



AROMA White Paper

RADIO ACCESS TECHNOLOGY SELECTION IN HETEROGENEOUS WIRELESS NETWORKS: AROMA'S VIEW

Abstract

The inherent human needs of independence and flexibility, leading to the 'connected everywhere, anytime, anyhow' philosophy applied to the mobile communications sector, are one of the forces driving the evolution of this sector and explain why it will continue to be one of the most dynamic technological drivers within comparative industries in the forthcoming years. In this evolution, the heterogeneous network notion arises as a flexible architecture capable of managing a large variety of coexisting Radio Access Technologies (RATs) along with applications and services comprising different Quality of Service (QoS) demands and protocol stacks. This can be achieved through the so-called Common Radio Resource Management (CRRM) functions, which consider the pool of resources in all RATs as a whole leading to a better overall performance than the aggregated performances of the stand-alone networks. CRRM has been one of the main pillars around which the work in the AROMA project [1] has been built. The proposed advanced solutions follow a range of deployment scenarios, technological functionalities, network topologies and service mixes targeting medium and long term evolutions. The operator driven approach selected for the project ensures the compliance with future developments as well as the provision of useful results at a suitable time, so that cost reduction for infrastructure equipment and applications of 3GPP based mobile networks can be envisaged. In this framework, the purpose of this White Paper is to provide the main results and conclusions devised for CRRM in general and RAT selection in particular as a result of the work in AROMA. Specifically, after identifying the main variables influencing the RAT selection procedure in heterogeneous networks, this White Paper presents the different methodological approaches that have been followed in the project. These include analytical studies, system-level simulations, real-time emulation, techno-economical analysis and implementation considerations. Then, the different identified aspects in terms of radio quality and coverage, perceived user throughput, operator and user preferences are analysed with the objective of devising a generic RAT selection framework. More specifically, focusing on radio coverage, the impact of the main parameters governing the inter-RAT cell reselection in 3GPP is analysed. Then, the impact in terms of user throughput of different CRRM strategies is analysed, in order to estimate the throughput gains with respect to a manual RAT selection. On the other hand, the user and operator preferences are captured by means of a cost function that enables the implementation of different RAT selection strategies. Finally, and trying to collect the main conclusions of the previous studies, a new generic framework is proposed based on a metric denoted as fitness factor, which takes into account all the diverse aspects influencing the RAT selection and comes up with suitable principles under any possible circumstance that may arise in a practical implementation. The proposed framework is evaluated by means of system level simulations and real-time emulations and both techno-economical analysis and implementation considerations are presented. It is obtained that significant investment savings can be achieved by the operator thanks to the proposed framework.

Authors

Jordi Pérez-Romero, Universitat Politècnica de Catalunya, jorperez@tsc.upc.edu
Ricardo Azevedo, Portugal Telecom Inovação, ricardo-a-pereira@ptinovacao.pt
Andrea Barbaresi, Telecom Italia, andrea.barbaresi@telecomitalia.it
Fernando Casadevall, Universitat Politècnica de Catalunya, ferranc@tsc.upc.edu
Luis M. Correia Inst. Sup. Técnico / Technical University of Lisbon, luis.correia@lx.it.pt
Robert Farotto, Telecom Italia, robert.farotto@telecomitalia.it
Ramon Ferrús, Universitat Politècnica de Catalunya, ferrus@tsc.upc.edu
Rickard Ljung, Telia Sonera (currently in Ericsson), rickard.ljung@ericsson.com
Miguel López-Benítez, Universitat Politècnica de Catalunya, miguel.lopez@tsc.upc.edu
Mats Magnusson, Telia Sonera, Mats.E.Magnusson@teliasonera.com
Oriol Sallent, Universitat Politècnica de Catalunya, sallent@tsc.upc.edu
António Serrador, Inst. Sup. Técnico / Technical University of Lisbon, aserrador@deetc.isel.ipl.pt
Anna Umbert, Universitat Politècnica de Catalunya, annau@tsc.upc.edu
Avelina Vega, Telefónica I+D, avn@tid.es

