end of the learning process.

Let X and Y be two stationary signals representing 2 quality indicators. The covariance function between X and Y is written as:

$$C_{xy} = \frac{1}{N}\sum_{n=1}^{N} x(t_n)y(t_n) - E[x(t)]E[y(t)] \tag{4.4}$$

The correlation between the variable X and Y is given by:

$$\rho(x,y) = \frac{Cxy}{\sqrt{CxxCyy}} \tag{4.5}$$

From the last formulae, it turns out that ρ is always between -1.0 and +1.0. If the correlation is near 1 or -1, the two indicator variables X and Y are highly correlated and when ρ tends to 0, the two variables are practically uncorrelated.

In table 4.1, we present an example of correlation between different global quality indicators. The indicator values are obtained by simulating a UMTS network for a long period. To get different values of the indicators, generated traffic, in the network, changes in time and in space from a simulation to another.

	CSR for RT	CSR for NRT	Throughput per mobile	Satisfaction
CSR for RT	1	0,63	-0,5	-0,39
CSR for NRT	0,63	1	-0,55	-0,44
Throughput per mobile	-0,5	-0,55	1	1
Satisfaction	-0,5	-0,55	1	1

Table 4.1. Correlation between quality indicators.

We note that according to table 4.1, RT and NRT call success rates are not very correlated with each other and with the other quality indicators. However, Throughput per mobile and average user satisfaction are highly correlated with each other. Consequently, for the controller inputs, we retain RT and NRT CSR. As a third indicator, we take the user satisfaction since it is more correlated to the throughput and implicitly to the qualitative user satisfaction.

4.3 Auto-tuning of resource allocation in UMTS

In UMTS systems, each BS has a limited soft capacity, defined as the total available transmission power in downlink and an acceptable threshold of interference in uplink. The capacity has to be

shared efficiently among users with different services and quality requirements. Generally, services can be time sensitive (voice, video streaming) or time tolerant (background applications). The former are RT services and the latter are NRT applications. In a traffic environment characterized by high variability and dynamicity, RT and NRT services can be in continuous competition. Hence, rules and methods have to be defined to establish a fair repartition of the radio resources between these service users.

The air interface capacity may be divided into two parts: the first is shared between RT and NRT services and the second is reserved for RT services as a guard band, noted here as X_{RT}, to prioritize RT users. Based on some quality indicators for both services, the Fuzzy Q-Learning controller dynamically regulates X_{RT} reserved for the RT. The objective of this section is then to evaluate the gain that can be achieved by auto-adapting the guard band according to the quality indicators of each service class.

4.3.1 Admission control strategy

As shown in figure 4.1, the downlink capacity in each BS, defined as the maximum transmitted power, is distributed among control channels and traffic channels. To avoid system saturation and excessive call dropping, we keep 5% of the power margin below the total BS transmitted power. 10% of capacity can be also reserved in order to minimize dropping of handoff calls since call dropping is perceived worse than call blocking. The rest of the capacity is split into two parts: The first, denoted X_{mix}, is shared between both RT and NRT services and the second, X_{RT}, is reserved as a guard capacity for RT calls and common channels.

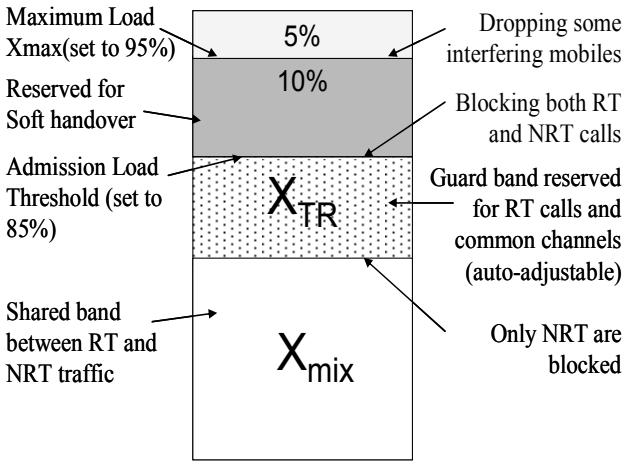


Figure 4.1. Capacity model for a UMTS base station [10].

The admission control strategy, used in this study, is based on a power threshold as described presently. Let *Load* be the radio load of a given BS. It is defined as the ratio between the instantaneous power and the total available BS power. The admission criteria of a call are described below [10]:

- If $Load < X_{mix}$, both RT and NRT calls are accepted.

- If $X_{mix} \leq Load < X_{mix} + X_{RT}$, RT calls are admitted into the network but NRT calls are blocked.
- If $X_{mix} + X_{RT} < Load \leq 0.95$, all new calls are blocked. Only link addition due to soft handover is authorized.
- When *Load* exceeds 0.95, the system drops out some mobiles, i.e. mobiles with high power consumption.

Since traffic distribution is different for each service and is space-time variable, setting a fixed X_{RT} leads to an inefficient use of available capacity. For instance, a high value of X_{RT} causes a quality degradation of *NRT* services when the traffic of the latter is high and the one of the RT is low. The need for X_{RT} auto-tuning can be illustrated by figure. 4.2. This figure is obtained by simulating a UMTS network composed of 24 cells. Traffic of RT and NRT services is generated independently and non-uniformly in the network map. We compare the evolution of the quality of *RT* and *NRT* services when varying the guard band X_{RT} for two different traffic scenarios. The used quality indicator is the call success rate, denoted CSR_{RT} and CSR_{NRT} for respectively *RT* and *NRT* service.

When the traffic arrival rates of RT and NRT equal 3 and 5 mobiles/s respectively, the best configuration of the guard band is about 25% of the available power if no priority is given to the RT service. If this parameter is fixed then, for the condition that RT traffic rate equals 5 mobiles/s and NRT traffic rate equals 2 mobiles/s, the quality of RT service is degraded compared to the NRT service. In the last traffic condition, the relative improvement in RT quality, when going from the configuration of 25% of the guard band to around 50%, comes at a price of a small NRT quality degradation. Therefore, the dynamic trade-off between RT and NRT quality can be well achieved by optimally auto-tuning the guard band.

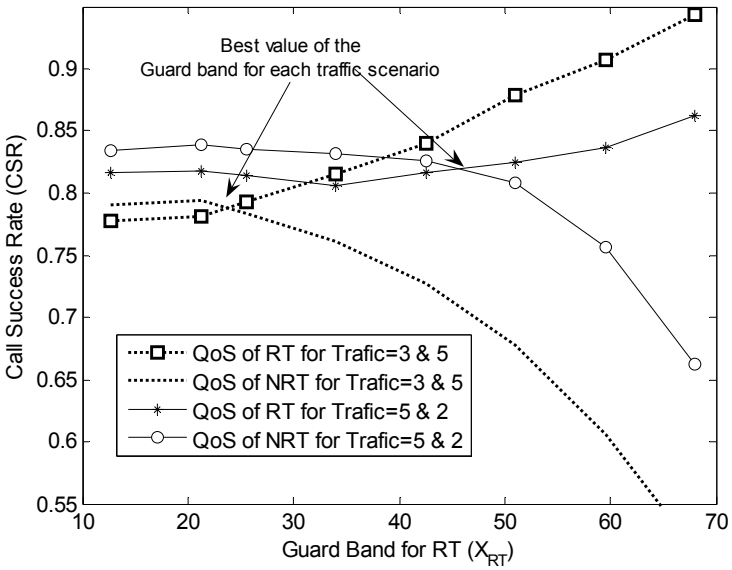


Figure 4.2. Call success rate of each service as a function of the RT guard band for two traffic situations: (1) RT=3 and NRT=5; and (2) RT=5 and NRT=2.

4.3.2 Quality indicators, actions and reinforcement function

The concept of using FQLC to adapt the guard band in each BS is depicted in figure. 4.3 Since each BS is responsible for managing its resources between different traffic classes independently from other cells, we assume that the controller is logically implemented in each BS. However the controller should be physically implemented in the RNC because the resource allocation algorithm is installed in the RNC.

From the previous chapter, recall that at each time step the controller observes the current BS state $s(t)$ (defined as a vector of quality indicators), and performs an action $a(t)$ (changing the guard band value) that shifts the network to another state. One step later, the controller receives from the network a reward signal or reinforcement function *r,*

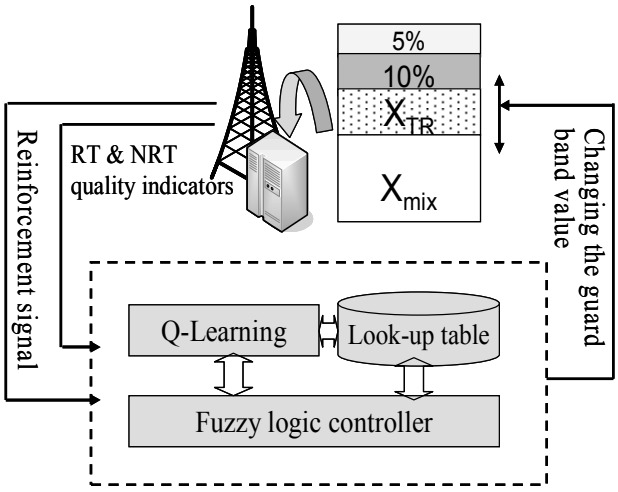


Figure 4.3. Fuzzy Q-learning controller for auto-tuning the RT guard band in each BS.

In this case study, the quality indicator vector that we have used as an input to the controller is $\boldsymbol{s}$= (CSR_{TR}, CSR_{NTR}, S_{NTR}). This choice is based on the correlation table given in the previous section. The average degree of satisfaction of mobiles in the BS S_{NTR} is more pertinent than throughput because it reflects the user satisfaction and it is coherent with the vector state. As can be observed, all quality indicators are variable in [0,1].

The action *a* of the controller modifies the guard band value according to the received quality indicators *s* (or system state). The controller can increase the guard band by 0.05, keep it constant or decrease it by 0.05. The possible actions in each fuzzy state belong then to the set {-0.05, 0, +0.05}. The guard band value must be bounded by fixed lower and upper bounds (set here to 20% and 50% respectively).

To guide the controller in finding a best trade-off between RT and NRT services, we use a reinforcement function that prioritizes the RT traffic in a highly loaded BS and rewards with fair consideration in normal condition

$$r(t) = \alpha CSR_{TR} + \beta CSR_{NTR} + \omega S_{NTR} \tag{4.6}$$

(α,β,ω) is the weighting vector that gives the desired importance to each quality indicator. In our study, we change this vector according to the traffic conditions as follows:

- If ($CSR_{RT} \leq 0.85$) then (α,β,ω)=(1,0,0);
- If ($0.85 < CSR_{RT} \leq 0.95$) then
- If ($S_{NRT} \leq 0.70$) then (α,β,ω)=(0.5,0,0.5);
    - Else if ($CSR_{NRT} \leq 0.9$) (α,β,ω)=(0.5,0.5,0);
    - Else (α,β,ω)=(1,0,0);
- If ($CSR_{RT} > 0.95$) then (α,β,ω)=(0,2/3,1/3).

It is noted that other choices for the weighting vector will lead to different optimal compromises between RT and NRT services.

4.3.3 Performance evaluation

To evaluate the performance of the proposed auto-tuning method, a Semi Dynamic Simulator (SDS) has been utilized. The SDS performs correlated snapshots to account for the time evolution of the network. After each time step that can typically vary from a tenth to a couple of seconds, the new mobile positions and the powers transmitted from/to the mobile are computed. The simulator is developed in France Telecom R&D. A compendium of the simulator is presented in appendix C.

Simulations have been carried out on a UMTS network composed of 24 sectors in a dense urban environment with high traffic level. Two services have been used in the simulations. The first is voice (RT) service, generated with a Poisson traffic model and having exponentially distributed communication duration with average of 100s. The second service is the FTP (NRT) service. An FTP call is generated by a Poisson process and its communication duration depends on the traffic condition. We model a buffer for the FTP service with unlimited capacity. The buffer is cleared whenever the system allocates the necessary bit rate. The FTP sources are modelled by ON-OFF events with exponentially distributed durations (an 'ON' event means that the FTP source is activated and vice versa for an 'OFF' event) [93]. The average duration of the 'ON' state is set to 100s and 1s for the 'OFF' state. FTP service remains activated until the file is completely downloaded. In the 'ON' state, the volume of the transmitted data follows a *Pareto* distribution [93]. Its mean length and its median variation are respectively set to 10 and 1.2 Kbytes.

The mobile can randomly change its direction within a limited angular interval. At the border of the area, the user does not leave the network but is reflected back. FTP users are pedestrians (3 km/h) or immobile whereas for voice users, the average speed is set to 20 km/h.

A comparison is made between the proposed dynamic version (adapted guard band) and the classic version (with fixed guard band) by varying the call arrival rate of the RT and NRT traffics. We keep the sum of their call arrival rates constant.

In figure 4.4, we present the global CSR (measured in the whole network) of each service versus the RT arrival rate for a high traffic condition (the sum of the RT and NRT arrival rate equals 8

mobiles/s). We observe that when the RT traffic is low (between 0 and 2 mobiles/s), setting a fixed guard band leads to a resource waste, especially for a high guard band value. With this proposed guard band adaptation, the gain of NRT capacity is high compared to a small degradation of RT service when the guard band is set to 45% (0.78 of NRT CSR with the guard band adaptation versus 0.68 of NRT CSR for the case where the guard band is fixed to 45%). When the RT traffic increases (between 3 and 5), dynamic adaptation gives slightly better CSR for RT service than the 25% fixed guard band. This can be explained by the choice of reinforcement function which prioritizes RT service over the NRT one, when both services coexist in the network with equal proportions. For high RT traffic, the proposed algorithm does not do anything since the RT service occupies all the resources.

Figure 4.5 and figure 4.6 show the distribution of CSR between BSs for each service in the case of traffic arrival rate equals to 2 mobiles/s for RT and 6 mobiles/s for NRT. As expected, when a BS is highly loaded, and the quality of NRT is mediocre compared with the RT one, the FQLC tends to degrade slightly the RT service quality by decreasing the guard band value and enhancing the quality of the NRT service.

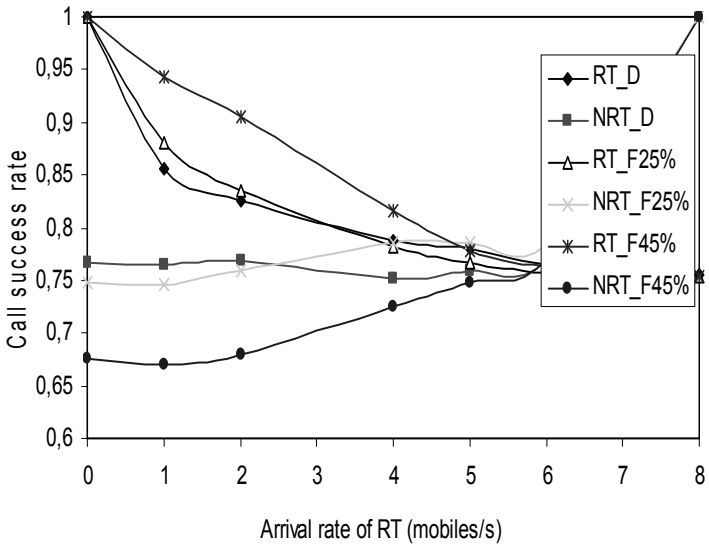


Figure 4.4. Call succes rate as a function of RT call arrival rate (RT_D means RT CSR in the dynamic version, RT_F25% means RT CSR for the fixed guard band of 25%).

As depicted in figure 4.5, 19 sectors have a RT CSR higher than 0.9 for the case with a guard band fixed to 45%, whereas for the case with a guard band fixed to 25% only 9 cells achieve the quality requirements. Dynamic adaptation increases the number of cells from 9 to 11 with a CSR higher than 0.9. As for the NRT quality, we observe that only 4 cells have a NRT CSR higher than 0.85 for the guard band fixed to 45%. Using dynamic optimization, we can achieve up to 10 cells having a NRT CSR higher than 0.85.

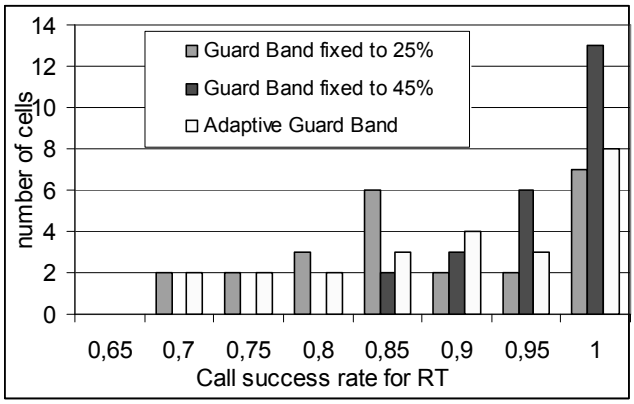


Figure 4.5. Histograms of RT CSR for all BSs with traffic arrival rate of RT=2 and NRT=6 mobiles/s.

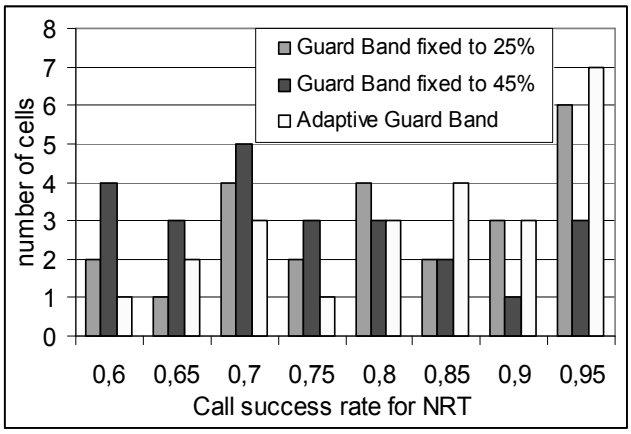


Figure 4.6. Histograms of NRT CSR for all BSs with traffic arrival rate RT=2 and NRT=6 mobiles/s.

4.4 Auto-tuning of UMTS soft handover parameters

Recall that, in UMTS system, each cell covers a geographical area and serves a limited number of mobiles in each service class with a target quality of service (E_b/N_0, *CSSR*, *CSR*,...). Continuous coverage over more than two BS-service areas is achieved by Soft HandOver (SHO) algorithm, called also macro diversity, which is the seamless transfer of a call from one BS to another. This inter-cellular call transfer is controlled by a set of parameters namely *active set size* (the maximum number of cells serving simultaneously the mobile), *Hysteresis_event1A* (for the addition of a link in the active set), *Hysteresis_event1B* (for the removal of a link from the active set) and *Hysteresis_event1C* (for the substitution of a link in the active set) [19]. A mobile, having more than two links in the active set, is said to be in SHO situation. Note that the parameters *Hysteresis_event1A, Hysteresis_event1B* and *Hysteresis_event1C* are called also respectively Addition Window (AddWin), Drop Window (DropWin) and Replacement Window (RepWin).

Each BS has its own parameters but the algorithm is implemented in the RNC since it has more visibility on the state of other BSs. A uniform setting of all BSs leads certainly to a sub-optimal parameterization. Each BS should be optimally parameterized with respect to the other BSs. SHO hysteresis parameters of one BS strongly impact its own radio load and that of its neighbours. Consequently, these parameters influence the downlink capacity and the quality of service of the network. For instance, by increasing the hysteresis parameter values, the downlink and the uplink loads increases and decreases respectively. However, a low value of these parameters may increase the risk of call dropping and may cause coverage holes [4].

This section presents the auto-tuning process of SHO parameters and the corresponding performance gain achieved. The auto-tuning is performed in each BS according to its measured downlink load as well as the load measured in its neighbouring cells. Like the previous case study, the auto-tuning process is based on a fuzzy Q-learning controller. Here, we improve the entire network quality as well as the quality of each individual BS.

4.4.1 SHO algorithm

In UMTS FDD mode, the mobile performs periodically measurements of the quality of its links presented in the active set as well as the quality of signals coming from the declared neighbouring cells of its best serving cell. The mobile compares the measurement results with SHO thresholds provided by the RNC, and sends a measurement report back to the RNC when the reporting criteria are fulfilled and the final decision for handover is made. Based on the measurement reports received from the mobile (either periodic or triggered by certain events), the RNC orders the mobile to add/remove cells to/from its active set. This procedure is called Active Set Update (ASU). SHO thresholds should be set to avoid excessive ASU.