Table of Contents

ABSTRACT	1
AUTHORS	2
TABLE OF CONTENTS	3
LIST OF ACRONYMS	4
1 INTRODUCTION	6
2 AROMA METHODOLOGICAL APPROACHES	8
2.1 ANALYTICAL APPROACHES	8
2.2 SYSTEM-LEVEL SIMULATIONS	8
2.3 REAL TIME EMULATION	9
2.4 TECHNO-ECONOMIC EVALUATION	10
2.5 IMPLEMENTATION CONSIDERATIONS	11
3 CRRM BASED ON RADIO QUALITY AND RADIO COVERAGE	14
4 CRRM BASED ON PERCEIVED THROUGHPUT	16
5 CRRM BASED ON COST FUNCTIONS	17
6 A GENERIC CRRM METRIC: THE FITTINGNESS FACTOR	19
6.1 FITTINGNESS FACTOR DEFINITION	19
6.2 FITTINGNESS FACTOR-BASED RAT SELECTION PROCEDURES	20
6.3 EVALUATION OF THE FITTINGNESS FACTOR-BASED CRRM ALGORITHM THROUGH SYSTEM-LEVEL SIMULATIONS	21
6.4 EVALUATION OF THE FITTINGNESS FACTOR-BASED CRRM ALGORITHM THROUGH REAL-TIME EMULATION	22
6.5 TECHNO-ECONOMIC EVALUATION OF THE FITTINGNESS FACTOR-BASED CRRM ALGORITHM	25
6.6 IMPLEMENTATION OF THE FITTINGNESS FACTOR-BASED CRRM ALGORITHM	27
7 CONCLUSIONS	28
8 ACKNOWLEDGEMENTS	30
9 REFERENCES	30

List of Acronyms

2G	2nd Generation
2.5G	2nd Generation and a half
3G	3rd Generation
3GPP	Third Generation Partnership Project
ABC	Always Best Connected
AROMA	Advanced Resource Management Solutions for Future All IP Heterogeneous Mobile Radio Environments
B3G	Beyond Third Generation
BCCH	Broadcast Control Channel
BO	Blocking Only
BS	Base Station
BSC	Base Station Controller
CAPEX	Capital Expenditure
CDMA	Code Division Multiple Access
CF	Cost Function
CN	Core Network
CPICH	Common Pilot Channel
CR	Cell Redirection
CRRM	Common Radio Resource Management
CS	Cell Selection
DCH	Dedicated CHannel
DiffServ	Differentiated Services
DL	DownLink
DR	Directed Retry
E2E	End to End
Eb/No	bit energy over noise power spectral density
Ec/No	chip energy over total received power spectral density
EDGE	Enhanced Data rates for GSM Evolution
EGPRS	Enhanced GPRS
FOMA	Freedom of Mobile Multimedia Access
GERAN	GSM/EDGE Radio Access Network
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
HHO	Horizontal HandOver
HO	HandOver
HSDPA	High Speed Downlink Packet Access
HSPA	High Speed Packet Access
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IP	Internet Protocol
KPI	Key Performance Indicator
LB	Load Balancing
MMS	Multimedia Messaging Service
MOS	Mean Opinion Score
MPLS	Multi Protocol Label Switching
NPV	Net Present Value
NRT	Non Real Time
O&M	Operation And Maintenance
OPEX	Operational Expenditure
PDC	Pacific Digital Cellular
PDU	Protocol Data Unit
QoS	Quality of Service
R5	Release 5
R99	Release 99
RAN	Radio Access Network
RAT	Radio Access Technology
RIM	RAN Information Management

RLA	Received Level Average
RNC	Radio Network Controller
RRC	Radio Resource Control
RRM	Radio Resource Management
RS	RAT Selection
RSCP	Received Signal Code Power
RSSI	Received Signal Strength Indicator
RT	Real Time
TCH	Traffic Channel
TDMA	Time Division Multiple Access
TD-SCDMA	Time Division-Synchronous Code Division Multiple Access
UE	User Equipment
UL	UpLink
UMTS	Universal Mobile Telecommunication System
UTRAN	UMTS Terrestrial Radio Access Network
UUT	User Under Test
VHO	Vertical HandOver
WCDMA	Wideband Code Division Multiple Access
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area Network
WWW	World Wide Web

1 INTRODUCTION

Today's wireless communications can be driven by a wide range of Radio Access Technology (RAT) standards. The success of second-generation (2G) cellular systems, e.g. GSM (Global System for Mobile Communications), cdmaOne and PDC (Pacific Digital Cellular), along with the IP data support provided by 2.5G technologies, such as the GPRS (General Packet Radio System), paved the way towards evolved systems with higher data rate capabilities. In this sense, technologies like the Enhanced Data rates for GSM Evolution (EDGE) offer high data-rates using inherited 2G network infrastructure and frequency spectrum. In order to supply even higher data rates, third generation (3G) systems arose with new assigned frequency bands along with the deployment of new network elements, especially in the radio access part. 3G systems comprise several standards such as the UMTS (Universal Mobile Telecommunications System), the FOMA (Freedom of Mobile Multimedia Access), CDMA2000, and the Time Division-Synchronous Code Division Multiple Access (TD-SCDMA) among others. Moreover, in parallel with the evolution of cellular systems, a number of Wireless Local Area Networks (WLANs) like, e.g., the IEEE 802.11 standard families, have emerged and become profusely used in home and office environments, as well as in public hotspots. In addition, Wireless Metropolitan Area Networks (WMANs) like the Worldwide Interoperability for Microwave Access (WiMax/IEEE 802.16) standard will extend communication ranges beyond those covered by WLANs.

In this framework, the heterogeneous network notion arises in order to propose a flexible architecture capable of managing this large variety of wireless access technologies along with applications and services comprising different Quality of Service (QoS) demands and protocol stacks. The deployment of such heterogeneous network topologies requires, however, a degree of interworking between the different network entities which may lead to open, loose, tight and very tight couplings [2]. In this way, heterogeneous networks may provide a larger set of available resources than individual networks, allowing users to seamlessly connect, at any time and any place, to the access technology that is most suitable according to some user/operator specified criteria. This notion has been coined as the *Always Best Connected* concept [3].

This coexistence of RATs in heterogeneous wireless scenarios introduces an additional dimension to achieve an efficient exploitation of the scarce available radio resources. RATs may differ from each other by air interface technology, services, price, access, coverage and ownership. The complementary characteristics offered by the different radio access technologies make possible to exploit the trunking gain leading to a higher overall performance than the aggregated performances of the stand-alone networks. Clearly, this potential gain of B3G (Beyond 3G) systems can only turn into reality by means of a proper management of the available radio resources. Common Radio Resource Management (CRRM) refers to the set of functions that are devoted to ensure an efficient and coordinated use of the available radio resources in heterogeneous networks scenarios [4]-[6]. More specifically, CRRM strategies should ensure that the operator's goals in coverage and Quality of Service (QoS) are met while providing as high as possible overall capacity.

In this sense, CRRM has been identified as an important issue by the 3GPP, which defines some recommendations and architectures for CRRM operation [5][6]. Efficient CRRM will then exploit the trunking gain that results from the common management of all the available radio resources of all networks rather than managing those radio resources considering stand-alone networks [7]. Then, the tighter the coupling between these networks, the better the resources are being utilized leading to an improved performance. Consequently, efforts in the definition and implementation of required interfaces must be developed in this direction.

From a functional point of view, CRRM extends the main RRM functionalities arising in the context of a single Radio Access Network¹ (RAN), such as admission and congestion control, horizontal (intra-system) handover, packet scheduling and power control. Specifically, when these functionalities are coordinated between different RANs in a heterogeneous scenario, they can be denoted as "common" (i.e. thus having common admission control, common congestion control, etc.) as long as algorithms take into account information about several RANs to make decisions. In turn, when a heterogeneous scenario is considered, a specific functionality arises, namely RAT selection, devoted to decide to

¹ In this White Paper, the term Radio Access Network is used to refer to a mobile access network built over a single Radio Access Technology.

which RAT a given service request should be allocated, thus ending up with a user-to-RAT association. After the initial RAT selection decision, taken at session initiation, vertical (inter-system) handover is the procedure that allows switching from one RAN to another. The successful execution of a seamless and fast vertical handover is essential for hiding to the user the underlying enabling infrastructure. In general terms, vertical handover procedures from one RAN to another may be useful to support a variety of objectives, such as avoiding disconnections due to lack of coverage in the current RAT, blocking due to overload in the current RAN, possible improvement of QoS by changing the RAT, support of user's and operator's preferences in terms of RATs usage or load balancing among RATs.