The decision for an ASU is mainly based on the measurement performed on the Common Pilot Channel (CPICH). The quantities that can be measured by the mobile from the CPICH are as follows [19]:

- Received Signal Code Power (RSCP), which is the received power on one code after despreading, defined on the pilot symbols.
- Received Signal Strength Indicator (RSSI), which is the wideband received power within the channel bandwidth.
- Ec/No, representing the RSCP divided by the total received power in the channel bandwidth, i.e. RSCP/RSSI.

According to figure 4.7, The SHO procedures are performed essentially in three events, depending on the mentioned SHO parameters [24]. The three events are the addition of a new link or *Event1A*, the removal of an existing link, or *Event1B*, and the replacement of an existing link, denoted as *Event1C*.

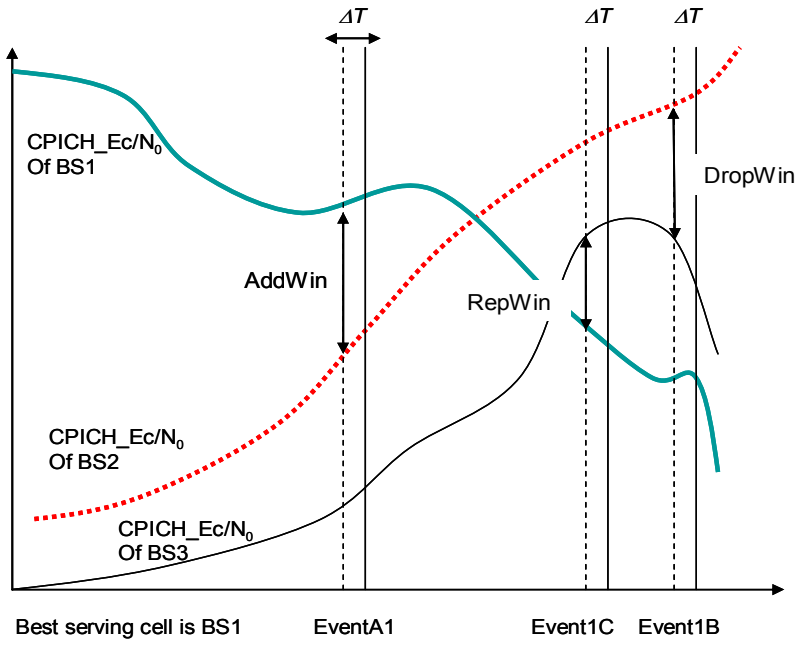


Figure 4.7. UMTS SHO algorithm (event 1A, 1B and 1C).

In the *Event1A* procedure, a BS is added to the active set if the signal from that BS is higher than that of the best serving BS minus the hysteresis window *AddWin* during a *time to trigger* period ΔT [6]:

$$\left(Ec/Io\right)_{CPICH}^{Best\ BS} - \left(\left(Ec/Io\right)_{CPICH}^{BS} + CIO^{BS}\right) \leq AddWin \tag{4.7}$$

CIO^{BS}, or *Cell Individual Offset* of the candidate BS, is an optional offset that can be added to the signal of a potential new link to favour the entry of this station to the active set.

With respect to the *Event1B*, a BS is removed from the active set if the corresponding signal is smaller than that of the best BS minus the hysteresis window *DropWin* during a period ΔT :

$$\left(Ec/Io\right)_{CPICH}^{Best\ BS} - \left(\left(Ec/Io\right)_{CPICH}^{BS} + CIO^{BS}\right) \geq DropWin \tag{4.8}$$

In the *Event1C*, a BS in the active set (superscript *In AS* in (4.9)) is replaced by a new BS if the corresponding signal from the new BS is bigger than that of a BS in the active set plus a hysteresis window *RepWin* during a period ΔT :

$$\left(\left(Ec/Io\right)_{CPICH}^{BS} + CIO^{BS}\right) - \left(Ec/Io\right)_{CPICH}^{In\ AS} \geq \mathrm{Re}\, pWin \tag{4.9}$$

By combining the received signals from the BSs of the active set using the maximum ratio combining mechanism [6], the radio link quality in downlink is improved, the coverage in the cell border is increased, and consequently the uplink capacity increases too. However, when the

number of links per mobile in one BS exceeds a certain threshold, the SHO mechanism becomes a handicap for the downlink capacity because each user consumes much more power than with only one link [94]. So in very loaded network, the SHO parameters should be set with a special care. The perfect setting is to adapt them to the load of each cell, so load balancing between cells can be reached and downlink capacity is optimized [4], [11].

4.4.2 FQLC-based auto-tuning of SHO parameters

The same concept of controller as used in the previous use-case and described in chapter 3 is utilised to tune SHO parameters. As mentioned previously, the change of the *AddWin* or the *DropWin* directly impacts the load of the best serving cell and the load of its neighbouring cell list. So, to achieve best performances from the adaptation of SHO parameters, we should adapt them according to the load of local cell (best serving cell) and to the load of each neighbouring cell. However using a large number of quality indicators (the load of all cells) in the controller input considerably increase the complexity of the learning process in finding the optimal controller. To avoid using all the neighbouring cells' loads, we model all the neighbouring cells of the best serving one as one equivalent neighbouring cell having a fictive load. The fictive load of a cell *k* is defined as the weighted average of the loads in the neighbouring cells.

$$\chi_f = \sum_{i \in NS(k)} \chi_i \omega_{k,i} \tag{4.10}$$

NS(*k*) is the cell-neighbouring set of cell *k*, and $\omega_{k,i}$ is a weighting coefficient defined as the flow of mobiles between cell *k* and cell *i* normalized with the total mobile flow involving the cell *k*. Of course, the load of each cell is filtered using filtering operator given in equation (4.3).

Using this modeling simplifies the system state or the controller input to $s = (\chi, \chi_f)$. This state is evaluated at each BS in each control loop of the controller. In order to keep the system as stable as we can, the adaptation of the parameters *AddWin* and *DropWin* is performed dependently: during the learning and the control process, the margin between both parameters is kept constant (=2dB).

The system quality that we want to improve here is the combination of both system blocking (*CBR*) and dropping rates (*CDR*). Then, the reinforcement function should contain these indicators or some others related to them. The following reinforcement function satisfies the desired criteria:

$$r = (CBR^* - CBR) + \beta \cdot (CDR^* - CDR) \tag{4.11}$$

where CBR^* and CDR^* are respectively the operator target blocking rate and the target dropping rate. β is the mixing factor between the dropping and blocking rate and depends on the operator preference. The dropping of a call is considered more penalizing than its blocking, and consequently β should be bigger than 1. In our study β equals 4 and the target blocking and dropping are set respectively to 0.05 and 0.01.

This reinforcement strategy gives the controller a punishment ($r \leq 0$) if the network quality is below the operator target and a reward ($r \geq 0$) otherwise. This policy allows the controller to carry out an action that decreases the combined blocking and dropping rate.

4.4.3 Performance evaluation

To evaluate the performance of the proposed auto-tuning method, the same simulator as in the previous use-case is used. The fuzzy Q-learning controller has been applied to auto-tune the SHO parameters of an UMTS network with 32 sectors in a dense urban environment. The studied network is extracted from a real situation where the propagation is calibrated according to a professional model that accounts for clutter effects. Each cell is surrounded by a set of neighbouring cells constructed dynamically by the simulator. To model the traffic and mobility in the system, we use the following assumptions:

- Call requests are generated according to a Poisson process with a rate λ that varies between 2 to 13 call requests per second in the present simulation. During one simulation, the traffic is stationary and then λ is kept constant. To model the non-uniformity of traffic, the generated call appears in the network area according to a pre-prepared traffic map. The communication time of each call is exponentially-distributed with a mean equal to 100 s.
- Recall that the user mobility is based on a two-dimensional semi-random walk model. 80% of the generated mobiles are in indoor situation and their speed is set to zero. The speed of the rest of the mobiles is set to 60 km/h.
- The maximum transmit power of each base station is set to 20 Watt. 20% of the power is assigned to the common channel including the common pilot channel (CPICH). 65% of the power is assigned to traffic channels. The admission control threshold is then set to 85% (20% plus 65%). When the cell load reaches 85%, requested calls are blocked. The considered soft handover criterion is the CPICH_Ec/. In a classis UMTS network without any control, *AddWin* and *DropWin* are set to 4 and 6 dB respectively. In the present method, these parameters are controlled according to system reactivity equal to 1/50 s^{-1} (i.e. the period of parameter regulation is set to 50 seconds) unless contrary indication.

We first illustrate the convergence of the proposed auto-tuning algorithm and the evolution of the state-action quality in the learning process. Once the algorithm converges and the stability of state-action quality is reached, we stop the learning process, as described in chapter 3, and we exploit the obtained optimal fuzzy logic controller. Next, we show how the online controller improves the network quality.

Figure 4.8 shows the convergence behaviour of the learning process. In each time step, we take the maximum variation of the quality (given in equation 3.19) of all states and actions $Q(s,a)$. As the UMTS network is a stochastic system, the maximum variation of the quality behaves as a random variable that tends to a zero for a long period of learning. By investigating the mathematical expectation of this random variable, we notice the convergence of the algorithm for

time superior to 400000 seconds, equivalent to 4 days and 7.5 hours of learning in a real network. This learning process is performed by computer simulation for around 20 minutes.

The convergence is clear in figure 4.9 which shows the time-evolution of the quality of some rules (states) and actions. Above 400000 seconds, all the state-action qualities become nearly constant. The learning process goes through three phases. At the beginning of learning, the controller ignores the best action in each rule. It seems that the quality is better because it reaches for short time a maximum situation which is a local optimum. Next and rapidly, the quality goes down since the local optimum is not good in long-term cumulative revenue (reward). Finally, the fuzzy Q-learning improves its behaviour and the states-action quality goes up again and reaches an optimal stable situation. Once the convergence of the algorithm is reached, we assess the impact of the obtained controller on the global and local system quality.

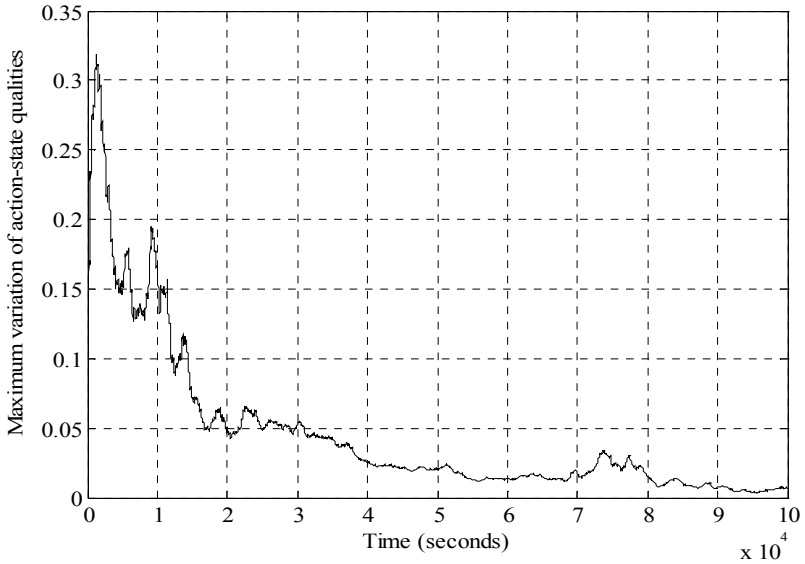


Figure 4.8. Evolution of the convergence criteria of the fuzzy-Q-learning controller.

Figure 4.10 plots the global CSR in the entire network as a function of the call arrival rate for the dynamically-controlled network compared to a classical network with fixed parameters. We observe that the proposed method significantly improves the system capacity for each traffic level. Although, the learning is performed for specific level of traffic, the obtained controller goes on improving network performances in other traffic situations. We can observe that for 95% of CSR (the conventional operator quality target) the controlled network serves 7.85 mobiles per second, whereas the classical network serves only 6 mobiles per second. So, the capacity improvement at 95% of CSR is equal to 30%.

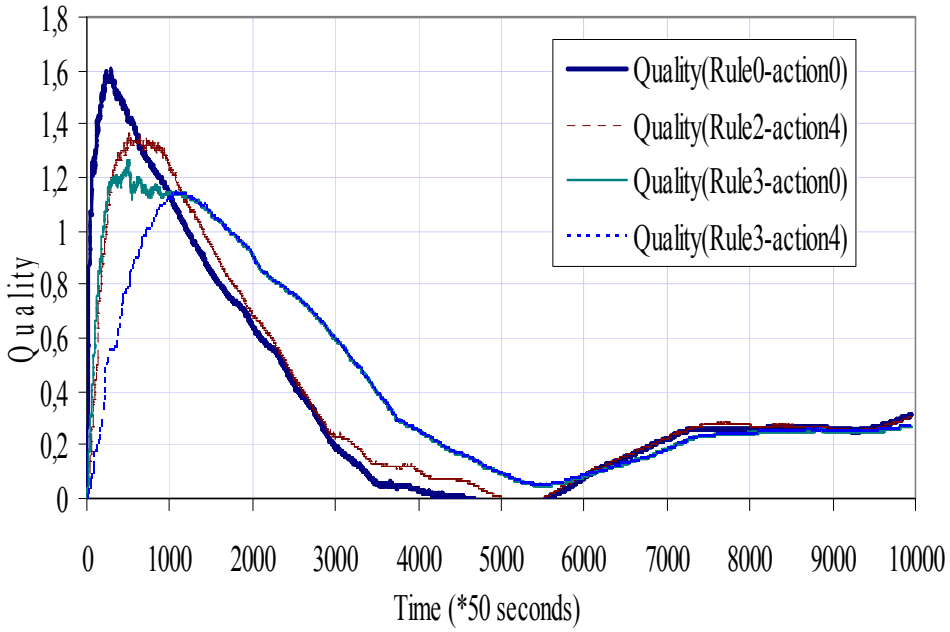


Figure 4.9. Evolution of the quality of rule-action pairs.

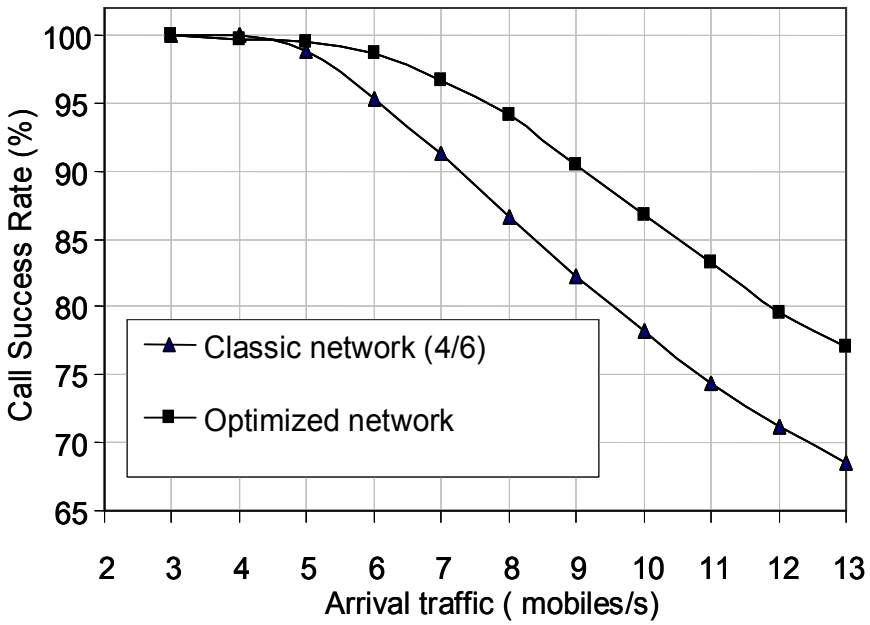


Figure 4.10. Call succes rate versus incomming tarffic for the optimized network with autonomic management compared to a classical network.

To further assess improvement in quality of service at each cell, we show, in figure 4.11, the cumulative distribution function of the CSR in each cell. As expected, the CSR has been also improved in the cells as in the entire network. For the controlled network, 75.8% of cells have a CSR superior to 95%. For the classical network, only 72.5% of cells have a CSR superior to 95%.

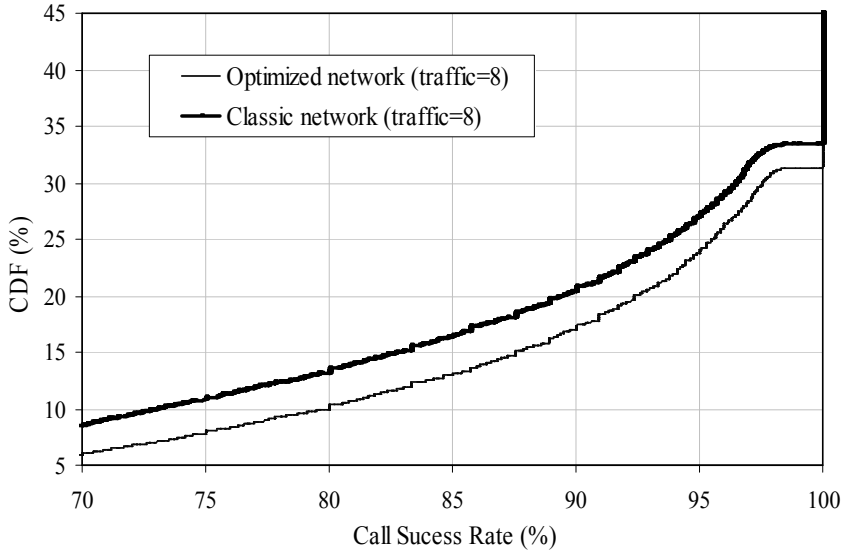


Figure 4.11. Cumulative distribution function of the cell call succes rate in the optimized network compared to the classical network.

Figure 4.12 shows the distribution of cell load for the optimized network compared to a network with fixed configuration. As expected, controlling dynamically and optimally SHO parameters can lead to a better load distribution between the cells since it allows a cell with high traffic to give in certain links to a neighbouring cell with low traffic. As can be seen in figure 4.12, the number of cells with medium load is higher for the optimized network than for the classic one. Also, the auto-tuning reduces the number of cells with very high and very low loads namely it has improved traffic balancing in the network.

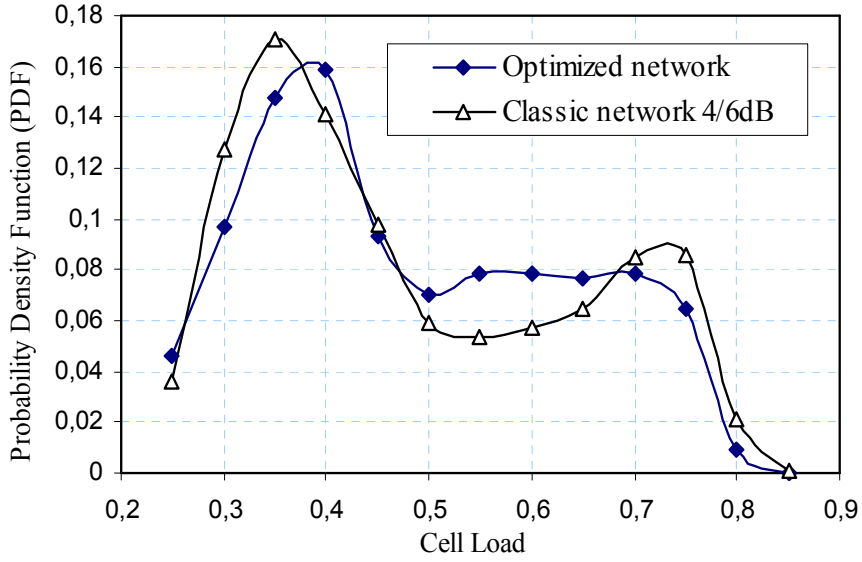


Figure 4.12. Distribution of cell load for the optimized network compared to a classic network with fixed configuration.

Figure 4.13 illustrate the percentage of mobiles in SHO situation as a function of traffic arrival rate for the network with auto-tuning compared to the case without auto-tuning. From the figure, we observe that for high traffic level, the auto-tuning process decreases faster the percentage of mobiles in SHO with respect to the network with fixed *AddWin* and *DropWin* parameters set respectively to 4 and 6 dB. The auto-tuning process tries to alleviate overloaded BSs which suffer from poor QoS and allows the network to provide better capacity. For low level traffic, this tendency is reversed, namely the auto-tuned network allows more mobiles to be in SHO. So the auto-tuning allows increasing the coverage when there is no limitation in capacity. However, when the system becomes loaded, the auto-tuning slightly decreases the coverage and highly increases the capacity compared to a classic network.

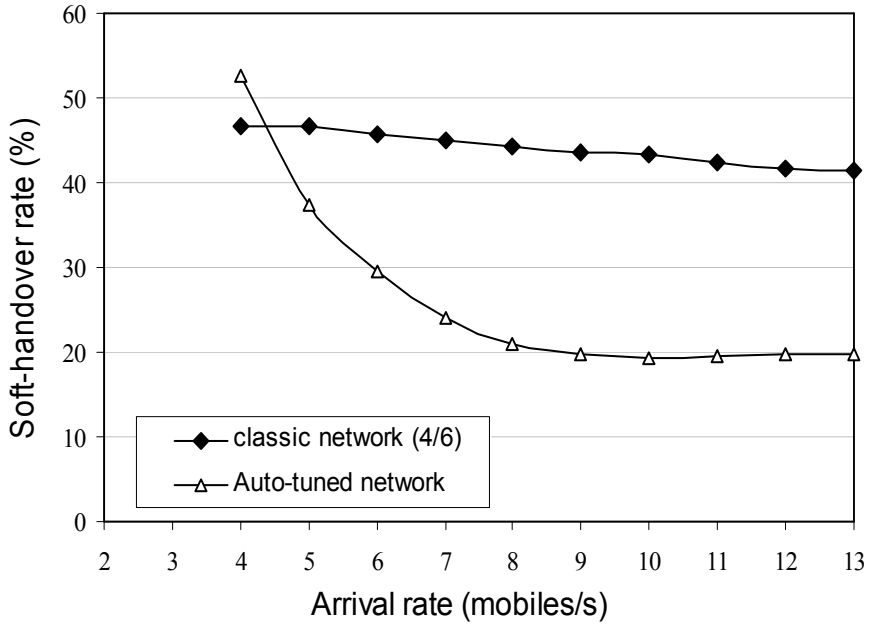


Figure 4.13. Percentage of mobiles in SHO situation as a function of arrival rate for the network without and with auto-tuning.

4.4.4 Signalling overload due to auto-tuning

During a triggering of SHO algorithm, different signalling procedures are involved [95]. The signalling messages for an active set update (link addition or deletion) are mainly transported in the DCCH (Dedicated Control Channel). The DCCH transports the RRC layer (Radio Resource Control) messages for SHO procedure such as measurement report and active set update. The quality of this logical channel is very important for the success of SHO procedures.

The channels DCCH and DTCH (Dedicated Traffic Channel) are mapped into DCH (Dedicated Channel) and SRB (Signalling Radio Bearer) respectively [19]. Sometimes, the SRB is a part of the DCH and its bit rate is usually fixed to 3.4kbit/s. For a call, the RRC signalling traffic is low compared to DTCH traffic but a high number of active set update generates more traffic in the

DCCH. So the question that can be arisen here is whether the auto-tuning influences the capacity in the signalling channel and the stability of the system.

Figure 4.14 shows the impact of the controller on the distribution of Ping-Pong effect. The Ping-Pong effect is related to the frequency of active set update since whenever a mobile add or delete a link in the active set, some signalling messages are involved in the radio and core interfaces. Here the frequency of the active set update is measured as the number of active set updates, generated by each mobile, over its sojourn time in the network. In order to give prominence to this effect, we simulate two network environments: low mobility (speed =3km/h for 20% of users) and high mobility (speed =60km/h for 20% of users) environment. When the controller performs a regulation at every 50s in a low-mobility environment, the frequency of active set update is increased by 10% for more than 10% of mobiles compared to a classical network. So the proposed controlling process increases the signalling messages by 10%. This ratio goes up to 14.3% in a high-mobility environment. By decreasing the reactivity of the controller to one regulation per 100s, the additional signalling introduced by the controller is decreased to 7.1% in a high-mobility environment. The adaptation of the controller or the auto-tuning reactivity needs to be carefully studied in each network environment. The trade-off between the capacity gain and the associate signalling introduced by the auto-tuning is out of scope.

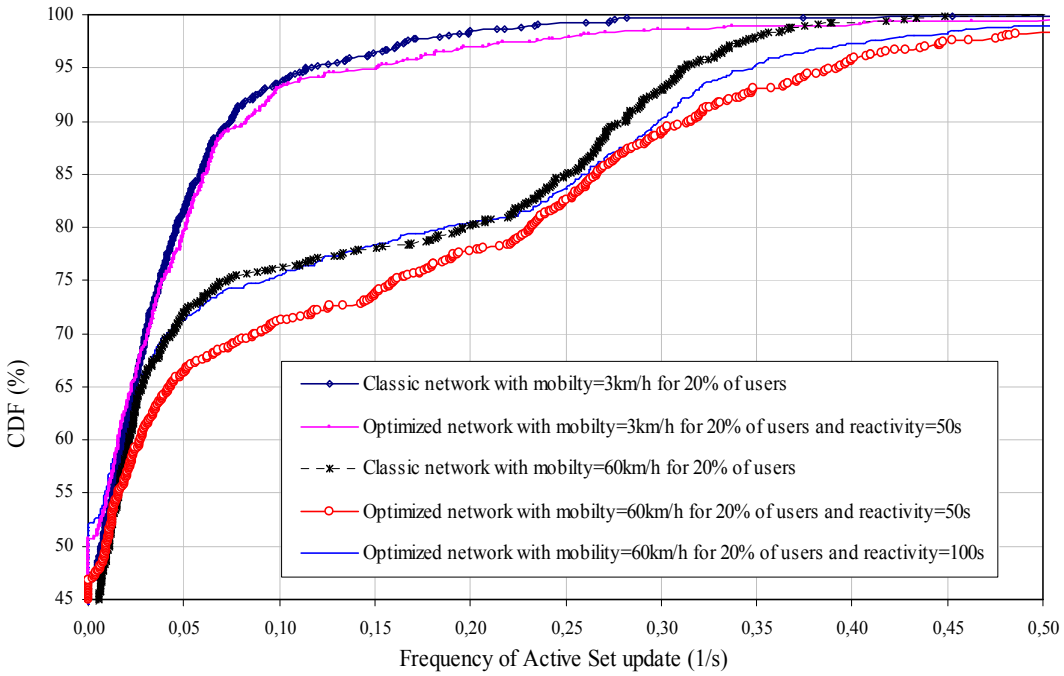


Figure 4.14. Cumulative distribution function of the active set update frequency.

4.5 Conclusions

This chapter has presented results for the application of auto-tuning and optimization algorithms in UMTS networks. The following scenarios have been studied:

The first case is the auto-tuning of resource allocation: the RT guard band is dynamically regulated to achieve optimal tradeoffs between QoS of RT and NRT users. Simulation results have shown an efficient compromise between the perceived QoS in RT and NRT services especially when the traffic is unbalanced. However, when the traffic of both services is very high, the auto-tuning does not improve the QoS for both services.

The second scenario is the mobility auto-tuning: The auto-tuning algorithm dynamically changes the SHO parameters setting as a function of traffic condition, and implicitly performs traffic balancing between cells. This last case study shows an important gain in the overall network performances. The global capacity gains brought by the auto-tuning of SHO parameters are important and typically reach 30% compared to a network with fixed parameters. We have shown that the auto-tuning increases the frequency of active set updates and then increases the signalling messages in the radio interface as well as in the core network. The Ping-Pong phenomena can be reduced by bringing down the reactivity of the auto-tuning. So the auto-tuning should be applied to a network with a special care since it impacts the stability of the system.

The study on auto-tuning of mobility parameters in UMTS network is currently investigated in the long term evolution (LTE) of UMTS in 3GPP TSG-RAN WG3. The next chapter deals with the LTE mobility auto-tuning. Instead of using fuzzy reinforcement learning for the auto-tuning, the LTE mobility adaptation is based on a predefined auto-tuning function.

Chap. 5 Self optimization of mobility algorithm in LTE networks

5.1 Introduction

3GPP (3rd Generation Partnership Project) organization has defined the requirements for an evolved UTRAN (e-UTRAN: UMTS Terrestrial Radio Access Network) [96] and are currently in advanced process of its specification [97]. The evolution of 3G UTRAN is referred to as the 3GPP Long Term Evolution (LTE). Different working groups are involved in defining the architecture and the technology of the radio access and the core network [98]. In the framework of the working group 3GPP TSG-RAN WG3, there have been discussions and studies on the use of self-configuration, self-tuning and self-optimization in the e-UTRAN system [99]. In the first phase of the network optimization/adaptation, neighbour cell list optimization and coverage and capacity control have been proposed [99]. The study of auto-tuning of mobility parameters has been further identified as a relevant case study of self-configuration and has been proposed in different technical reports [100] [101].

The purpose of this chapter is to present an approach for auto-tuning of handover algorithm in LTE system and to present through a case study the performance achieved by the proposed auto-adaptation approach. Unlike in UTRAN where soft-handover is used for mobility, in e-UTRAN a hard handover solution for mobility has been adopted. The handover algorithm has not been specified by 3GPP for the e-UTRAN and for this reason, we adapt an algorithm similar to the one used in GSM networks.

The chapter is organized as follows: in the second section, an overview on the LTE system is presented, including the system requirements, architecture and the physical layer. The third section develops system and interference models. The fourth section deals with the LTE mobility management algorithms including its auto-tuning scheme. Simulation results are given in the fifth section. Finally, a conclusion summarizes the chapter.

5.2 Overview of LTE system

The objective of the LTE is to introduce a new mobile-communication system that will meet the needs and challenges of the mobile communication industry the coming decade [98] [101]. LTE has been often referred to a 4th generation technology. It is characterized by a flat architecture; a new radio access technology with an OFDM (Orthogonal Frequency Division Multiplexing) based physical layer; and considerably enhanced performance with respect to current 3G networks, including delays, high data rates and spectrum flexibility. The LTE technology is specified by 3GPP and is developed in parallel with the evolved HSPA. Unlike the evolved HSPA that comprise a smooth evolution of 3G networks, LTE is fully based on packet switched transmissions with IP based protocols and will not support circuit switched transmissions. The LTE radio access can be deployed in both paired and unpaired spectrum, namely it will support

both frequency- and time-division based duplex arrangements. In Frequency Division Duplex (FDD) downlink and uplink transmission are carried out on well separated frequency bands whereas in Time Division Duplex (TDD) downlink and uplink transmissions take place in different non-overlapping time slots. A special attention is given in LTE to efficient multicast and broadcast transmission capabilities. This transmission is denoted as the Multicast-Broadcast Single-Frequency Network (MBSFN). Standardization of LTE has been carried out in Release 7 and 8, and will continue in 2009. First commercial deployments are expected from 2010.

5.2.1 System requirements

3GPP has defined ambitious performance targets to the LTE system, and the important ones are summarized below [96] [101]. Some of the performance targets are given relative to those of HSPA Release 6 [22]. At the base station, one transmit and two receive antennas are assumed and at the mobile terminal side, one transmit and maximum two receive antennas are assumed.

- Peak data rate of 100 Mbit/s and 50 Mbit/s in downlink and uplink transmissions respectively in a 20 MHz bandwidth,
- Improvement of mean user throughput with respect to HSPA Release 6: 3-4 times in downlink; 2-3 times in uplink; and 2-3 times in cell-edge throughput measured at the 5th percentile,
- Significantly improved spectrum efficiency: 2-4 times that of Release 6, achieved for low mobility, between 0 to 15 km/h, but should remain high for 120 km/h, and should still work at 350 km/h,
- Significant reduction of user and control plane latency with a target of less than 10 ms user plane round-trip time and less than 100 ms for channel setup delay,
- Spectrum flexibility and scalability, allowing to deploy LTE in different spectrum allocations: 1.25, 1.6, 2.5, 5, 10, 15 and 20 MHz,
- Enhanced Multimedia Broadcast/Multicast Service (MBMS) operation.

5.2.2 System architecture

The requirements of reducing latency and cost have led to the design of simplified network architecture, with a reduced number of nodes. The RAN has been considerably simplified. Most functions of the RNC in UMTS have been transferred in the LTE to the eNodeBs (eNB) that constitute now the RAN part, and denoted as the e-UTRAN. The e-UTRAN consists of eNBs interconnected with each other by means of the X2 interface (see Figure 5.1). The eNBs are also connected by means of the S1 interface to the *Evolved Packet Core* (EPC), and more specifically, to the Mobility Management Entity (MME) via the S1-MME interface, and to the Serving Gateway (S-GW) via the S1-U interface. The S1 interface supports a many-to-many relation between MMEs / Serving Gateways and eNBs.

Among the functions of the eNBs are RRM functions, such as radio admission control, radio bearer control, connection mobility control, dynamic resource allocation (scheduling) to the User

Equipment (UE) in both uplink and downlink; IP header compression and encryption of user data stream; routing of user data towards the Serving Gateway (S-GW); scheduling and transmission of paging messages; and scheduling and transmission of broadcast information [97].

The MME is responsible for the following functions: distribution of paging messages to the eNBs; security control; idle state mobility control; SAE bearer control; and Ciphering and integrity protection of Non-Access Stratum (NAS) signalling. The term SAE, or *System Architecture Evolution* has been given by 3GPP to the evolution of the core network, and was finally denoted as the EPC. The Serving Gateway is the mobility anchor point. The different functions of the eNB, MME and the S-GW are depicted in Figure 5.2.

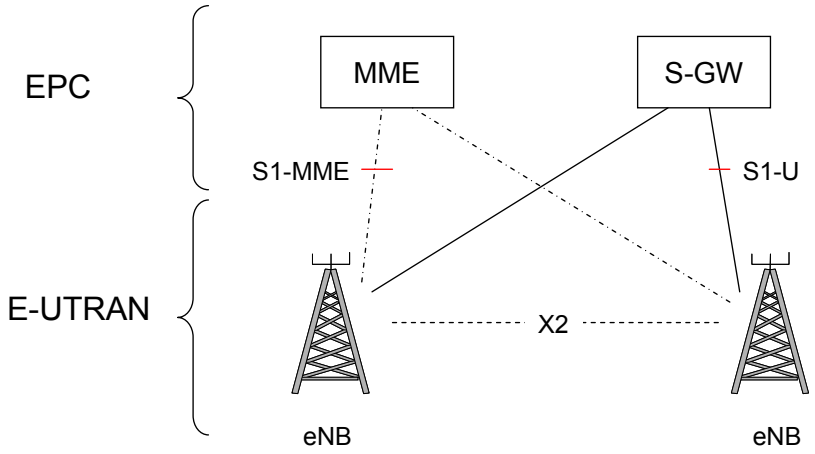


Figure 5.1. LTE architecture.

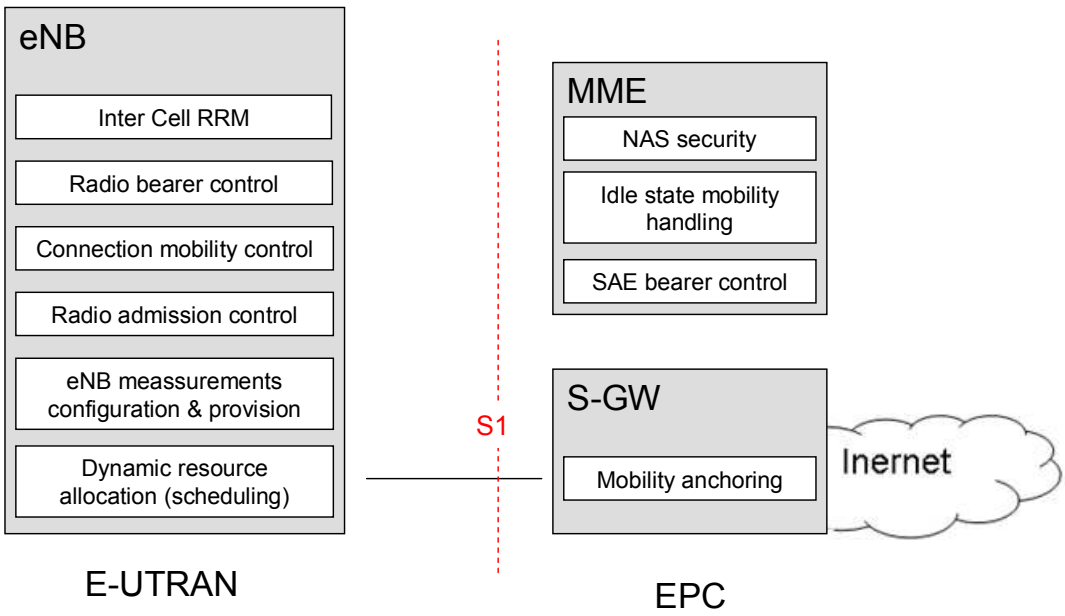


Figure 5.2. E-UTRAN (eNB) and EPC (MME and S-GW).

5.2.3 Physical layer

LTE uses OFDMA (Orthogonal Frequency Division Multiple Access) as the downlink transmission scheme [98] [103]. OFDMA uses a relatively large number of narrowband subcarriers, tightly packed in the frequency domain. The subcarriers are orthogonal, hence without mutual interference. The OFDMA scheme can be rendered robust to time-dispersive channel by the *cyclic-prefix* insertion, namely the last part of the OFDM symbol is copied and inserted at the beginning of the OFDM symbol. Subcarrier orthogonality is preserved as long as the time dispersion is shorter than the cyclic-prefix length. To achieve frequency diversity, channel coding is used, namely each bit of information is spread over several code bits. The coded bits are then mapped via modulation symbols to a set of OFDM subcarriers that are well distributed over the overall transmission bandwidth of the OFDM signal [104]. In the uplink, LTE uses the Single-Carrier FDMA (SC-FDMA: Frequency Division Multiple Access) transmission scheme. This scheme can be implemented using a DFTS-OFDM, namely an OFDM modulation preceded by a DFT (Discrete Fourier Transform) operation. It allows flexible bandwidth assignment and orthogonal multiple-access in the time and frequency domains.

The OFDM transmission scheme allows dynamically sharing time-frequency resources between users. The scheduler controls at each instant to which user allocate the shared resources. It can take into account channel conditions in time and frequency to best allocate resources. According to channel variation, in addition to choosing the mobiles to be served, the scheduler determines the data rate to be attributed to each link by choosing the appropriate modulation. Hence rate adaptation can be seen as part of the scheduler. In the downlink, the smallest assignment resolution of the scheduler is 180 kHz during a 1 ms which is called a resource block. Any combination of resource blocks in a 1 ms interval can be assigned to a user. In uplink, for every 1 ms, a scheduling decision is taken in which mobile terminals are allowed to transmit during a given time interval, on a contiguous frequency region, with a given attributed data rate. Scheduling in LTE is a key element to enhance network capacity.

To enhance the RAN performance, fast hybrid ARQ (Automatic Repeat-reQuest) with soft combining will be used to allow the terminal to rapidly request retransmissions of erroneous transport blocks [98]. From the first release, LTE will support multiple antennas in both eNB and the mobile terminal. Multiple antennas are among the features that will allow the LTE to achieve its ambitious targeted performances, including multiple receive antennas, multiple transmit antennas, and MIMO (Multiple-Input Multiple-Output) for spatial multiplexing.

5.2.4 Self optimizing network functionalities

Within 3GPP Release 8, LTE considers Self Optimizing Network (SON) functions. Some of the SON functions have already been standardized and others are in still being studied. It is noted that the term Organizing is sometimes used instead of Optimizing, but have the same meaning. SON concerns both self-configuration and self optimization processes. Self configuration process is defined as the process where newly deployed nodes are configured by automatic installation procedures to get the necessary basic configuration for system operation [97]. The determination of automatic neighbour cell relation list [105] [106] is an example of self-configuration process

that is being standardized in LTE Release 8. Self-optimization process is defined as the process where user equipment and eNB measurements and performance measurements are used to auto-tune the network. The problem of auto-tuning of mobility parameters as a mean to achieve traffic balancing and to considerably enhance the network capacity, has been discussed within 3GPP [100], and will be presented in details in this chapter.

5.3 Interference in e-UTRAN system

Little material addressing performance and capacity analysis of e-UTRAN is available today. This motivates us to present in this chapter an interference model for the e-UTRAN. The interference is given based on a system model which includes the eNBs distribution and the propagation model. The interference model is used in a second step by the network level simulation.