RAT selection becomes a key CRRM issue to exploit the flexibility resulting from the joint consideration of the heterogeneous characteristics offered by the available radio access networks. This RAT selection can be carried out considering different criteria (such as, e.g., service type, load conditions, etc.) with the final purpose of enhancing overall capacity, resource utilisation and service quality. The scenario heterogeneity is also present from the customer side, because users may access the requested services with a variety of terminal capabilities (e.g. single or multi-mode) and different market segments can be identified (e.g. business or consumer users) with their corresponding QoS levels. Then, selecting the proper RAT and cell is a complex problem due to the number of variables involved in the decision-making process, as reflected in Figure 1 with some possible inputs. Furthermore, some of these variables may vary dynamically, which makes the process even more difficult to handle.

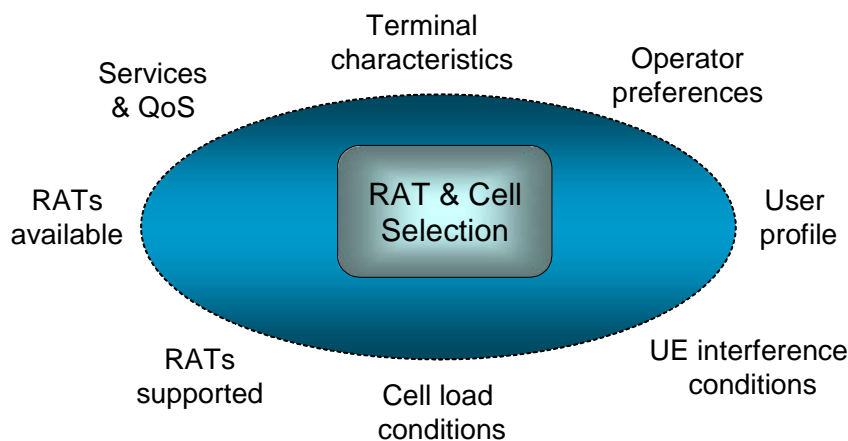


Figure 1: Factors influencing the RAT and Cell selection.

In this context, CRRM in general and RAT selection mechanisms in particular have received a lot of attention in recent years, clearly acknowledging the key role that these strategies will have for a full realization of Beyond 3G (B3G) scenarios. Research efforts have been oriented either to propose and assess the performance of heuristic algorithms [8]-[15] or to identify architectural and functional aspects for CRRM support [5][6][16][17]. From an algorithmic point of view, in [10] and [11], mechanisms to balance the load in different RATs by means of vertical handover decisions are analyzed. However, the service-dimension is not captured in the problem because only real time services are considered. In turn, in [12] the authors compare the load balancing principles with respect to service-based CRRM policies, as described in [8]. Similarly, Lincke discusses the CRRM problem from a more general perspective in e.g. [13][14] and references therein, comparing several substitution policies and including the multi-mode terminal dimension with speech and data services. Finally, in [15] and [18] the authors propose a RAT allocation methodology that considers the specific radio network features of a CDMA network to reduce the interference by allocating users to RATs depending on the total measured path loss and capturing also the service dimension but considering that all terminals support the available RATs and that the involved RATs support the same services.

The abovementioned studies usually approach the problem of RAT selection from a system-level simulation point of view. The analytical approach to the RAT selection problem, however, has received less attention in the literature. Fewer analytical proposals have been developed up to date, among those e.g. [19]-[22]. In [19], Lincke et al. propose an analytical approach to the problem of traffic overflowing between several RATs using an M-dimensional Markov model. Nonetheless, in order to

derive a closed form solution by means of applying independence between service types, Markov states in this model indicate the number of sessions of each service that are being carried in whole composite network, but not on which RAT each session is being carried out. In [21], Koo et al. evaluate the separate and common Erlang capacity of a multi-access/multi-service system. A Markovian approach is also assumed and a closed product form expression is provided [22]. However, this implies that the fractional traffic loads of each service over each system are known. In [20], a near-optimum service allocation is proposed in order to maximize the combined multi-service capacity. The authors assumed a-priori knowledge of the services that need to be allocated, rather than modelling user arrival process.