5.3.1 System model and assumptions

In this section, we analyze only the interference in the downlink. For the uplink, the same concept should be followed. In downlink, each terminal reports an estimate of the instantaneous channel quality to the cell. These estimates are obtained by measuring on a reference signal, transmitted by the cell and used also for demodulation purposes. Based on the channel-quality estimate, the downlink scheduler grants an arbitrary combination of 180 kHz wide resource blocks in each 1 ms scheduling interval. Since the time scale of scheduling is very small, we will not take into account the scheduling process in the interference model and in the system level simulation. Only propagation loss and shadow fading, namely channel variations over large time scales are considered. However, small-scale variations (multi-path fading) are considered in the link level simulation which serves as an input to the present study. The link level simulation returns a link curve which represents the throughput as a function of the received Signal to Interference plus Noise Ratio (SINR). The *Ukumara-Hata* propagation model is used in the 2 GHz band. The attenuation L is given by $L = l_o d^{\gamma} \zeta$, where l_o is a constant depending on the used frequency band, d is the distance between the eNB and the mobile, γ is the path loss exponent and ζ is a log-normal random variable with zero mean and standard deviation σ representing shadowing losses.

5.3.2 Interference model

In e-UTRAN system, user signals are orthogonal in the same cell thanks to the OFDMA access technology. As a consequence there is no intra-cell interference. On the other hand, the same frequency band can be used by a given (central) cell and by some other neighbouring cells. This generates inter-cell interference which limits the performance of the e-UTRAN system. In the downlink for instance, inter-cell interference occurs at a mobile station when a nearby eNB transmits data over a subcarrier used by its serving eNB. The interference intensity depends on user locations, frequency reuse factor and loads of interfering cells. For instance, with a reuse factor equals 1, low cell-edge performances are achieved whereas for reuse factor higher than 1, the cell-edge problem is resolved on the expenses of resource limitation. To make an optimal

trade-off between inter-cell interferences and resource utilization, different interference mitigation schemes are proposed in the standard [98].

One of the techniques for interference mitigation is the inter-cell interference coordination. It is a scheduling strategy in the frequency domain that allows increasing the cell edge data rates. Basically, inter-cell interference coordination implies certain (frequency domain) restrictions to the uplink and downlink schedulers in a cell to control the inter-cell interference. By restricting the transmission power of parts of the spectrum in one cell, the interference seen in the neighbouring cells in this part of the spectrum is reduced. This same part of the spectrum can then be used to provide higher data rates for users in the neighbouring cell. This mechanism is called also partial (or fractional) frequency reuse because the frequency reuse factor is different in different areas of the cell (Figure 5.3). The partial frequency reuse scheme proposed for e-UTRAN is a combination between reuse 1 and 3. In a fractional frequency reuse, an admitted user gets resource blocks from the portion of the bands as a function of its position in the cell. Of course, if the user is located in the cell edge, he gets resources from the cell edge band, (denoted also as protected band).

Assuming that the spectral band is composed of C resource blocks, one third of the band is reserved for the cell edge users and the rest is for cell centre users.

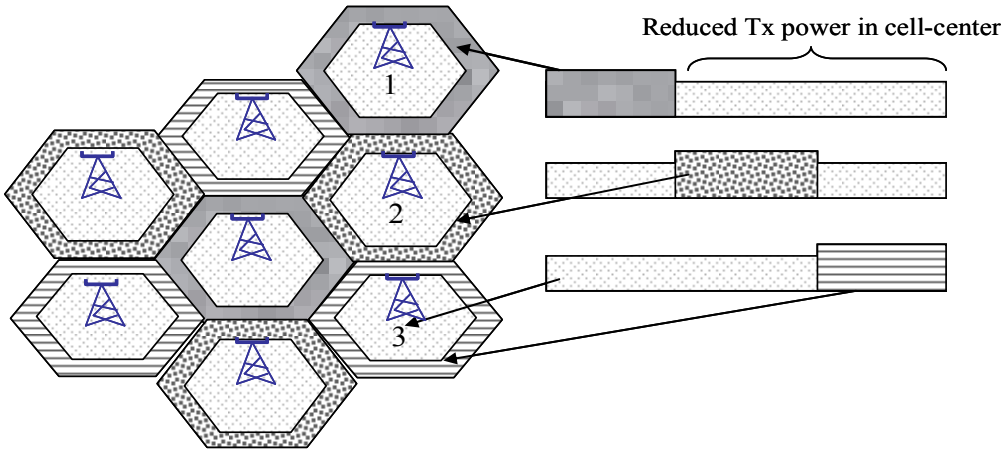


Figure 5.3. Inter-cell interference coordination scheme.

The resource allocation is made according to users' positions. Let L_{th} be the path loss threshold separating users from the two different sub-bands. A user with a path loss higher than L_{th} is granted resource blocks in the cell-edge band and otherwise he gets resources in the cell center band. The eNB transmit power in each cell-edge resource block equals the maximum transmit power P. To reduce intercell interference, the eNB transmit power in the cell-center band must be lower than P. Let εP (where $\varepsilon < 1$) be the transmit power in the cell-center band.

The interference should be determined for two different users according to their positions: the cell-center user and the cell-edge user. Let m_c and m_e be two users connected to a cell k. the mobile m_c uses the central band whereas m_e uses the cell-edge band. Let Λ denote the

interference matrix between cells, where the coefficient $\Lambda(i,j)$ equals 1 if cells i and j use the same cell-edge band and zero otherwise.

For cell-edge user m_e, the interference comes from users in the cell center of the closest adjacent cells and from the cell-edge user in other cells. The mobile m_e connected to the cell k and using one resource block in the cell-edge band, receives an interfering signal from a cell i equals

$$I_{i,m_e} = \left((1 - \Lambda(k,i))\beta_i^c \varepsilon P_i + \Lambda(k,i)\beta_i^e P_i\right)\frac{G_{i,m_e}}{L_{i,m_e}} \tag{5.1}$$

where P_i is the downlink transmit power per resource block of the cell i. G_{i,m_e} and L_{i,m_e} are respectively the antenna gain and the path loss between cell i and the mobile station m_e. The factor β_i^c (respectively β_i^e) is the probability that the same resource block in the center-cell band (respectively the cell-edge band) is used at the same time by another mobile connected to the cell i.

Using analysis given in appendix B, the total interference perceived by user m_e is the sum of all interfering signals

$$I_{m_e} = \sum_{i \neq k} \tilde{\Lambda}_e(k,i)\chi_i \frac{P_i G_{i,m_e}}{L_{i,m_e}} \tag{5.2}$$

the term $\tilde{\Lambda}_e(k,i) = 3\left((1-\Lambda(k,i))\frac{(1-\alpha_i)}{2}\varepsilon + \Lambda(k,i)\alpha_i\right)$ is interpreted as a new interference matrix, denoted here as the fictive interference matrix for cell-edge users.
The factor α_i is defined as the proportion of traffic served in the cell-edge band of cell i,

For the cell-center user m_c, the interference comes from users in the cell-edge and cell-center of closest adjacent cells and also from the cell-center and cell-edge users in other cells. Similarly to the cell-edge users, the mobile m_c connected to the cell k and using one resource block in the cell-center band, receives an interfering signal equal to

$$I_{m_c} = \sum_{i \neq k} \tilde{\Lambda}_c(k,i)\chi_i \frac{P_i G_{i,m_e}}{L_{i,m_e}} \tag{5.3}$$

Here, the fictive interference matrix in the cell-center band is given by

$$\tilde{\Lambda}_c(k,i) = 3\left(\tfrac{1}{2}(1-\Lambda(k,i))\left(\frac{1-\alpha_i}{2}\varepsilon + \alpha_i\right) + \Lambda(k,i)\frac{1-\alpha_i}{2}\varepsilon\right) \tag{5.4}$$

The downlink *SINR* is then given by

$$SINR_m = \frac{P_k G_{k,m}}{L_{k,m}\left(I_m + N_{th}\right)} \tag{5.5}$$

In equation (5.5), the subscript m stands for m_e if the mobile considered is a cell-edge user and m_c for the cell-center user. N_{th} is the thermal noise per resource block.

For more details about the LTE interference model, reader is invited to see the appendix B.

In e-UTRAN system, an adaptive modulation and coding scheme are used [103] [104]. So, the choice of the modulation depends on the value of the SINR through the perceived Bloc Error Rate (*BLER*). The decrease of the *SINR* will increase the *BLER*, forcing the eNB to use a more robust (less frequency efficient) modulation. The latter may have negative impact on the communication quality. For instance, a lower modulation efficiency results in a lower throughput and a larger transfer time for elastic data connections. In the present chapter, the throughput per resource block for each user is determined by link level curves. The user physical throughput is N_m times the throughput per resource block, where N_m represents the number of resource blocks allocated to the user m.

5.4 Auto-tuning of e-UTRAN handover algorithm

Mobility in e-UTRAN is based on hard handover rather than on soft handover as in UMTS [11]. The mechanism of hard handover has been used in 2nd generation GSM networks and has shown to be efficient for mobility management. In a hard handover, the user keeps the connection to only one cell at a time, breaking the connection with the former cell immediately before making the new connection to the target cell. The basic concept of handover as in GSM is likely to be implemented in e-UTRAN except for the handover preparation phase, which requires new mechanisms.

The reason for abandoning soft handover is related to the extra-complexity involved in its implementation, and the fact that it is not suitable for inter-frequency handover. Furthermore, as explained in the previous chapter, soft handover handicaps system capacity in highly loaded network condition and with high number of users in soft handover situation. To guarantee seamless and lossless hard handover in e-UTRAN, the handover triggering time should be as low as possible [98].

In general, there are different causes for handover: a user can move to another cell because he verifies a power budget condition or because he experiences bad channel quality. In this work we consider only the first case, namely power budget based handover. This handover is based on the comparison of the received signal strength from the serving cell and from the neighbouring cells.

5.4.1 E-UTRAN handover algorithm

In order to study the performance of the auto-tuning of e-UTRAN hard handover, some assumptions for the call admission control (CAC) and resource allocation are made. In the CAC algorithm, a user can be admitted to the network only when the following conditions are fulfilled:

- Good signal strength: the mobile selects the cell that offers the maximum signal. If this signal is lower than a specified threshold then the mobile is blocked because of coverage shortage. This condition is in fact a selection criterion. It is noted that in 3GPP, there is no specification for LTE cell selection and reselection.
- Resource availability in the selected cell: the mobile can be granted physical resources in terms of resource blocks between a minimum and maximum threshold. When the signal strength condition is satisfied, the eNB checks for resource availability. If the available resource is lower than a minimum threshold, the call is blocked.

Hard handover is performed in this study using a similar algorithm to the one used in GSM: while in communication, the mobile periodically measures the received power from its serving eNB and from the neighbouring eNBs. The mobile, initially connected to a cell k, triggers a handover to a new cell i if the following conditions are satisfied:

- The Power Budget Quantity (PBQ) is higher than the handover margin:

$$PBQ = P_i^* - P_k^* \geq HM(k,i) + Hysteresis \tag{5.6}$$

  where P_k^* is the received power from the eNB k expressed in dB; $HM(k,i)$ is the handover margin between eNB k and i; the *Hysteresis* is a constant independent of the eNBs and mobile stations and is fixed in this study to 0.
- The received power from the target eNB must be higher than a threshold. This is the same condition as in the CAC process.
- Enough resource blocks are available in the target eNB.

The last condition requires information exchange between eNBs because the original cell has to know a priori the load of the target cell; otherwise the handover is blind and the communication risks to be dropped. In an inter-eNB handover procedure, the source eNB is responsible for performing handover preparation to the target eNB based on measurement report transmitted by the mobile. If the highest ranked cell listed in the mobile measurement report is congested and can not admit incoming calls especially real-time service calls, the source eNB performs handover preparation to the cell with the next highest rank. On the other hand, if the source eNB has the load knowledge of its neighbour's eNBs, it could efficiently decide to which cell the handover should be performed before initiating handover preparation procedures [107]. Therefore, the source eNB needs to perform handover decision considering load information.

In 3GPP proposals, there are mainly two solutions for the source eNB to know whether the highest ranked cell can admit the incoming call or not. The first solution is to standardize load

information exchange on the interface X2 as a common measurement, whereas the second solution concerns the implication of the target cell in the handover preparation phase. The target cell can reply a *handover preparation failure* for example if it is loaded.

The first solution achieves shorter delay in handover preparation phase; however it requires the definition of a new eNB measurement for load information. The exact definition of eNB load is still under discussions in 3GPP. The second solution seams to be similar to the existing handover procedures (in UMTS and GSM) where base stations are not aware of the load status of their neighbours. In this second solution, the target eNB has to respond to each handover request message regardless of its congestion state. As a result, the processor load and signaling messages in the X2 interface can rise. If the highest ranked cell can not accept the call, the source eNB is compelled to try handover preparation to another eNB resulting in a longer delay for handover preparation phase. The first solution is preferred to the second one because of the requirement that the hard handover be lossless and the handover triggering time should be as low as possible.

5.4.2 Handover adaptation and load balancing

From the previously described handover algorithm, a low handover margin allows users to be connected to the closest cell everywhere in the cell, but ping-pong effect may occur frequently. On the other hand, a high handover margin generates high interference for cell-edge users especially with the use of low frequency reuse factor, but ping pong effect problems are avoided. Adapting the handover margin allows to achieve interesting trade-offs between different network states.

Figure 5.4 presents a typical handover situation, whereas Figure 5.5 shows a situation where handover thresholds are adapted to the relative cell load, rather than being constant. In Figure 5.5 Some of users in the handover zone that would otherwise be served by the congested cell (cell k) are now handed over to the less congested cell. This can be achieved by delaying the handover to the congested cell and advancing the handover from the congested cell or in other words decreasing the handover margin from cell k to cell i and increasing the handover margin for the other direction. In fact, decreasing handover margin from cell k to cell i allows more users to verify the power budget handover condition. So, it leads to a decrease of the service area of congested cells and to an increase of the service area of less-loaded cells.

Auto-tuning in a non-uniformly loaded network could be particularly beneficial in e-UTRAN. It implements a simple load balancing mechanism which increases the overall capacity of the system, by simply distributing the load more evenly between the neighbouring cells.

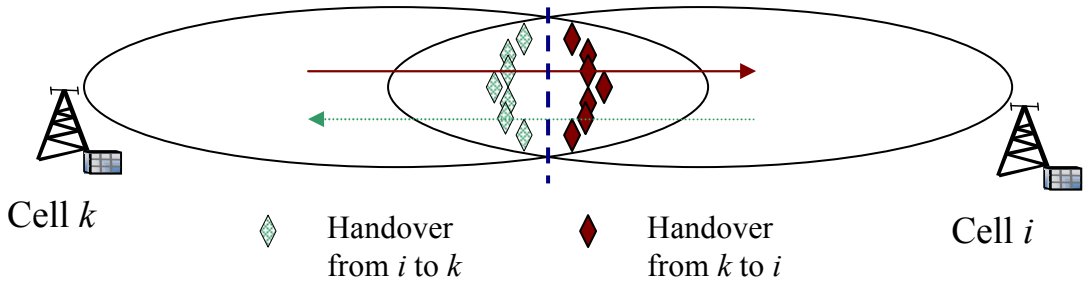


Figure 5.4. Typical pattern of geographical distribution of HO procedure.

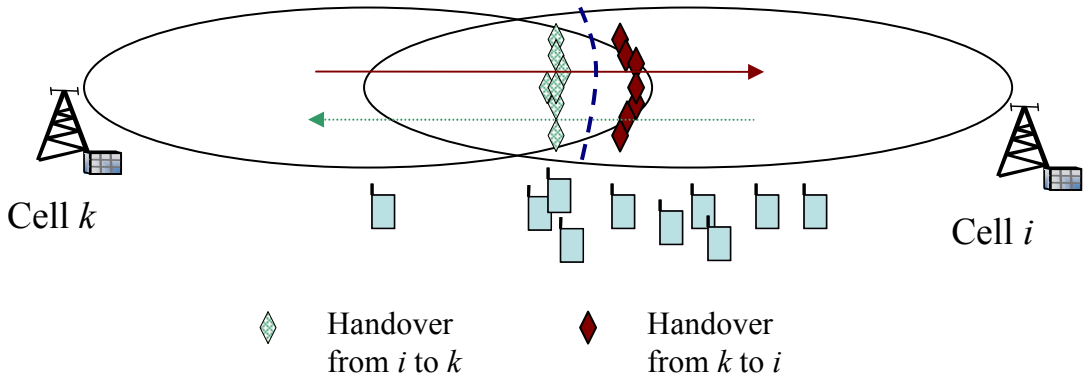


Figure 5.5. Example of geographical distribution of HO procedure with traffic balancing.

5.4.3 Auto-tuning of handover margin

The auto-tuning aims at dynamically adapting handover margins between cells as a function of their loads, to optimize network performance. Each coefficient of the matrix *HM* governs the traffic flows between two cells. The coefficient *HM*(*k,i*) depends only on the difference between the load of cell *i* and *k*. Define the handover margin matrix *HM* as

$$HM(k,i) = f(\chi_k - \chi_i) \tag{5.7}$$

The function *f* should satisfy:

(*i*) *f* is a decreasing function from the interval [-1,1] to [HM_{min}, HM_{max}],

(*ii*) $f(x) + f(-x) = 2f(0), \forall x \in [-1,1]$,

where HM_{min} and HM_{max} are respectively the minimum and the maximum values of the handover margin. $f(0)$ is the value of the planned handover margin since the planning process assumes the uniformity of the cell loads.

The first condition implies that when the cell k is fully loaded and i does not serve any mobile, (i.e. $\chi_k - \chi_i$ approaches 1) it is worth keeping the handover margin $HM(k,i)$ to the lowest value. For the second condition, let x be defined as the difference between loads of cell k and i (i.e. $x=\chi_k - \chi_i$); this condition is used to avoid ping pong effect. It implies that when the cell k is over loaded and the cell i is less-loaded, cell k pushes mobiles to cell i and conversely, cell i delays handover to cell k.

The function f can be approximated by a polynomial (with Taylor series expansion). The polynomial coefficients can be dynamically determined using learning techniques. The concept of fuzzy reinforcement learning presented earlier in previous chapters could be well suited. In the present study, we restrict the development of $f(x)$ to the order 1:

$$f(x) = f(0) + (f(0) - HM_{\max})\, x \tag{5.8}$$

The development of order 0 corresponds to the classical case without any auto-tuning. The simulations aim at comparing the development of order 1, namely with auto-tuning to the case without auto-tuning.

5.5 Simulations and results

To evaluate the performance of the proposed auto-tuning method, a dynamic simulator developed using *Matlab* tool has been utilized. The simulator is conceptually similar to the UMTS simulator, presented in appendix C.

Simulations have been carried out on a 3G LTE network composed of 45 eNBs (Figure. 5.6). Each eNB has a fixed capacity equal to 25 resources blocks (corresponding to a 5 MHz bandwidth). The studied scenario uses a non-uniform traffic distribution resulting in unbalanced cell loads. Only an FTP service class is considered. An FTP call is generated by a Poisson process and the communication duration of each user depends on its bit rate. Each user is allocated at least one resource block and at most 4 resource blocks to download a file of 5 Mbytes. The value of the function f in 0 equals 6 dB. The minimum and the maximum handover margin values, HM_{min} *and* HM_{max}, are set respectively to 0 dB and 12 dB.

Figure 5.7 presents the access probability (the complementary of the blocking rate) versus the traffic intensity for the case of auto-tuning compared with the classic case without auto-tuning. As expected, the gain of using auto-tuning is important when the traffic intensity is low because the dispersion of cell loads is still high. For high traffic intensities, all cell loads approach 1 and the load difference of adjacent cells becomes too small to benefit from traffic balancing. According to the auto-tuning of order 1, the handover margin tends to the default handover margin, $f(0)$, when the traffic increases and all loads tend to 1.

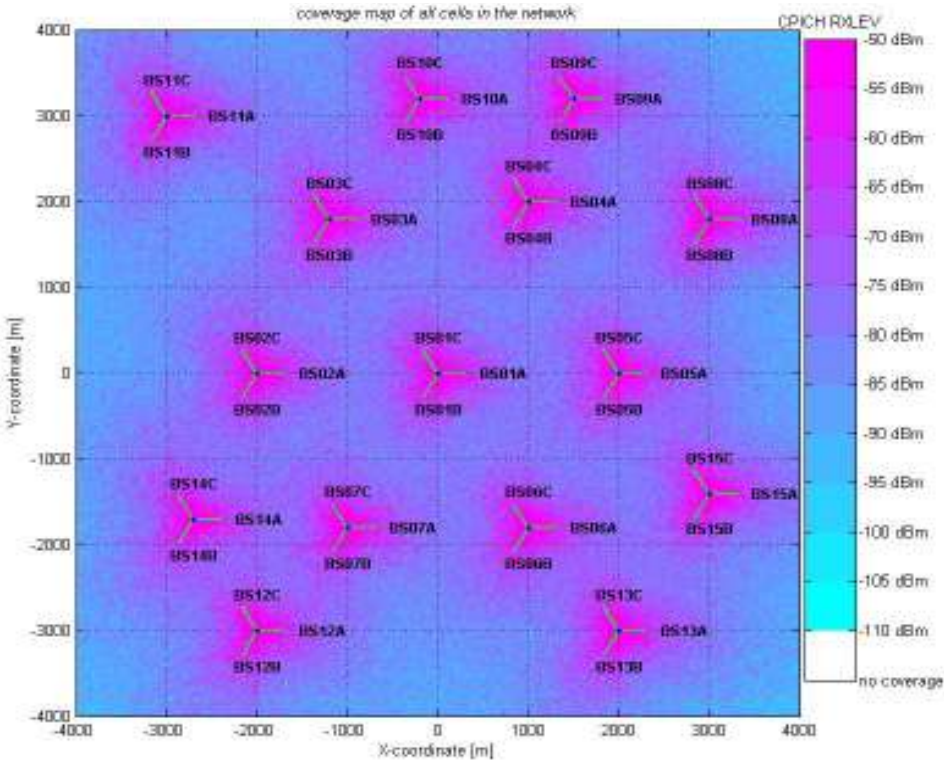


Figure 5.6. The network layout including coverage of each eNB.

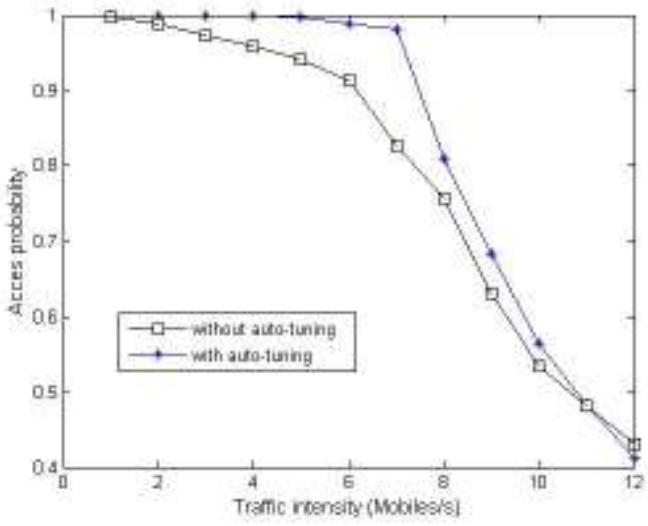


Figure 5.7. Admission probability as a function of the traffic intensity for auto-tuned handover compared with fixed handover margin network (6dB).

In figure 5.8, we present the connection holding rate (the complementary of the dropping rate) as a function of the traffic intensity. We notice that the variation of the holding rate is small when the traffic increases. This is due to the fact that a mobile is not dropped when there is not enough resources but instead its throughput decreases (i.e. the number of allocated resource blocks

decreases). The auto-tuning gain for this quality indicator is modest. The trend of the two curves for the high traffic intensity, confirms again that the auto-tuning tends to the classic case in very high traffic condition.

Figure 5.9 shows the average throughput per user as a function of the traffic intensity. The throughput per user is a decreasing function of the traffic rate since it is an increasing function of the SINR. The achieved gain of the auto-tuning is high. For instance, for the traffic intensity equals 5 users/s, the throughput per user is approximately 1.15Mbyte/s whereas for the classic case, it is only 0.975Mbyte/s. This gain is explained by the following two reasons:

1. The implementation of resource allocation: when there are enough resources in the cell, the user gets the maximum number of resource blocks. So its bit rate is high and the user ends quickly its communication. As a consequence, it rapidly releases resources for new users. This explains the gain in successful access rate brought by the auto-tuning.
2. Interference diversity: due to the auto-tuning, the distribution of inter-cell interference becomes more or less the same in each eNB since the interferences experienced by each user depend on the load of the neighbouring cells. Hence, load balancing leads to interference diversity.

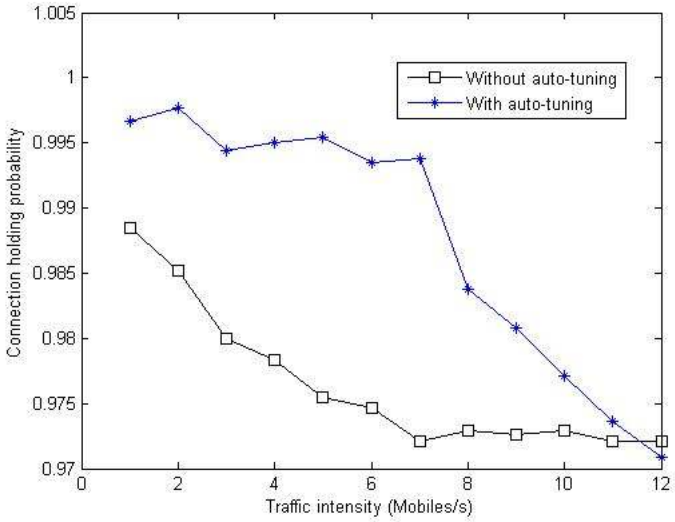


Figure 5.8. Connection holding probability as a function of the traffic intensity for auto-tuned handover compared with fixed handover margin network (6dB).

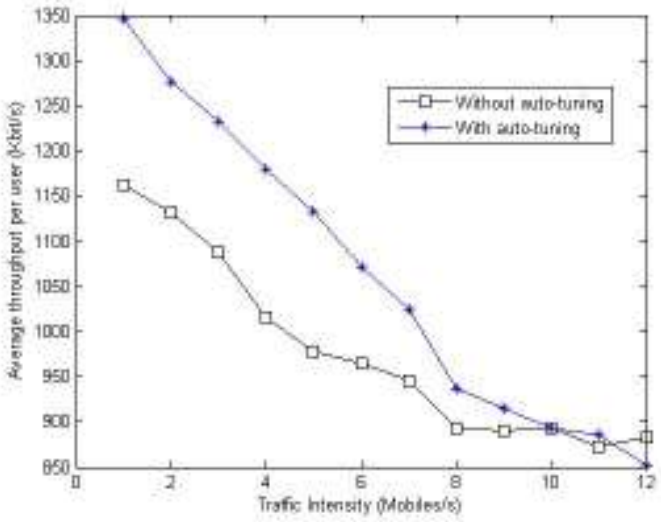


Figure 5.9. Average throughput per user versus the traffic intensity for auto-tuned handover compared with fixed handover margin network (6dB).

In figure 5.10, the cumulative distribution for SINR is presented for both the auto-tuning and the classical cases. The traffic intensity has been set to 8 mobiles/s. The interference diversity generated by the auto-tuning mechanism leads to an increase of the perceived SINR.

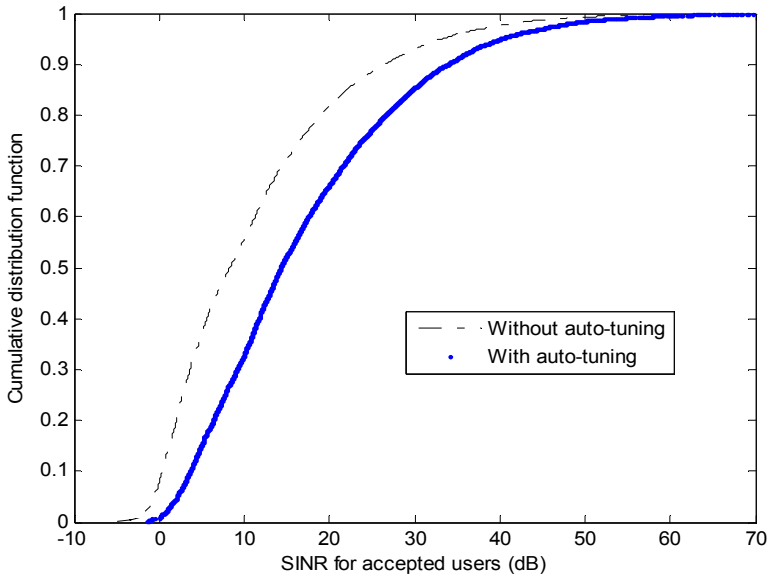


Figure 5.10. Cumulative distribution function of the SINR for network with and without auto-tuning, for traffic intensity equals to 8 mobiles/s.

5.6 Conclusion

This chapter has investigated auto-tuning of mobility algorithm in e-UTRAN system. The mobility is based on hard handover. The handover margin involving each couple of eNB governs the hard handover and its value directly affects the radio load distribution between the cells. The auto-tuning of the handover margin parameters balances the traffic between neighbouring cells. As a consequence, the system capacity is increased and the user perceived quality of service, namely the user throughput, is enhanced. The auto-tuning functionality has been incorporated into a dynamic system level simulator and has been implemented to a non regular (e.g. cells are not hexagonal) LTE network. Significant improvement in the cumulative distribution of the signal over interference has been achieved, and more than 15 percent increase in user throughput has been attained. These results show the importance of mobility auto-tuning to the performance of the e-UTRAN system.

Chap. 6 UMTS-WLAN load balancing by auto-tuning inter-system mobility

6.1 Introduction

WLAN networks become the most popular wireless technology to cover hot spot areas. WLAN are high-capacity networks that can offer high bit rates for users with low mobility. They are used by cellular network operators to absorb traffic in localized high traffic zone and to relieve the wide area cellular networks. The integration of WLANs within a cellular network is difficult and challenging. Standardization bodies, such as 3GPP, IETF, IEEE and ETSI, are actively working on RATs' inter-working [108].

The aim of this chapter is to propose an efficient UMTS-WLAN algorithm for load balancing by means of intersystem call admission and forced handover. To further improve the network performance, the Vertical Handover (VHO) algorithm is dynamically optimized using auto-tuning process as described in chapter 3 and used in chapter 4. The traffic balancing strategy is the following: If a mobile having a packet-based application demands access in a WLAN coverage zone, it is admitted to the network that can offer a high bit rate. When a UMTS base station or a WLAN access point gets congested, a VHO towards the other system is triggered to balance the traffic between the two systems. The Joint Radio Resource Management (JRRM) algorithm compares the UMTS or the WLAN load to a target threshold to decide whether or not to perform a VHO. Numerical simulations using a semi-dynamic network simulator illustrate the effectiveness of the proposed approach.

The chapter is organized as follows: in the second section, we present the assumptions used in the study. These assumptions cover the UMTS-WLAN inter-working mode and the inter-system selection and admission control. The third section deals with the proposed vertical handover and its auto-tuning. In section 4, we present the system performances in terms of capacity and throughput. Finally, a conclusion ends the chapter.

6.2 Assumptions

6.2.1 UMTS-WLAN inter-working mode

In the present chapter, a WLAN network based on the IEEE 802.11b specification [108] is considered. The UMTS network is supposed to be deployed according to 3GPP release 5 or 6.

As depicted in figure 6.1, we use an access network which is a very tightly coupled UMTS/WLAN network. Recall that the very tight coupling mode involves establishing the WLAN network as a second radio network system, integrated at the UMTS RNC. The benefit of this mode is that the resource control of WLAN is co-located with the resource control of UMTS. In this way, it is possible to manage the WLAN hotspot as a UMTS cell or as a part of it.

Furthermore, the very tight coupling mode allows a seamless handover between both technologies.

We assume that the network environment is composed of UMTS cells with WLAN hotspots arranged in a hierarchical cell structure. This means that all WLAN Access Points (AP) are under UMTS coverage. WLAN APs are associated with specific UMTS location areas and managed by JRRM procedures installed in the RNC. The deployment of WLAN under the UMTS converge implies that coverage-based handover is allowed only from WLAN to UMTS.

Over time, the mobile terminal moves from an area supporting only UMTS coverage, to an area with both UMTS and WLAN. The mobile should be capable of reporting the WLAN RSSI values to the RNC, to update the Common RRM entity with regard to the radio conditions and facilitate an inter-system load-based handover. Of course, mobile terminals are assumed to have both UMTS and WLAN capabilities. In addition, mobile users are assumed to have Subscriber Identity Module (SIM) cards authorized to get access to both technologies.

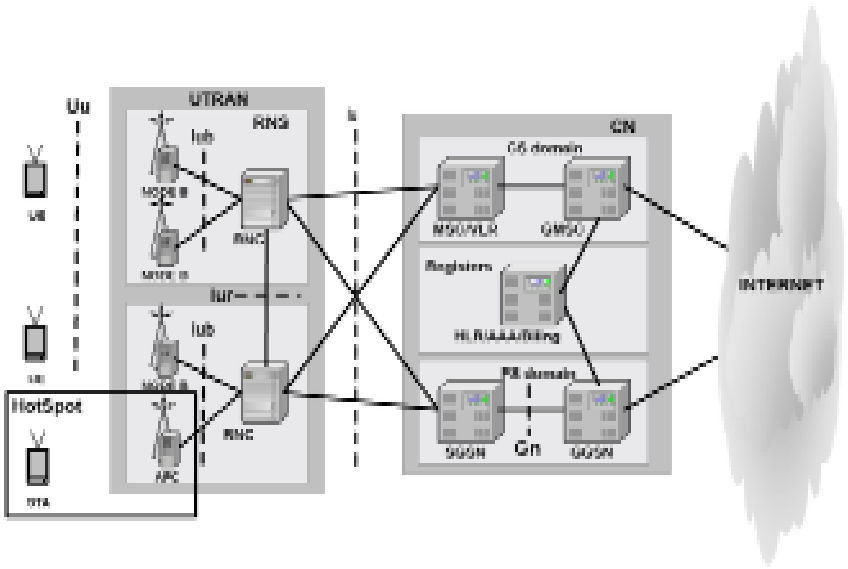


Figure 6.1. Very tightly coupled UMTS/WLAN network.

6.2.2 Technology selection and admission control

The selection procedures concern the idle mode state of a mobile. Selection or reselection is triggered on either initial power-up or on change of available network coverage but does not concern the active call. Since the UMTS coverage is always available in the network, the mobile selects the cell as in the case of a UMTS network alone without any other integrating technology.

Once the mobile selects a UMTS cell, it starts to receive network information in signalling messages. Camping in a cell allows the mobile to register in a UMTS location area. While it is still in the coverage of the UMTS technology, the mobile can detect a WLAN hotspot by scanning periodically the medium, by actively transmitting Probe Requests to identified access

points or by passively waiting to receive Beacon Frames. The RSSI values on the link can be measured when the management frames are received. The WLAN SSID is transmitted as part of the Beacon Frame and the mobile can use this information to initiate the association procedure. The mobile that initially camps in the UMTS cell, selects a WLAN AP and associates it to its best UMTS cell. This allows the mobile to register and access WLAN services while still being able to receive signalling, paging and system information from UMTS.

The user is identified within the UTRAN and Core Network, and signalling bearers are set up for him. To change from idle to connected mode, the mobile user must send an RRC connection request. This procedure is common to all call types, as the signalling has to be set up prior to actual call establishment. Being in connected mode allows the mobile user to establish call sessions.

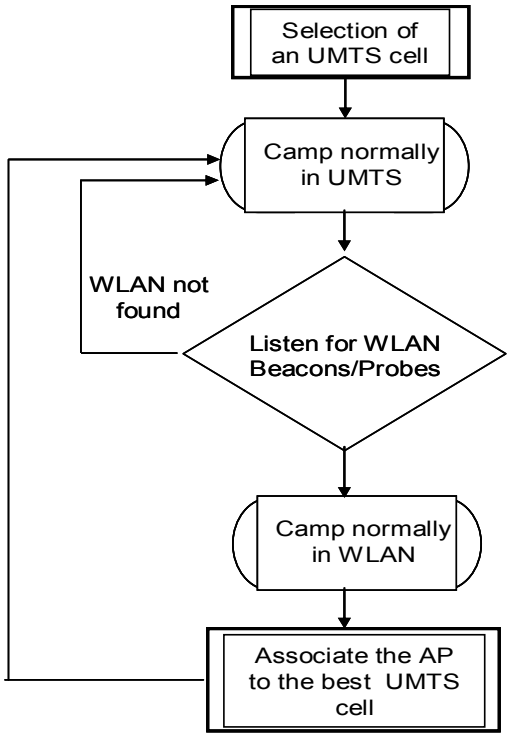


Figure 6.2. Selection procedures in very tightly coupled UMTS/WLAN network

For the admission control process it is assumed that calls requiring tight QoS requirements such as voice/video conferencing and voice calls (RT service) use UMTS system as the preferred network and are not allowed to connect to the WLAN system. However, streaming and background services (NRT calls) are enabled over WLAN.

When a user generating a data call is under the WLAN coverage and gets a minimum bit rate, it uses the WLAN as its default technology. It is noted that other CAC strategies can be implemented as well. The main difference between establishing a call on WLAN as opposed to UMTS is the indication of where the radio bearer is established. When the JRRM entity is

interrogated during the admission control procedure, the reply for a data call is to establish the bearer over WLAN. With signalling for the data call being provided by UMTS, the message sequence is similar to that of the originating voice call. The differences in this case are introduced when the bearer is being set up and negotiated. The bearer results in data being transported over WLAN rather than UMTS, so the signalling must reflect this.

6.3 UMTS-WLAN vertical handover and its auto-tuning

6.3.1 Vertical handover description

The load based VHO algorithm between UMTS BSs and WLAN APs is presented in figure 6.3. Let $T_{U\text{-}W}$ be the load threshold of UMTS BS that is used to trigger VHO from UMTS to WLAN, and $T_{W\text{-}U}$ the threshold of minimum throughput offered by an AP that is used to trigger a VHO between WLAN to a UMT BS. Denote by L_{UMTS} the load of a UMTS BS and let L_{WLAN} be the throughput offered by a WLAN AP. The minimum throughput that can be offered by an AP is a good indicator of the AP load. So it is necessary to keep the minimum offered throughput higher than $T_{W\text{-}U}$ to guarantee user-quality of service. The UMTS BS is loaded if $L_{UMTS} > T_{U\text{-}W}$ but an AP is loaded if $L_{WLAN} < T_{W\text{-}U}$. To avoid BS congestion, $T_{U\text{-}W}$ is assumed to be lower than the UMTS admission threshold.

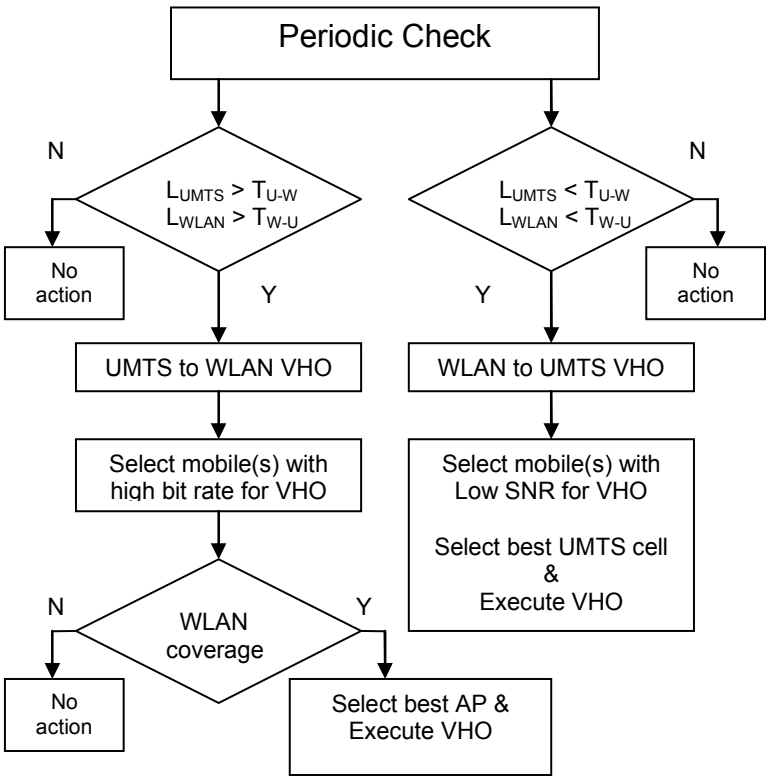


Figure 6.3. Load based VHO algorithm between UMTS and WLAN networks.

Periodically loads of BSs are checked. If a BS load exceeds $T_{U\text{-}W}$, mobiles connected to that BS are sorted according to their bit rate. The ones with highest bit rate are handed over to a neighbouring AP of the BS until BS-load becomes lower than $T_{U\text{-}W}$. Of course, coverage of handed-over mobiles and AP load are checked in the destination AP before handover triggering. Likewise, the minimum throughput that can be offered by an AP is checked periodically. If an AP is unable to offer a throughput higher than $T_{W\text{-}U}$, mobiles with low SNR are handed over to a non-loaded neighbouring BS. When the thresholds $T_{U\text{-}W}$ and $T_{W\text{-}U}$ are well parameterised, the traffic balancing enhances network capacity.