Under this framework, this White Paper presents the main results and conclusions devised for CRRM in general and RAT selection in particular as a result of the work in the AROMA project [1]. In particular, Section 2 presents the methodological approaches followed in the project. Then, the different identified aspects influencing on RAT selection in terms of radio quality and coverage, perceived user throughput, operator and user preferences are analysed in Sections 3, 4 and 5, respectively, with the objective of devising a generic framework, based on a new metric denoted as fittingness factor, which will be presented and evaluated in Section 6. Finally, conclusions are summarised in Section 7.

2 AROMA METHODOLOGICAL APPROACHES

This section describes different methodological approaches to develop and evaluate RAT selection strategies, based on the work carried out in the AROMA project.

2.1 Analytical approaches

Analytical models can be regarded as a first step towards the objective of gaining insight in the RAT selection problem. In spite of the necessary simplifying assumptions, they can help in identifying the relevant parameters to be considered by the RAT selection algorithm. In the framework of the AROMA project, one of the first works in this line considered a generic CDMA/TDMA heterogeneous network with a single service and developed the criteria to decide the optimum traffic splitting that minimizes the uplink outage probability [23][24]. Two different functions were identified depending on the propagation loss statistical distribution, on the capacities of each RAT and on the corresponding sensitivity levels. From the analysis of these functions and their mathematical properties it was possible to decide the optimum number of users in TDMA and in CDMA.

One of the main outcomes of the previous work is the identification of the relevant role played by the propagation loss distribution in the two access technologies, mainly because of the limited-interference nature of CDMA. As a result of that, in [24] an enhanced methodology has been proposed making use of the specific measured propagation loss, so that those terminals having the lowest propagation loss in a given moment were allocated to CDMA. This approach could be combined either with a simple load balancing approach in which the total load in the CDMA and TDMA RATs is kept at similar levels [10], or with an optimization mechanism like the one presented in [23]. In both cases significant outage probability reductions and consequent capacity increases are achieved.

In order to introduce the service component in the RAT selection procedure, a flexible framework for evaluating generic RAT selection policies in CDMA/TDMA scenarios with two different services has been built by using a 4-dimensional Markov model, [25]. Given a total offered traffic to the network, the fractional traffic arriving to each RAT was dependant on the chosen RAT selection scheme which is fully embedded in the state transitions of the Markov chain. In this way, the model allows the evaluation of different RAT selection schemes accounting for different principles, ranging from the simplest ones like service-based selection or load balancing, up to more sophisticated schemes accounting for the amount of multi-mode terminals in the scenario or trying to minimize the resulting congestion probability in each RAT.

2.2 System-level simulations

Although the analytical models are very useful to get a first insight into the problem and to identify the relevant aspects in the RAT selection process, they usually require some simplifying assumptions in terms of traffic generation patterns, mobility, multi-cell structures and detailed procedures like

measurement averaging, power control, scheduling, etc., so that it is difficult to assess the performance of RAT selection schemes in more realistic scenarios. Then, it is a usual approach to execute detailed dynamic system-level simulations, supported by link-level simulation results, to evaluate the performance of the considered algorithms under realistic conditions in multi-user, multi-cell and multi-service scenarios.

The simulator must include an adequate modeling of the relevant aspects that have an impact over the performance of the strategies being evaluated. In the context of CRRM, these aspects include the traffic generation, the user mobility as well as the different network procedures in the radio interface like e.g. the power control in CDMA, the packet scheduling, the link adaptation, the measurement reporting, etc [26].

In the framework of the AROMA project, different RAT selection schemes have been evaluated by means of the available system-level simulators, taking into consideration the scenarios defined in the project [27] and including detailed characterization of UTRAN, considering both release 99 (R99) channels and High Speed Packet Access (HSPA), GERAN and WLAN technologies. The set of selected scenarios includes both theoretical and realistic scenarios, where the service mix, RAT and environment characterization as well as cell deployment are defined.

2.3 Real Time Emulation

The third approach to demonstrate performance of RAT selection strategies in future heterogeneous networks is real time emulation, which appears as a powerful tool to perform real time trials in order to study the effect of the algorithms on specific applications and is becoming essential as a step forward towards the implementation of the algorithms in real B3G systems. For that reason, a real time testbed, currently emulating a B3G network, has been developed in AROMA project [28]. The network implemented in the testbed is composed of several RAN emulators, namely UTRAN (including R99, and HSPA in both uplink and downlink), GERAN and WLAN. These RANs are connected to IP Core Network (CN) based on DiffServ technology and Multiprotocol Label Switching (MPLS).