6.3.2 Auto-tuning of vertical handover parameters

To dynamically and optimally set the vertical handover parameters $T_{U\text{-}W}$ and $T_{W\text{-}U}$, we use the fuzzy Q-learning controller as described in chapter 3 and used in chapter 4. For this case study, we use quality indicators from both system and optimize jointly both parameters $T_{U\text{-}W}$ and $T_{W\text{-}U}$. The controller is performed only in UMTS cells but the change affects also WLAN APs. When the controller performs a change of the parameter $T_{U\text{-}W}$ in a UMTS cell the parameter $T_{W\text{-}U}$ is changed in all associated APs.

The used controller input is the vector

$$s = \left(CSR^{u}_{RT}, CSR^{u}_{NRT}, CSR^{nu}_{NRT} \right) \tag{6.1}$$

where CSR_{RT} and CSR_{NRT} stand for call success rate of real time (RT) and non-real time (NRT) services respectively in an UMTS base station. The indicator CSR^{nu}_{NRT} of a UMTS base station is defined as

$$CSR^{nu}_{NRT} = \sum_{i \in NS(u)} \omega_{u,i} CSR^{i}_{NRT} \tag{6.2}$$

where $NS(u)$ is the neighbouring set of access points associated to the UMTS BS u, and $\omega_{u,i}$ is a weighting coefficient defined as the normalized traffic flux of mobiles originating at BS u and handed over the AP i. All quantities in (6.1) and (6.2) are filtered using an averaging sliding window as presented in equation (4.3).

The controller produces as output a simultaneous modification to the load threshold $T_{U\text{-}W}$ governing the handover from UMTS to WLAN and the throughput threshold $T_{W\text{-}U}$ responsible for the VHO from WLAN to UMTS.

The reinforcement function is defined similarly to equation (4.6), with ω=0:

$$r(t) = \alpha CSR_{TR} + \beta CSR_{NTR} \tag{6.3}$$

(α,β) is the weighting vector that gives the desired importance to each quality indicator. The change of the importance vector is carried out according to the traffic conditions as follows:

- ✓ If ($CSR_{RT} \leq 0.95$) then $(\alpha,\beta) = (0.5, 0.5)$;

- ✓ If (CSR_{RT}>0.95) then (α,β) = (0, 1).

6.4 Simulations and performance evaluations

A multi-system network with full UMTS coverage is considered with 59 UMTS cells and 22 access points (APs) in a dense urban environment (see Figure 6.4). Recall that the WLAN network considered is based on the IEEE 802.11b specification [108]. The APs are located in the higher traffic zones within the UMTS covered area. Each BS has a fixed downlink capacity defined as the maximum transmit power and each AP has a fixed capacity defined as the maximum bit-rate that can be granted to a NRT user. This bit-rate depends on two factors. The first is the user conditions, since mobile with low SNR gets lower bit rate due to the link adaptation [108]. The second factor is the number of users: in the CSMA/CA access medium mechanism, when the number of users increases, collision probability increases, resulting in the reduction of the average offered bit-rate. The WLAN network supports only NRT FTP traffic, whereas UMTS supports both RT voice and NRT FTP traffic. The FTP data calls arrive in the network according to a Poisson process and each user downloads data traffic file of one Mbytes. Arrival rate of voice mobiles is 6 mobiles/sec. Both indoor and outdoor traffic is present, with 40 and 60 percent respectively. The outdoor users are mobile with 3km/h speed.

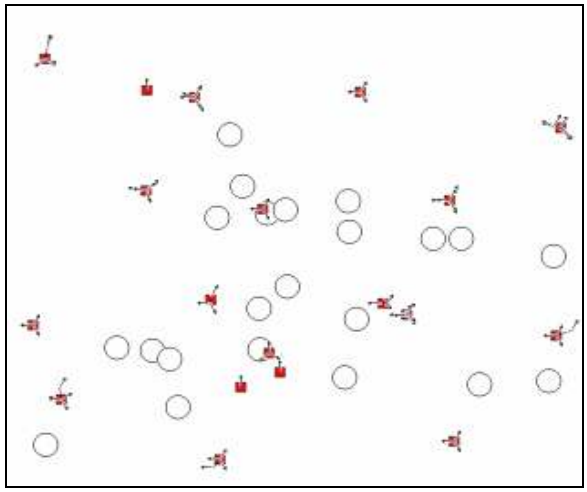


Figure 6.4. Heterogeneous network layout with 59 UMTS cells (squares with arrows) and 22 WLAN APs (circles).

Two scenarios are considered: the first uses the traffic balancing algorithm. The second scenario both traffic balancing and auto-tuning are combined. The auto-tuning is performed station by station in the network. In both scenarios, the real time voice traffic is kept fixed to 6 mobiles per second, whereas, the packet switched (PS) FTP traffic is increased in a series of distinct simulations.

Figures 6.5 and 6.6 compare the call success rate (CSR) of RT and NRT traffic respectively as a function of the FTP traffic intensity, with (squares) and without (diamonds) auto-tuning. The

auto-tuning of the traffic balancing algorithm considerably improves the CSR and hence the overall network capacity.

Figure 6.7 presents the average throughput of the network as a function of traffic intensity, for the two scenarios, with and without auto-tuning. The auto-tuning process increases the overall network throughput and hence the network profitability. By correctly balancing the network load, and by adapting the algorithm to the network traffic, both CSR and throughput are improved.

Define the VHO execution rate as the ratio between the number of VHO triggered to the total number of mobiles admitted to the network; and the VHO success rate as the ratio between the number of successful VHO to the number of triggered VHO, both during the entire simulation. We present in figures 6.8 and 6.9 the execution rate and the success rate of VHO from UMTS to WLAN, respectively. The initial network has an excess of VHO execution rate, including ping-pong handovers. The auto-tuning process reduces the number of VHOs and improves the success rate of the executed ones. So, the auto-tuning process considerably avoids unnecessary VHO generated by the proposed load balancing algorithm.

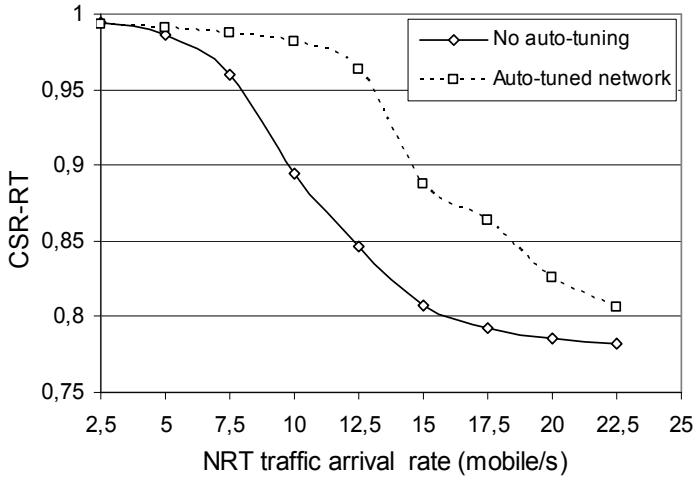


Figure 6.5. Call success rate of RT traffic as a function of NRT traffic arrival rate for the network with (squares) and without (diamonds) auto-tuning.

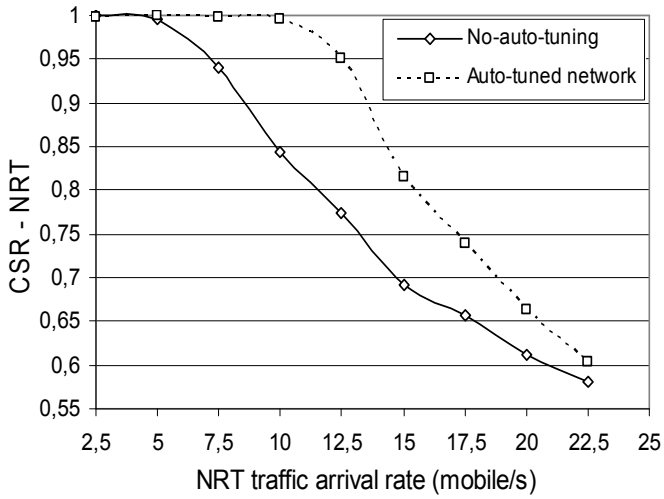


Figure 6.6. Call success rate for NRT traffic as a function of NRT traffic arrival rate for the network with (squares) and without (diamonds) auto-tuning.

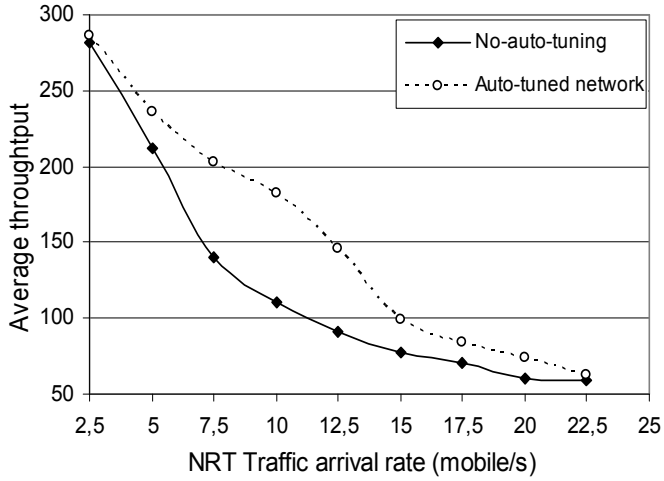


Figure 6.7. Average throughput as a function of NRT traffic intensity.

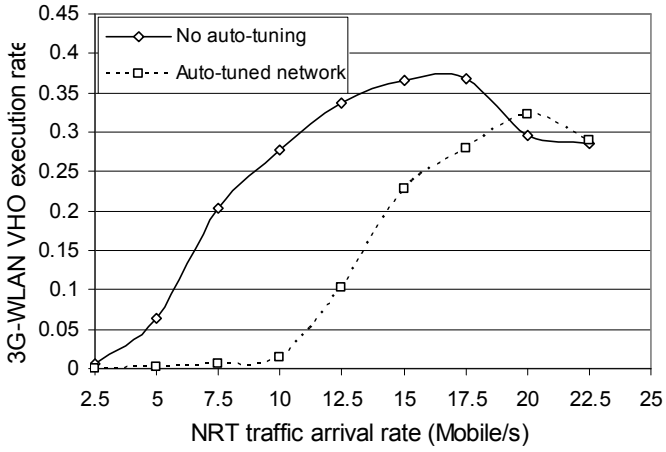


Figure 6.8. Impact of auto-tuning on the execution rate of UMTS to WLAN vertical handovers.

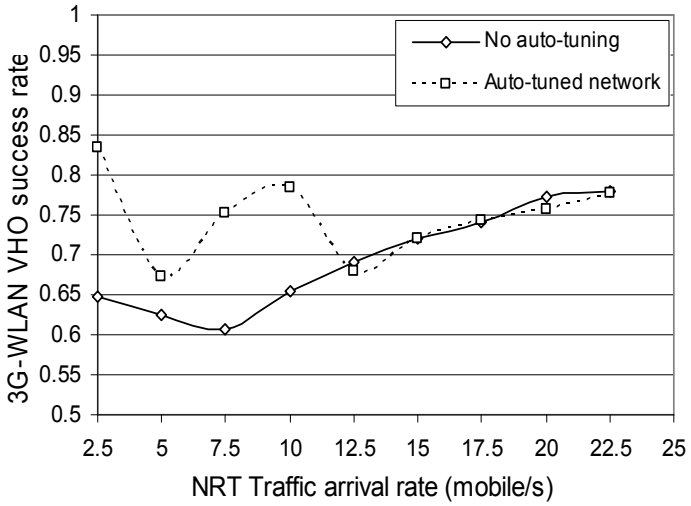


Figure 6.9. Impact of auto-tuning on the success rate of UMTS to WLAN vertical handovers.

6.5 Conclusion

This chapter has presented a load balancing algorithm between a UMTS network and WLAN access points, and its dynamic auto-tuning. The auto-tuning process is performed using a fuzzy logic controller that is optimized using a fuzzy Q-learning algorithm. The controller adapts the UMTS load threshold of the VHO algorithm for each BS according to quality indicators from the BS and its neighbouring WLAN access points. Simulation results show that the auto-tuned traffic balancing algorithm controls the amount of VHOs between the two sub-systems, and prevent unnecessary handovers. It considerably increases the overall call success rate for both RT and NRT traffic.

Chap. 7 Conclusions and perspectives

7.1 Conclusions

In this dissertation, a state of the art for multi-system auto-tuning and self-optimization is presented. An overview of each technology and the target parameters that could be auto-tuned is given to define the different scenarios and use-cases considered in the thesis. The given overview includes also multi-system management, and the advanced (or joint) radio resource management. The auto-tuning architecture has been described for both online and off-line mode of operation. The requirements for efficient auto-tuning architecture are highlighted by pointing out the relation between network entities and the auto-tuning engine. An example of signalling messages between the network and the auto-tuning and optimization engine (AOE) has been described. The composition of the auto-tuning and optimization module is presented with a special emphasis on used auto-tuning tools. In this thesis we have mainly used the fuzzy Q-learning algorithm and that is why we have described in depth Fuzzy Logic Control (FLC) and reinforcement learning. The fuzzy logic controller is shown to be a simple and effective framework for designing a controller that orchestrates the auto-tuning process. In the design of this controller, engineering rules are presented in terms of linguistic predicates that are directly translated into mathematical form. The optimization of the controller is shown to be essential to guarantee a high quality auto-tuning process, and may be required when the conditions of utilization of the controller change, i.e. network environment or traffic composition. The Q-learning implementation of the Reinforcement Learning does not require a system model and is suitable to fully automate the FLC design process and to optimize network parameters. In order to efficiently auto-tune network parameters, the FQLC requires decorrelated input indicators that are related to the parameters, targets for auto-tuning.

The auto-tuning of resource allocation in UMTS network is studied. In this use-case, the RT guard band is dynamically adapted to achieve optimal tradeoffs between QoS of RT and NRT users. Simulation results have shown an efficient compromise between the perceived QoS of RT and NRT services, especially when the traffic is unbalanced. However, when the traffic of both services is very high, the auto-tuning does not improve the QoS for both services. The design and the optimization of the UMTS SHO algorithm have been described in details. Two parameters are optimally and dynamically adapted, namely the SHO *addwin* and *dropwin*. The proposed auto-tuning algorithm utilizes downlink load of a cell and of its neighbours as input indicators and has shown to be simple and effective. The auto-tuning brings about a capacity increase of around 30 percents for a network in a dense urban environment. This example illustrates the importance of auto-tuning in UMTS engineering and has motivated us to investigate self-optimizing handover algorithm in the 3GPP LTE system.

LTE handover optimization is tackled using a predefined auto-tuning function instead of the FQLC employed in the UMTS mobility use-case. A GSM-like hard HO algorithm governed by a matrix handover margin is considered. A margin relates each cell with one neighbour and is responsible for controlling traffic flux between them. Simulation results show that the

optimization of handover margin in a LTE network can improve basic KPIs, namely blocking rates and system throughput by a few percent. The optimization of the handover margin balances the traffic between the eNBs in the network and spatially smoothes interferences in the network.

The last contribution of the thesis deals with the dynamic adaptation of an inter-system vertical handover algorithm. The VHO algorithm is jointly based on the coverage and the system load. Like the previous mentioned use-cases, the VHO is governed by a set of thresholds. Using FQLC, the auto-tuning adapts VHO thresholds that control traffic flux between UMTS and WLAN systems and perform traffic balancing. The WLAN access technology is assumed to be completely managed by the UMTS RNC. A FQLC having double outputs has been formulated and has shown to be well adapted to the optimization of the auto-tuning process when several network subsystems coexist.

7.2 Limitations and perspectives

Several aspects related to the auto-tuning and self-optimization processes need further investigation. Auto-tuning impacts system signalling and hence it is worth making a best trade-off between the capacity gain and the signalling overhead introduced by the auto-tuning process. The auto-tuning may require new supplementary channels in addition to the existent signalling channels reserved for users in the radio and core network.

Although most of the results obtained from the simulations and case studies carried out in this thesis have been very useful to illustrate auto-tuning mechanisms influencing the network performance, these results provides trends. This has various reasons. Some uncertainties and inaccuracies have been introduced during the modelling phase. Limitations of the used simulation tool, assumptions and approximations made when building the system models affect the degree to which the reality is reflected. Testing on real experimental network is required to provide a rich source of information and clear knowledge of self-optimization behaviour.

In future research, more attention could be focused to multi-criteria self-optimization aspects. In this thesis, we have limited ourselves to the use of mono-objective optimization in the framework of the Q-learning algorithm by aggregating the various criteria. Finally, the problem of simultaneously activating several auto-tuning controllers that adapt different RRM parameters is of particular interest as a mean to further boost the capacity gain. Simulating such auto-tuning processes requires complex system modelling allowing to describe the coordination between the different controllers. Such concept is known as a multi-agent problem.

References

[1] Z. Altman, R. Skehill, R. Barco, L. Moltsen, R. Brennan, A. Samhat, R. Khanafer, H. Dubreil, M. Barry, B. Solana, "The Celtic Gandalf Framework", 13th Mediterranean Electronical Conference, MELECON 2006, May 16-19, 2006. Malaga, Spain.

[2] R. Nasri, A.E. Samhat, Z. Altman, “Procédé de calcul de l'état des éléments du réseau sans fil pour la gestion de ses ressources à partir des indicateurs de qualité”, Brevet No 06529.

[3] Orange, "Self-optimization use-case: self-tuning of handover parameters", 3GPP TR3-071262.

[4] R. Nasri, Z. Altman and H. Dubreil, “Fuzzy-Q-learning-based autonomic management of macro diversity algorithms in UMTS networks,” Annals of Telecommunications, Vol. 61, N°9-10, Septembre - octobre 2006.

[5] H. Dubreil, R. Nasri, Z. Altman, “Ingénierie Automatique des Réseaux mobiles,” chapitre dans le livre "L'autonomie dans les réseaux", Traité Hermès IC2, Septembre 2006.

[6] Z. Altman, H. Dubreil, R. Nasri, O.B. Amor, J.M. Picard, V. Diascorn and M. Clerc “Auto-tuning of RRM parameters in UMTS networks,” book chapter in “Understanding UMTS Radio Network Modelling, Planning and Automated Optimization: Theory and Practice,” Wiley & Sons 2006.

[7] R. Nasri, Z. Altman, "Handover adaptation for dynamic load balancing in 3GPP Long Term Evolution Systems", 5th International Conference on Advances in Mobile Computing & Multimedia (MoMM2007), Jakarta, December 2007.

[8] R. Nasri, A.E. Samhat, Z. Altman "A new approach of UMTS-WLAN load balancing; algorithm and its dynamic optimization", IEEE WoWMoM Workshop on Autonomic Wireless Access 2007 (IWAS07), Helsinki, Finland, June 18th, 2007.

[9] A.E. Samhat, R. Nasri, Z. Altman, "Joint Mobility RRM Algorithms for Heterogeneous Access Networks", 13th European Wireless Conference (EWC2007), Paris, France, April 1st 2007.

[10] R. Nasri, Z. Altman, H. Dubreil, "Optimal tradeoff between RT and NRT services in 3G-CDMA networks using dynamic fuzzy Q-learning", IEEE PIMRC’06, Helsinki, 11-14 Sept., 2006.

[11] R. Nasri, Z. Altman and H. Dubreil, “WCDMA downlink load sharing with dynamic control of soft handover parameters,” IEEE VTC2006 Spring, Melbourne, Australia, 7-10 Mai 2006.

[12] 3GPP TS 05.08 V8. 11.0 "Radio Access Network; Radio subsystem link control", (Release 1999), 2001.

[13] 3GPP TS 03.22, "Digital cellular telecommunications system (Phase 2+); Functions related to Mobile Station (MS) in idle mode and group receive mode".

[14] 3GPP TS 23.034, "High Speed Circuit Switched Data (HSCSD) - Stage 2", (Release 1999), 2000-12

[15] 3GPP TS 23.060 v 5.2.0, "General Packet Radio Service (GPRS); Service description; Stage 2" (Release 5), June 2002.

[16] X. Lagrange, P. Godlewski, S. Tabbane, "Réseaux GSM-DCS, des principes à la norme", Hermès Science publication, 1999.

[17] 3GPP Technical Specification 25.401 "UTRAN Overall Description".

[18] M. Nawrocki, M. Dohler, H. Aghvami, "Understanding UMTS Radio network modelling, planning and automated optimization: theory and practice" john Wiley & sons, 2006.

[19] H. Holma and al., "WCDMA for UMTS", john Wiley & sons 2004, third edition.

[20] 3rd Generation Partnership Project (3GPP) website, http://www.3gpp.org.

[21] 3GPP TS 23.107, "Quality of Service (QoS) concept and architecture", Release 6.

[22] 3GPP TR 25.855: "High Speed Downlink Packet Access (HSDPA): Overall UTRAN Description".

[23] H. Holma, A. Toskala, "HSDPA/HSUPA for UMTS: High Speed Radio Access for Mobile Communications", John Wiley & Sons 2006.

[24] 3GPP TR 25.922 V6.0.1, "Radio resource management strategies", (Release 6), 04-2004.

[25] G. Edwards, R. Sankar, "Handoff using fuzzy logic," IEEE GlobeCom, Singapore (November 1995) pp. 520-524.

[26] G. Edwards, R. Sankar, "Microcellular Handoff Using Fuzzy Techniques," Wireless Networks, Vol. 4, No. 5, pp. 401-409, 1998.

[27] P. Magnusson, J. Oom, "An Architecture for self-tuning cellular systems", Proc. of the 2001 IEEE/IFIP International symposium on Integrated Network Management, 2001, pp. 231-245.

[28] P. Gustas, P. Magnusson, J. Oom, N. Storm, "Real-time performance monitoring and optimization of cellular systems", Ericsson Review, n. 1, 2002, pp. 4-13.

[29] S. Choi, K.G. Shin, "A comparative study of bandwidth reservation and admission control schemes in QoS-sensitive cellular networks," Wireless Networks 6(4): 289-305, 2000.

[30] C. Oliveria, J.B. Kim, T. Suda, "An adaptive bandwidth reservation scheme for high-speed multimedia wireless networks," IEEE Journal on Selected Areas in Communications, Vol. 16, No 6, pp. 858-874, 1998.

[31] C. Lindemann, M. Lohmann, A. Thümmler, "Adaptive Call Admission Control for QoS/Revenue Optimization in CDMA Cellular Networks," ACM Journal on Wireless Networks (WINET), Vol 10, pp. 457-472, 2004.

[32] P. Ramanathan et al. "Dynamic Resources Allocation Schemes During Handoff for Mobile Multimedia Wireless Networks" JSAC July 1999

[33] Jiongkuan Hou, "Mobility-based call admission control schemes for wireless mobile networks"; Wireless Communication Mobile Computing, Jul-Sept 2001; Wiley

[34] B.M.Epstein, M Schwartz, "Predicting QoS-Based Admission Control for Multiclass Traffic in cellular Wireless Networks", JSAC March 2000

[35] Tao Zhang et al. "Local Predictive Resource Reservation for Handover Multimedia Wireless IP networks", JSAC 2001, October 2001, pp 1931-1941.

[36] H. Holma, J. Laakso "Uplink Admission Control and Soft Capacity with MUD in CDMA", IEEE VTC Fall 1999, Vol. 1.

[37] 3GPP TS 25.304: "UE Procedures in Idle Mode and Procedures for Cell Reselection in Connected Mode"

[38] 3GPP TS 25.133: "Requirements for Support of Radio Resource Management (FDD)".

[39] 3GPP TS 25.123: "Requirements for Support of Radio Resource Management (TDD)".

[40] 3GPP TS 23.122: "NAS functions related to Mobile Station (MS) in idle mode ".

[41] R. Guerzoni, I. Ore, K. Valkeahlati D. Soldani, "Automatic Neighbor Cell List Optimization for UTRA FDD Networks: Theoretical Approach and Experimental Validation", WPMC2005.

[42] K. Valkealahti, A. Höglund, J. Parkkinen, A. Flanagan, "WCDMA Common Pilot Power Control with Cost Function Minimization", VTC-Fall 2002, September 2002.

[43] A. Hämäläinena, K. Valkealahtia, A. Höglunda, J. Laaksob, "Auto-tuning of Service-specific Requirement of Received EbNo in WCDMA", VTC-Fall 2002, September 2002.

[44] A. Höglund, K. Valkealahti, "Quality-based Tuning of Cell Downlink Load Target and Link Power Maxima in WCDMA", VTC-Fall 2002, September 2002.

[45] J. A. Flanagan, T. Novosad, "WCDMA Network Cost Function Minimization for Soft Handover Optimization with Variable User Load", VTC-Fall 2002, September 2002.

[46] B. Homnan, W. Benjapolakul, "QoS-controlling soft handoff based on simple step control and a fuzzy inference system with the gradient descent method", IEEE Transactions on Vehicular Technology, Vol. 53, pp. 820-834, May. 2004.

[47] V. Kunsriraksakul, B. Homnan, W. Benjapolakul, "Comparative evaluation of fixed and adaptive soft handoff parameters using fuzzy inference systems in CDMA mobile communication systems"**,** IEEE VTS 53rd Vehicular Technology Conference, 2001. VTC 2001 Spring. Volume 2, 6-9 May 2001 Page(s):1017 - 1021 vol.2.

[48] J. Ye, X. Shen, J.W. Mark, "Call Admission Control in Wideband CDMA Cellular Networks by using Fuzzy Logic," IEEE Transactions on Mobile Computing, Vol 4, No2, pp. 129- 141, April. 2005.

[49] P. Dini, S. Guglielmucci, "Call admission control strategy based on fuzzy logic for WCDMA systems, IEEE International Conference on Communications, June 2004, Page(s):2332- 2336.

[50] R.N.S. Naga, D. Sarkar, "Call admission control in mobile cellular CDMA systems using fuzzy associative memory", IEEE International Conference on Communications, June 2004, Page(s):4082 - 4086 Vol.7.

[51] H. Dubreil, Z. Altman, V. Diascorn, J.M. Picard, M. Clerc, "Particle Swarm optimization of fuzzy logic controller for high quality RRM auto-tuning of UMTS networks," IEEE International Symposium VTC 2005, Stockholm, Sweeden, 29 May-1 June 2005.

[52] H. Dubreil, "Méthodes d'optimization de contrôleurs de logique floue pour le paramétrage automatique des réseaux mobiles UMTS", thèse de doctorat, ENST 2005.

[53] L. Jouffe, "Fuzzy Inference System Learning by reinforcement Methods", IEEE Transactions on Systems, Man, and Cybernetics, Vol. 28, pp. 338-355, Aug. 1998.

[54] S.M. Senouci, A.L. Beylot, G. Pujolle, "Call admission control in cellular networks: a reinforcement learning solution", International Journal of Network Management, No. 14, pp 89-103, 2004.

[55] F. Yu, V.W.S. Wong, V.C.M. Leung, "Efficient QoS Provisioning for Adaptative Multimedia in Mobile Communication Networks by reinforcement Learning," First International Conference on Broadband Networks, BROADNETS'04 IEEE, 2004.

[56] F. Yu, V.W.S. Wong, V.C.M. Leung, “A New QoS Provisioning Method for Adaptive Multimedia in Cellular Wireless Networks,” IEEE Conference on Computer Communications (INFOCOM’04), Hong Kong, China, Mar. 2004.

[57] Y. S. Chen, C. J. Chang, F. C. Ren, "A Q-learning-based multi-rate transmission control scheme for RRM in multimedia WCDMA systems," IEEE Transactions on Vehicular Technology, Vol. 53, No. 1, pp. 38-48, Jan. 2004.

[58] 3GPP TR 25.881 V5.0.0, “Improvement of RRM across RNS and RNS/BSS”, 2001-12.

[59] 3GPP TR 25.891 V0.3.0, “Improvement of RRM across RNS and RNS/BSS (Post Rel-5),” Release 6, 2003/2.

[60] 3GPP TR 22.934 V6.2.0, "Feasibility study on 3GPP system to Wireless Local Area Network (WLAN) interworking".

[61] 3GPP TR 22.234, "Requirements on 3GPP system to Wireless Local Area Network (WLAN) interworking"

[62] 3GPP TR 23.234, "3GPP system to Wireless Local Area Network (WLAN) interworking; System description"

[63] A. K. Salkintzis, “Interworking Techniques and Architectures for WLAN/3G Integration Towards 4G mobile Data Networks,” IEEE Wireless Communications, June 2004, pp 50-61.

[64] J. Perez-Romero et al, "Common radio resource management: functional models and implementation requirements", PIMRC 2005.

[65] M. K. Starr, M. Zeleny, "Multiple Criteria Decision-Making", Mc.Graw-Hill.

[66] C.L. Hwang, K. Yoon, "Multiple Attribute Decision Making", Springer-Verlag, Berlin, 1981.

[67] R. Steuer, "Multiple Criteria Optimization: Theory, Computation and Application", , John Wiley & Sons, 1986.

[68] K. Shum and C.W. Sung, "Fuzzy Layer Selection Method in Hierarchical Cellular Systems”, IEEE Trans. Vehicular Technology, vol. 48, pp. 1840-1849, Nov. 1999.

[69] N.D. Tripathi, J.H. Reed, J. H. VanLandingham, "Adaptive handoff algorithms for cellular overlay systems using fuzzy logic", Proc. 49th IEEE Vehicular Technology Conference, vol. 2, 1999, pp. 1413 – 1418.

[70] A. Majlesi, B.H. Khalaj, "An Adaptive Fuzzy Logic Based Handoff Algorithm For Interworking Between Wlans And Mobile Networks", Proc. IEEE 49th PIMRC, pp. 2446-2451, 2002.

[71] P.M.L. Chan, R.E. Sheriff, Y.F. Hu, P. Conforto C. Tocci, "Mobility Management Incorporating Fuzzy Logic for a Heterogeneous IP Environment", IEEE Communication Magazine, pp. 42-51, Dec. 2001.

[72] P.M.L. Chan, Y.F. Hu, R.E. Sheriff, "Implementation of Fuzzy Multiple Objective Decision Making Algorithm in a Heterogeneous Mobile Environment", Proc. IEEE Wireless Communications and Networking Conference, Orlando, FL, USA, 17-21, pp. 332-336, March 2002.

[73] R. Agusti, O. Sallent, J. Perez-Romero, L. Giupponi, "A fuzzy-neural based approach for joint radio resource management in a beyond 3G framework", Proc. Of 1st Quality of Service in Heterogeneous Wired/Wireless Networks conference, pp. 216 – 224, Oct. 2004.

[74] L. Giupponi, J. Perez-Romero, R. Agusti, O. Sallent, "A Novel Joint Radio Resource Management Approach with Reinforcement Learning Mechanisms", IEEE International Workshop on Radio Resource Management for Wireless Cellular Networks, Apr. 2005.

[75] W. Zhang, "Handover Decision Using Fuzzy MADM in Heterogeneous Networks", IEEE WCNC 2004, Atlanta, pp. 653-658, March 2004.

[76] R.R. Yager, Multiple Objective Decision Making using Fuzzy Sets, International Journal Man Machine Studies, no. 9, 1977, pp. 375-382.

[77] C.T. Lin, C:S.G. Lee, Neural-Network-Based Fuzzy Logical Control and Decision System, IEEE Trans. Computers, vol. 40, no. 12, pp. 1320-1336, Dec. 1991.

[78] Mamdani, E.H. Applications of fuzzy logic to approximate reasoning using linguistic synthesis. IEEE Transactions on Computers, Vol. 26, No. 12, pp. 1182-1191, 1977

[79] Sugeno, M. Industrial applications of fuzzy control. Elsevier Science Pub. Co., 1985

[80] Astrom, K.J.; Wittenmark, B. Adaptive control, 2nd Ed. Addison-Wesley 1995

[81] Jantzen, J. A tutorial on fuzzy adaptive control. Proc. EUNITE 2002

[82] P. Y. Glorennec, Apprentissage par renforcement et logique floue, in Actes des rencontres francophones sur la logique floue et ses applications (LFA'01), November 2001.

[83] A Bayesian approach for automated troubleshooting for UMTS networks", *the 17th Annual IEEE Intern. Symp. PIMRC'06*, Helsinki, 11-14 Sept., 2006.

[84] Estimating GPRS Link Bit Rates in TEMS Investigation, Ericson white paper, 2000.

[85] Gandalf Deliverable D3.2, "Overview and Definition of the Global Architecture for Multi-System Self-Tuning", September 2005.

[86] K. Tomsocic, M.Y. Chow, "Tutorial on Fuzzy Logic: Applications in Power Systems", IEEE-PES Winter meeting in Singapore, January 2000

[87] J.M. Mendel, "Fuzzy logic systems for engineering: a tutorial" Proceedings of the IEEE, V83, n3, Mar 1995, pp. 345 –377.

[88] S. R. Sutton and A.G. Barto, " Reinforcement Learning: An Introduction", MIT Press, Cambridge, MA, 1998.

[89] H.S. Chang, M.C. Fu, J. Hu, S.I. Marcus, " Simulation-based algorithms for Markov decision processes", Springer, 2007.

[90] S.P. Meyn, R.L. Tweedie , "Markov Chains and Stochastic Stability", Springer-Verlag, 1993.

[91] T. Jaakkola, M.I. Jordan, S.P. Singh, "On the Convergence of Stochastic Iterative Dynamic Programming Algorithms", Neural Computation, 6(6):1185-1201, 1994.

[92] C.J. Watkins, "Learning from Delayed Rewards", PhD Thesis, University of Cambridge, England, 1989.

[93] B. J. Prabhu, E. Altman, K. Avrachenkov, J. A. Dominguez, "A simulation study of TCP performance over UMTS downlink," in IEEE VTC 2003.

[94] A. Rebai, "Conception et développement d'un simulateur UMTS," training report, SupCom, 2003.

[95] Sébastien Baret and al., "Analysis of signalling procedures for end-to-end QoS modelling in UMTS" technical report July 2004.

[96] 3GPP TR 25.913, V7.1.0 Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN), (Release 7), (2005-09).

[97] 3GPP TS 36.300, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 3 (Release8), Sept. 2007.

[98] E. Dahlman, S. Parkvall, J. Sköld and P. Beming, 3G Evolution, HSPA and LTE for Mobile Broadband, Academic Press, London 2007.

[99] 3GPP TR3-061487 "Self-Configuration and Self-Optimization".

[100] 3GPP TR3-071262, "Self-optimization use-case: self-tuning of handover parameters", Orange.

[101] 3GPP TR3-071438, "Load Balancing SON Use case", Alcatel-Lucent.

[102] H. Ekström, A. Furuskär, J. Karlsson, M. Meyer, S. Parkvall, J. Torsner, and M. Wahlqvist, "Technical Solutions for the 3G Long-Term Evolution," IEEE Commun. Mag., vol. 44, no. 3, March 2006, pp. 38–45.

[103] 3GPP TR 25.814, V7.1.0, "Physical Layer Aspects for evolved Universal Terrestrial Radio Access (UTRA) (Release 7)" (2006-09).

[104] 3GPP TS 36.211, V0.3.1; "Physical Channels and Modulation (Release 8)" (2007-02).

[105] 3GPP TSG RAN WG3, R3-071494, "Automatic neighbour cell configuration", Aug. 2007.

[106] 3GPP TSG RAN WG3, R3-071819, "On automatic neighbour relation configuration", Octobre. 2007.

[107] 3GPP TR3-072191, "Measurements for handover decision use case", NTT DoCoMo, Orange, Telecom Italia, T-Mobile, Telefonica, November 2007.

[108] *ANSI/IEEE Std* 802.11, “Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications”, 1999 Edition.

Appendix A: Convergence proofs of the reinforcement learning algorithm

Proposition 3.1

Let π be a history-dependent policy. For any initial state *x*, there exists a Markov policy π', such that: $V^{\pi'}(x) = V^{\pi}(x)$.

Proof

Let π' be the Markovian policy defined from the history-dependent policy π by:

$$\forall t \in T, \forall s \in S, \forall a \in A \quad q_{\pi'}(a_t = a, s_t = s) = P^{\pi}(a_t = a / s_t = s, s_0 = x)$$

Then

$$\forall t \in T, P^{\pi'}(a_t = a / s_t = s) = P^{\pi}(a_t = a / s_t = s, s_0 = x)$$

We can prove by recurrence over *t* the relation:

$$\forall t \in T, P^{\pi}(s_t = s, a_t = a / s_0 = x) = P^{\pi'}(s_t = s, a_t = a / s_0 = x)$$

By construction, the relation is verified for *t=0*. We assume that the equality is verified until *t-1* and we prove that it remain true for *t*.

$$\begin{aligned} P^{\pi}(s_t = s / s_0 = x) &= \sum_{i \in S} \sum_{a \in A} P^{\pi}(s_{t-1} = i, a_{t-1} = a / s_0 = x) p(s / i, a) \\ &= \sum_{i \in S} \sum_{a \in A} P^{\pi'}(s_{t-1} = i, a_{t-1} = a / s_0 = x) p(s / i, a) \\ &= P^{\pi'}(s_t = s / s_0 = x) \end{aligned}$$

Therefore

$$\begin{aligned} P^{\pi'}(s_t = s, a_t = a / s_0 = x) &= P^{\pi'}(a_t = a / s_t = s) P^{\pi'}(s_t = s / s_0 = x) \\ &= P^{\pi}(a_t = a / s_t = s, s_0 = x) P^{\pi}(s_t = s / s_0 = x) \\ &= P^{\pi}(s_t = s, a_t = a / s_0 = x) \end{aligned}$$

So the proposition is proved.

Remarks

iv) From the previous proposition, we deduce that every history-dependent policy can be replaced by a *Markovian* policy having the same value function if the initial state is given. From now on, we use only the markovian policy, unless contrary mentioned.

v) If the policy is *markovian*, the process (s_t) is itself a *markovian* process with a transition matrix P_{π}, defined by:

$$\forall s, s' \in S \quad P_{\pi s, s'} = \sum_{a \in A} q_{\pi}(a, s) p(s' / s, a)$$

because

$$P^{\pi}(s_{t+1}/s_0,s_1,...,s_t)=\sum_{a\in A}P^{\pi}(a_t=a/s_0,s_1,...,s_t)P^{\pi}(s_{t+1}/s_0,s_1,...,s_t,a_t=a)$$
$$=\sum_{a\in A}q_{\pi}(a,s_t)P^{\pi}(s_{t+1}/s_t,a_t=a)$$
$$=P^{\pi}(s_{t+1}/s_t)$$

vi) According to the previous notation, the value function can be expressed as:

$$V^{\pi}(x)=\sum_{t=0}^{+\infty}E^{\pi}\left[\gamma^t r_t(s_t,a_t)/s_0=x\right]$$
$$=\sum_{t=0}^{+\infty}\sum_{s\in S}\sum_{a\in A}\gamma^t r_t(s,a)P^{\pi}(s_t=s,a_t=a/s_0=x)$$

Theorem 3.1

Let r_{π} be the reward vector whose elements are $\sum_{a\in A}q_{\pi}(a,s)r(s,a)$ and V^{π} (the same notation of the value function) the value vector whose elements are $V^{\pi}(s)$. The size of r_{π} and V^{π} is equal to the number of states. The matrix expression of the value function V^{π} is then:

$$V^{\pi}=(I-\gamma P_{\pi})^{-1}r_{\pi}$$

Proof

We have from the previous remarks that

$$V^{\pi}(s)=\sum_{t=0}^{+\infty}\sum_{s'\in S}\sum_{a\in A}\gamma^t r(s',a)P^{\pi}(s_t=s',a_t=a/s_0=s)$$
$$=\sum_{t=0}^{+\infty}\gamma^t\sum_{s'\in S}\sum_{a\in A}r(s',a)q_{\pi}(a,s')P^{\pi}(s_t=s'/s_0=s)$$
$$=\sum_{t=0}^{+\infty}\gamma^t\sum_{s'\in S}P^{\pi}(s_t=s'/s_0=s)r_{\pi}(s')$$

$P^{\pi}(s_t=s'/s_0=s)$ is the probability of going from state s to state s' in t time steps. Using *Chapman-Kolmogorov equation* [90], that for any t' such that $0 < t' < t$,

$$P^{\pi}(s_t=s'/s_0=s)=\sum_{i\in S}P^{\pi}(s_{t'}=i/s_0=s)P^{\pi}(s_{t-t'}=s'/s_0=i)$$

As the space state is finite, the t-step transition probability is computed as the t'th power of the transition matrix [90]. So $P^{\pi}(s_t=s'/s_0=s)$ is the element of the matrix P_{π}^t.