Coordination and interworking of the considered RATs in terms of CRRM is stated as a key driver to be studied with the testbed. Besides, the progressive introduction of IP technology in the RAN also constitutes a main pillar in the way towards the definition of more efficient and less complex network architectures capable of accommodating the radio access heterogeneity. Therefore, the control and interworking of these IP-based functionalities like QoS-aware mobility with CRRM arises as one of the testbed goals.

The AROMA testbed has been conceived as an evaluation platform for testing real multimedia applications in real time so that performance metrics and experienced QoS can be extracted. In consequence, it can be used to foresee the behaviour of a specific application when different configuration parameters and algorithms are set in both the radio and CN parts. Selected applications in the testbed cover the services envisaged for heterogeneous networks (i.e. conversational, streaming, interactive and background), and at the same time are widely available and up-to-date (i.e. Quick Time Player, Microsoft NetMeeting, Mozilla Firefox).

Just to illustrate the testbed capabilities and the type of results that can be obtained, in Figure 2, a sample of a videostreaming made with QuickTime Player and Darwin Streaming Server is shown. In the example, videos are QCIF size with average bit rate equal to 128 kb/s and the user under test has full coverage of UTRAN and WLAN. The user starts a session in UTRAN, then after certain time, a vertical handover is decided and it is passed to WLAN. The user is connected to both access technologies with 384 kb/s, so the bandwidth is not restrictive. However, the execution of the VHO procedure introduces some delay and packet losses during the session duration, as it can be observed in the right side of the figure, which shows the resulting image during the VHO, as opposite to the left side of the image, which shows the original image. With this trial it is possible to measure the impact of the loss of IP packets during a VHO on the Mean Opinion Score (MOS). Thus, these trials are able to reflect perceived QoS aspects, which can hardly be captured by means of simulations.

2.4 Techno-economic evaluation

In order for a specific CRRM solution to be attractive for the different players in the wireless communications arena, it is important not only to achieve technical benefits, which can be stated from the evaluation methodologies described in the previous two sub-sections, but also the potential economic benefits that this solution can bring need to be studied. Taking this into account, and following the technical analysis by means of both simulations and real-time emulation, the techno-economic impact of the developed CRRM algorithms in the AROMA project has been analysed. In the following, the considered methodology in that respect is described.



Figure 2: Comparison between two sample videos in a videostreaming session.

The techno-economic evaluations have been carried out by assuming a short- or medium- term increase of the data traffic and by analyzing the potential savings offered by the addressed solutions with respect to the total investment (i.e. CAPEX and OPEX) needed to increase the capacity of the network [29]. Hence, these evaluations are based only on the estimation of the total costs faced by a network operator for upgrading the already deployed network in order to support the expected amount of traffic. In principle, an alternative way of calculating the economic value of the solutions taken into account could be based also on the estimation of the extra revenues related to the additional data traffic supported by the network. Within the context of the work, this approach has been considered less appropriate, since it would require as much exact as possible assumptions on the revenues deriving from the services. Unfortunately, market forecasts on revenues could be very subjective and are usually affected by a higher degree of uncertainty with respect to the estimation of the network investments, since these are strictly related to the willingness to pay of the users for new services. Moreover, revenues from the offerings of new services are strictly dependent also on specific marketing strategies carried out as far as the end-user pricing policies is concerned, that can find justifications on many reasons (e.g. promotion of a specific new service by means of flat-rate prices, volume discounts to boost the usage, etc.).

Within the context of a complex scenario like the one made possible by the all IP heterogeneous network, pricing policies are evidently a very complex issues (especially compared to the voice only traditional scenario), since different pricing and charging schemes are possible (e.g. daily, weekly, monthly flat rate, per subscription rate, data traffic rate, etc.). Moreover service revenues are not always proportional to the load generated in the network, because these respond to different (marketing based) mechanisms with respect to the technical ones. Thus, the value of the service can be not related at all the traffic generated. As a consequence, the revenues typically do not grow linearly with the amount of data exchanged, and this make hard to estimate the economic impacts of the addressed solutions on the basis of the capacity increase achieved within the heterogeneous network. For the above mentioned reasons, considerations on potential revenues have been avoided in the study, which mainly focuses on the potential investment savings.