The expression of the value function becomes then:

$$V^{\pi}(s)=\sum_{t=0}^{+\infty}\gamma^t P_{\pi}^t r_{\pi}(s)$$

Now the matrix γP_{π} is a stochastic matrix and all their eigenvalues have a complex modulus lower than $\gamma < 1$. Therefore the matrix $I-\gamma P_{\pi}$ is invertible and its inverse is

$$\left(I-\gamma P_\pi\right)^{-1}=\sum_{t=0}^{+\infty}\gamma^t P_\pi^t$$

The matrix expression of the value function is then

$$V^\pi(s)=\left(I-\gamma P_\pi\right)^{-1} r_\pi(s)$$

Theorem 3.2: Bellman equation

If S and A are finite sets, then V^* is the unique solution of the equation

$$V=LV$$

Proof

To prove the existence of the solution, we use the *contraction mapping theorem* which involves *Banach* spaces.

Theorem: contraction mapping theorem

Let B be a non-empty *Banach* space (i.e. complete normed vector spaces) and T be a contraction mapping on B (i.e. there is a nonnegative real number $0\le\lambda<1$ such that $\forall x,y\in B \quad \|Tx-Ty\|\le\lambda\|x-y\|$) then

i) The map T admits one and only one fixed point x^* in B (this means $Tx^*=x^*$).

ii) For any starting point $x_0\in B$ the iterative sequence $\{x_n\}$, defined by $x_{n+1}=Tx_n=T^{n+1}x_0$, converges to the point x^*.

The space Ω with the norm max, given in *definition 3.5*, is a complete normed vector space. We now show that the DPO operator is a contraction mapping on Ω. For that, consider U and V two value functions in Ω and s a state in S.

Assume that $LV(s)\ge LU(s)$. Let $a_s^*\in\arg\max_{a\in A}\left(r(s,a)+\gamma\sum_{s'\in S}p(s'/s,a)V(s')\right)$. It follows from the definition of the DPO operator that

$$\begin{aligned}
0\le\left|LV(s)-LU(s)\right| &= LV(s)-LU(s)\\
&= r(s,a_s^*)+\gamma\sum_{s'\in S}p(s'/s,a_s^*)V(s')-\max_{a\in A}\left\{r(s,a)+\gamma\sum_{s'\in S}p(s'/s,a)U(s')\right\}\\
&\le r(s,a_s^*)+\gamma\sum_{s'\in S}p(s'/s,a_s^*)V(s')-r(s,a_s^*)-\gamma\sum_{s'\in S}p(s'/s,a_s^*)U(s')\\
&= \gamma\sum_{s'\in S}p(s'/s,a_s^*)\left(V(s')-U(s')\right)\\
&\le \gamma\sum_{s'\in S}p(s'/s,a_s^*)\|V-U\|\\
&\le \gamma\|V-U\|
\end{aligned}$$

Therefore $\|LV-LU\|=\max_{s\in S}\left|LV(s)-LU(s)\right|\le\gamma\|V-U\|$. The mapping L admits then only one fixed point.

We show now that the fixed point is exactly V^*, defined as $V^*(s) = \max_{\pi \in \Pi} V^\pi(s) \quad \forall s \in S$. We want to show that if $V = LV$ then $V = V^*$.

Let V such that $V \geq LV$. We have then

$$\forall \pi \in \Pi \quad V \geq \max_{\pi \in \Pi} \{r_\pi + \gamma P_\pi V\} \geq r_\pi + \gamma P_\pi V .$$

If we apply the inequality $V \geq r_\pi + \gamma P_\pi V$ over V n times, we obtain

$$\begin{aligned} V &\geq r_\pi + \gamma P_\pi V \\ &\geq r_\pi + \gamma P_\pi (r_\pi + \gamma P_\pi V) = (I + \gamma P_\pi) r_\pi + \gamma^2 P_\pi^2 V \\ &\geq .... \\ &\geq \sum_{k=0}^{n-1} \gamma^k P_\pi^k r_\pi + \gamma^n P_\pi^n V \end{aligned}$$

The right hand term equals $V^\pi - \sum_{k=n}^{+\infty} \gamma^k P_\pi^k r_\pi + \gamma^n P_\pi^n V$. Therefore

$$V - V^\pi \geq \gamma^n P_\pi^n V - \sum_{k=n}^{+\infty} \gamma^k P_\pi^k r_\pi$$

The term $\gamma^n P_\pi^n V - \sum_{k=n}^{+\infty} \gamma^k P_\pi^k r_\pi$ tends to 0 for n large enough because $\left\| \gamma^n P_\pi^n V \right\| \leq \gamma^n \left\| V \right\|$ and $\left\| \sum_{k=n}^{+\infty} \gamma^k P_\pi^k r_\pi \right\| \leq \frac{\gamma^n}{1-\gamma} \max_{s \in S, a \in A} \{r(s,a)\}$.

This leads to the inequality $V \geq V^\pi$ for every $\pi \in \Pi$, and in particular for $\pi^* = \arg\max_{\pi \in \Pi} V^\pi$.

Hence, $V \geq LV \Rightarrow V \geq V^* = \max_{\pi \in \Pi} V^\pi$

Conversely, consider V such that $V \leq LV$. Let π^* be the policy that maximizes the value function over all policies. We have then

$$\begin{aligned} V &\leq r_{\pi^*} + \gamma P_{\pi^*} V \\ &\leq r_{\pi^*} + \gamma P_{\pi^*} (r_{\pi^*} + \gamma P_{\pi^*} V) \\ &\leq .... \\ &\leq \sum_{k=0}^{n-1} \gamma^k P_{\pi^*}^k r_{\pi^*} + \gamma^n P_{\pi^*}^n V \end{aligned}$$

Then

$$V - V^{\pi^*} \leq \gamma^n P_{\pi^*}^n V - \sum_{k=n}^{+\infty} \gamma^k P_{\pi^*}^k r_{\pi^*}$$

The same way as in the previous case, the hand right term converges to 0 for n large enough.

Hence $V \leq LV \Rightarrow V \leq V^*$.

By combining both cases, we obtain that $V = LV \Rightarrow V = V^*$: every solution of the *Bellman equation* is necessarily equals the optimal value function V^*.

Appendix B: LTE interference model

Starting from the interference coordination scheme, presented in section 5.3.2, we assume that the spectral band is composed of C resource blocks, one third of the band is reserved for the cell edge users and the rest is for cell centre users.

The resource allocation is made according to users' positions. Let L_{th} be the path loss threshold separating users from the two different sub-bands. A user with a path loss higher than L_{th} is granted resource blocks in the cell-edge band and otherwise he gets resources in the cell centre band. The eNB transmit power in each cell-edge resource block equals the maximum transmit power P. To reduce intercell interference, the eNB transmit power in the cell-centre band must be lower than P. Let εP (where $\varepsilon < 1$) be the transmit power in the cell-centre band.

The interference should be determined for two different users according to their positions: the cell-centre user and the cell-edge user. Let m_c and m_e be two users connected to a cell k. the mobile m_c uses the central band whereas m_e uses the cell-edge band. Let Λ denote the interference matrix between cells, where the coefficient $\Lambda(i,j)$ equals 1 if cells i and j use the same cell-edge band and zero otherwise.

For cell-edge user m_e, the interference comes from users in the cell centre of the closest adjacent cells and from the cell-edge user in other cells. The mobile m_e connected to the cell k and using one resource block in the cell-edge band, receives an interfering signal from a cell i equals

$$I_{i,m_e} = \left(\left(1 - \Lambda\left(k,i\right)\right)\beta_i^c \varepsilon P_i + \Lambda\left(k,i\right)\beta_i^e P_i\right)\frac{G_{i,m_e}}{L_{i,m_e}} \qquad \text{(B.1)}$$

where P_i is the downlink transmit power per resource block of the cell i. G_{i,m_e} and L_{i,m_e} are respectively the antenna gain and the path loss between cell i and the mobile station m_e. The factor β_i^c (respectively β_i^e) is the probability that the same resource block in the cell-centre band (respectively the cell-edge band) is used at the same time by another mobile connected to the cell i.

Since the analysis considers a long time scale (of the order of seconds), the interference is averaged. So, the factor β_i^c is the percentage of users using cell-centre band and β_i^e is the percentage of those using cell-edge band

$$\beta_i^c = \frac{\#occupied\ resource\ blocks\ in\ cell\ center\ band}{total\ capacity\ of\ cell\ center\ band} = \frac{M_c}{2C/3}$$

$$\beta_i^e = \frac{\#occupied\ resource\ blocks\ in\ cell\ edge\ band}{total\ capacity\ of\ cell\ edge\ band} = \frac{M_e}{C/3}$$

M_c and M_e are the number of resource blocks used in the cell centre and cell edge respectively, and the sum $M_c + M_e$ is the total number of resource blocks used in the cell.

Let χ_i be the load of cell i given by

$$\chi_i = \frac{M_c + M_e}{C} \tag{B.2}$$

Define the factor α_i as the proportion of traffic served in the cell-edge band, $\alpha_i = M_e/(M_c + M_e)$. The factors β_i^c and β_i^e become respectively

$$\beta_i^c = \frac{3(1-\alpha_i)(M_c + M_e)}{2C} = \frac{3(1-\alpha_i)\chi_i}{2}$$

$$\beta_i^e = \frac{3\alpha_i(M_c + M_e)}{C} = 3\alpha_i\chi_i$$

Remark

It is very hard to find an exact expression of the factor α_i. However, based on the assumptions that:

i. the cell is approximated by a circle and the eNB is located in the cell centre,

ii. the traffic is uniformly distributed in the cell and

iii. the effect of shadow fading is neglected

the factor α_i can be approximated by

$$\alpha_i \approx \max\left(\frac{1}{3}, 1 - \left(\frac{L_{th}}{L_{\max}}\right)^{2/\gamma}\right)$$

where L_{max} is the maximum path loss defining the cell surface.

Proof

Assume that a cell can serve users distributed uniformly in a circle with a radius R_{max}. The radius of the inner circle served by the resource blocks of the cell-centre band is denoted R_{th}. Then R_{max} and R_{th} are determined by the propagation model as:

$$R_{\max} = \left(\frac{L_{\max}}{l_0}\right)^{1/\gamma}; \quad R_{th} = \left(\frac{L_{th}}{l_0}\right)^{1/\gamma}.$$

With the uniform traffic assumption and if we don't take into account the limit of physical resources (maximum of one third of the capacity can be assigned to the cell-edge users), the factor α_i is the ratio between the area of the cell-edge and the total area of the cell. Hence

$$\alpha_i \approx 1-\left(\frac{R_{th}}{R_{\max}}\right)^2 = 1-\left(\frac{L_{th}}{L_{\max}}\right)^{2/\gamma}$$

In the following, α_i is calculated numerically using the relation $\alpha_i = M_e/(M_c + M_e)$ as mentioned above. Based on the previous assumptions, the expression of equation (5.1) becomes

$$I_{i,m_e} = \frac{3\chi_i P_i G_{i,m_e}}{L_{i,m_e}}\left((1-\Lambda(k,i))\frac{(1-\alpha_i)}{2}\varepsilon + \Lambda(k,i)\alpha_i\right) \qquad \text{(B.3)}$$

The term $\tilde{\Lambda}_e(k,i) = 3\left((1-\Lambda(k,i))\frac{(1-\alpha_i)}{2}\varepsilon + \Lambda(k,i)\alpha_i\right)$ can be interpreted as a new interference matrix, denoted here as the *fictive interference matrix* for cell-edge users. The total interference perceived by user m_e is the sum of all interfering signals

$$I_{m_e} = \sum_{i\neq k}\tilde{\Lambda}_e(k,i)\chi_i\frac{P_i G_{i,m_e}}{L_{i,m_e}} \qquad \text{(B.4)}$$

For the cell-centre user m_c, the interference comes from users in the cell-edge and cell-centre of closest adjacent cells and also from the cell-centre and cell-edge users in other cells. Similarly to the cell-edge users, the mobile m_c connected to the cell k and using one resource block in the cell-centre band, receives an interfering signal from a cell i equal to

$$I_{i,m_c} = \left(\tfrac{1}{2}(1-\Lambda(k,i))\left(\beta_i^c\varepsilon P_i + \beta_i^e P_i\right) + \Lambda(k,i)\beta_i^c\varepsilon P_i\right)\frac{G_{i,m_c}}{L_{i,m_c}} \qquad \text{(B.5)}$$

The total interference is then

$$I_{m_c} = \sum_{i\neq k}\tilde{\Lambda}_c(k,i)\chi_i\frac{P_i G_{i,m_e}}{L_{i,m_e}} \qquad \text{(B.6)}$$

where the fictive interference matrix in the cell-centre band is given by

$$\tilde{\Lambda}_c(k,i) = 3\left(\tfrac{1}{2}(1-\Lambda(k,i))\left(\frac{1-\alpha_i}{2}\varepsilon + \alpha_i\right) + \Lambda(k,i)\frac{1-\alpha_i}{2}\varepsilon\right) \qquad \text{(B.7)}$$

Appendix C: Network system level simulator

The multi-system simulator architecture is depicted in figure C.1, which shows the main blocks representing the system functionalities and their interactions. The simulation involves cooperation between these blocks according to the investigated scenarios and configurations. In a multi-system scenario, JRRM interacts with the simulator core and with the RRM of each system or RAN (Radio Access Network). In such a scenario, the auto-tuning can be applied utilizing filtered KPIs. The traffic generation and the simulator core and at least one of the RRM modules are required to run a simulation with one RAN. The auto-tuning process may be applied to one system and in this case, the JRRM functionalities are neutral, i.e. the JRRM block is transparent. The typical arrival processes, as well as the packet length distributions of the traffic, are supported by the traffic generator block. In addition, the traffic generation module can be fed by measurement using observation tools.

An object-oriented architecture is utilized to implement the simulator blocks. Generic objects are developed with flexible extensibility in each system. In addition, each module or algorithm can be easily replaced by an equivalent module: For example the admission control algorithm in a system can be replaced by another admission control algorithm without modifying the rest of the simulator. This feature is particularly attractive for testing new RRM algorithms and mobility schemes.

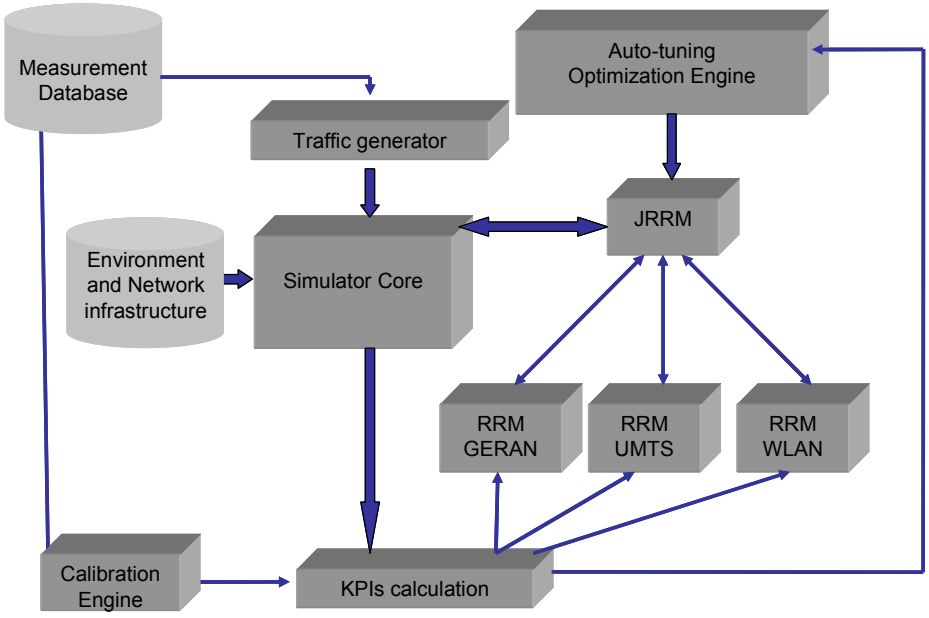


Figure C.1. Main blocs of the multi-system simulator architecture

The semi-dynamic simulator allows at monitoring the time evolution of the network. To achieve a fast computation time required in the dynamic paradigm, the simulator performs correlated snapshots to account for the time evolution of the network with a time resolution, T_c, of the order

of a second. A general scheme describing the network evolution between two correlated snapshots is depicted in figure C.2. At each snapshot, the new user positions (due to intra- and inter-RAN mobility with different speeds) are determined, and quality indicators are computed as in static simulator (powers, interference, etc). Between two snapshots, arrival and departure of users can occur and the corresponding station loads, necessary for admission control, are updated. The network performance statistics for each RAN is determined every T_c.

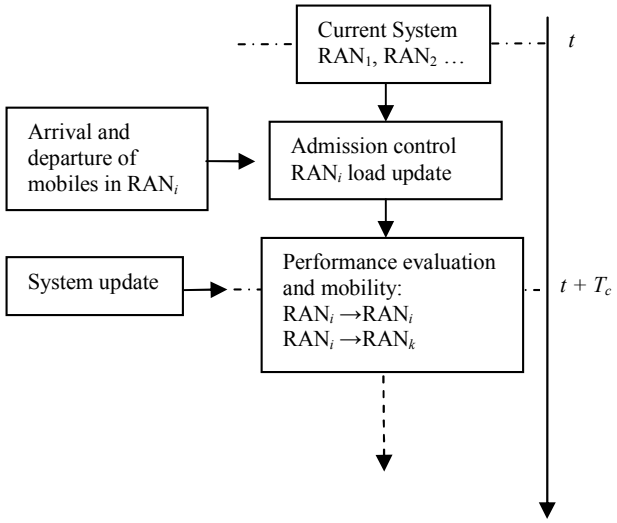


Figure C.2. Time evolution of the multi-system simulator

At the end of each time interval, the simulator performs the following operations for each MS:

- The traffic transmitted in the network is computed
- New mobile position is determined in a meshed surface network. The MS movement is implemented according to several scenarios covering typical mobile velocities and mobile orientations.
- Radio conditions of each mobile are updated (SIR, Power, SNR, etc.)
- Mobility between systems is executed if requested.

Between two time steps, i.e. during the period T_c, the following events occur:

- Arrival and departure of mobiles.
- Execution of the CAC algorithm at each MS arrival.
- Recording the volume of the traffic transmitted for the outgoing mobile.

For the UMTS sub-system simulator, the following procedures are implemented:

- CAC: This algorithm estimates the load increase (both in UL and DL) that the acceptance of a new mobile will cause in the radio network. A Target Load (TL) is specified so that when the network load exceeds TL, it allows no more admissions in

order to preserve a good quality for mobiles which are already in the network. The resources are shared between real-time (RT) and non-real-time (NRT) traffic with different implementations:

- ✓ A reserved band is allocated to RT traffic and a second one is shared between RT and NRT traffic,
- ✓ Two independent bands are reserved for RT and NRT.

- Load control: To decide whether the network is in congestion or not, a criterion is introduced based on a load threshold defined above the TL. If the load of the network increases above this threshold for certain duration, new mobiles are blocked and certain interfering mobiles are dropped by the load control mechanism.
- Power control: The power control is implemented by limiting the power of the mobile to the minimum power required to maintain a given SIR for required system performance. Therefore, the power of each mobile is calculated so that the transmitted power reaches the target SIR.
- Macro-diversity: the soft handover mechanism is implemented in the simulator based on the 3 events: event1A, event1B and event1C.
- Selection and re-selection: the algorithms for the selection/ reselection procedures implemented in the simulator are those specified in the 3GPP standards for 3G networks and described in the second chapter of the thesis.

With respect to WLAN sub-system simulator, basic RRM algorithms are implemented. Other mechanisms are also developed for the inter-system handover, i.e. WLAN-UMTS inter-system mobility and admission control, in the case of non real time traffic. The basic RRM algorithms are:

- Selection of modulation scheme: this algorithm, based on the SNR value, selects the modulation so that the packet error rate remains below a given threshold.
- CAC: when a new MS seeks access to the network, the algorithm estimates the network load and the throughput to be allocated to satisfy this request. Based on this expected throughput, the MS is either admitted or blocked out.
- Physical bit rate adaptation: after admission to the system, the radio conditions of a MS vary due to mobility. This is seen in terms of the fluctuations of the SNR value at each transmission. To maximize the system performance and to avoid the QoS degradation of other MSs, the bit rate adaptation algorithm is executed to adapt the physical bit rate of the MS, i.e. the modulation scheme.