The techno-economic investigations carried out in AROMA are based on the assumptions that a not negligible increase of data traffic will be demanded by users of mobile heterogeneous network in the next years, especially in dense populated areas. Several market forecasts agree on this assumption on the basis of the recent trends observed in European countries where 3G systems are more diffused

nowadays. Even though the increase of data traffic demands is evident, it is however very difficult to say to what extent, and when, the potential savings related to this market trend can be realized. This uncertainty rises from the difficulty to make reliable traffic forecasts on a per year base, and thus it is difficult to say how big the demand for future network capacity will be. Then, in order to limit the sensitivity of the results on the traffic forecasts, the approach based on the comparison of the solution with respect a “reference case” has been followed. In this way, economic impacts have been highlighted according to a “what-if” approach, apart from the absolute values achieved. This means that the obtained results should not be considered relevant in an absolute way but should be considered useful to compare the different scenarios.

Another important aspect taken into account in the techno-economic investigations of RAT selection approaches consists in the analysis of the market penetration of multi-mode terminals. How much relevant this aspect could be within the context of a heterogeneous network scenario is clear: by means of the CRRM mechanisms addressed by the AROMA projects, different services are supposed to be offered by means of different radio access networks and technologies, in a transparent way for the users, with the aim of improving the QoS and optimizing the network. It is evident that this objective can be accomplished only if a not negligible percentage of users own terminals capable of using most of the radio technologies taken into account.

2.5 Implementation considerations

Finally, and after the evaluation through different means of the proposed algorithms, the targeted nature of the AROMA project is reflected on the interest to deal with algorithm implementation aspects. In this respect, the basic methodology to be followed in such implementation feasibility studies is identified, defining a systematic process valid to move from conceptual algorithm proposals towards their practical implementation, which is illustrated with a number of examples on realistic network deployments.

A basic methodology aimed at defining a systematic process valid to move from conceptual algorithms proposals towards their implementation has been reported in [30]. The proposed methodology is intended to be used as a generic framework accounting for the main items to be dealt with when addressing implementation feasibility studies in AROMA. The proposed methodology comprises the following items:

1. Algorithm selection
2. Functional description of the algorithm
3. Development of practical implementation approaches and identification of critical aspects.
4. Performance assessment for practical implementations.
5. Potential algorithm enhancements and/or new system capabilities requirements

Attending to this methodology, practical implementation aspects of specific algorithms, or sets of algorithms having similar implementation requirements, have been analysed in accordance to the current status of standardisation (i.e. 3GPP, IEEE 802 and IETF specifications). In particular, the development of practical implementation approaches (item 3 within the proposed methodology) has been targeted to identify the main building blocks in terms of:

- Availability of the information and measurements for the decision-making process of a given algorithm in the network element where the logic is allocated.
- Availability and triggering of appropriate execution mechanisms required to enforce decisions in all involved network elements.

Focusing in particular on the CRRM implementation, the following four aspects have been considered as the key CRRM building blocks needed to effectively deploy most of considered CRRM strategies:

- Availability of **inter-RAT traffic steering mechanisms** (e.g. inter-RAT cell reselection and handover) capable to distribute traffic according to the considered CRRM strategy within the different coordinated RATs.
- **Inter-RAT measurement reporting from UEs and base stations.** Some of these measurements shall be used in the decision-making process of the CRRM algorithm.

- **Distribution of inter-RAT information to UEs** so that UEs connected or camping in a given RAT can be aware of the existence and characteristics of potential cells in other RATs. This information can be used to feed autonomous decisions in the terminals (e.g. cell reselection) but also to force the reporting of measurements for a given cell that could be relevant for the CRRM decision.
- **Inter-RAT information and measurement exchanges between radio network controllers** (i.e. RNC, BSC) or equivalent functionality in the different radio access networks so that proper coordination can be achieved in the network side.

Figure 3 illustrates these four key CRRM building blocks. Notice that this classification fits with the execution and monitoring building blocks mechanism that have been identified as the foundation for addressing the development of practical implementation approaches and identification of critical aspects.

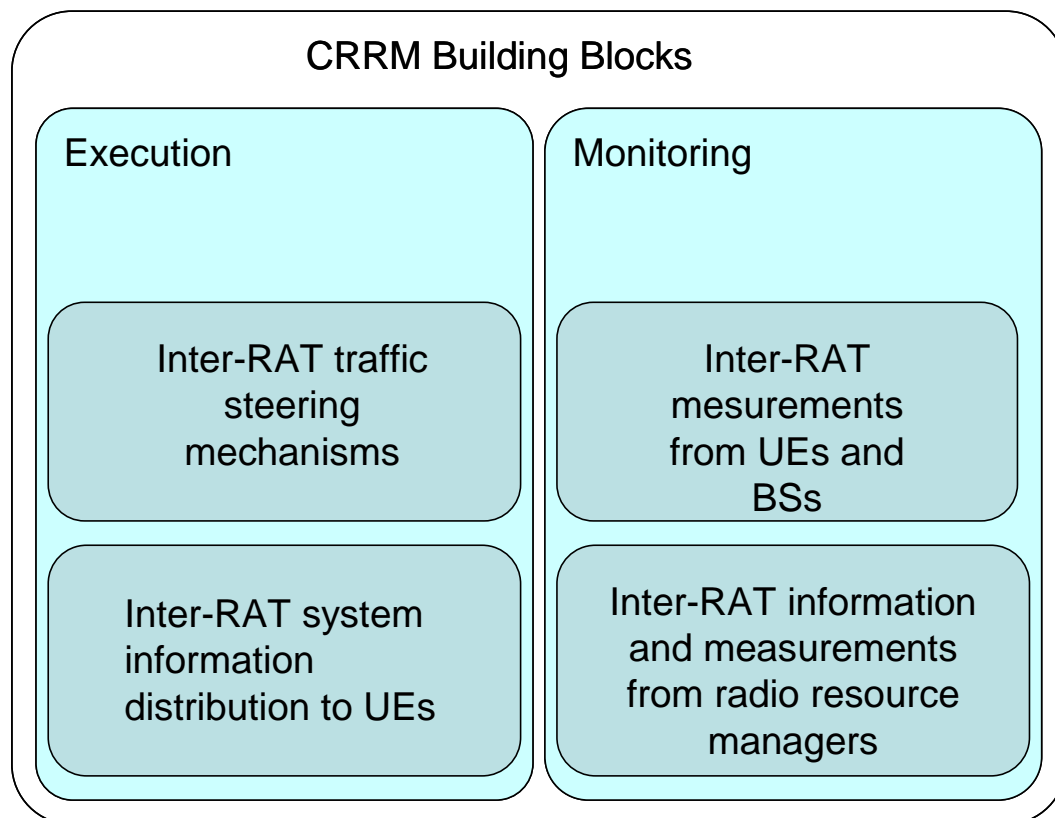


Figure 3: CRRM Implementation Building Blocks.

3GPP specifications already incorporate an important support for the deployment of CRRM strategies between GERAN and UTRAN in terms of inter-RAT system information distribution to UEs, inter-RAT measurements and reporting from UEs, inter-RAT procedures relevant for traffic steering and inter-RAT information and measurements from radio resource managers [31]-[33]. Table 1 summarises the main aspects of the support of the different CRRM building blocks in UTRAN-GERAN networks.

Attending to the current support for CRRM deployment in UTRAN-GERAN networks, some of the proposed algorithms in AROMA can already be implemented without requiring modifications in the specifications, simply based on both directed retry and inter-system handover procedures for distributing connected terminals between RATs. On the other hand, the development of other algorithms would require the modification of the specifications in order to incorporate new metrics.

Finally, it has been considered appropriate to address some considerations about the two major approaches devised in [5][6] to support CRRM in 3GPP networks, namely integrated CRRM and CRRM server. In this regard, it is worth noting that the current approach taken in 3GPP can be understood as a CRRM integrated implementation [6]. Under such approach, CRRM functionality is allocated in existing network elements jointly with local RRM functionality, that is, RNC in UTRAN and BSC in GERAN, without requiring an additional and separated entity to coordinate the overall process.

On the other hand, the CRRM server approach considers the existence of a stand-alone CRRM server that collects/distributes measurements from/to radio network controllers and provides decision support functionalities to the local RRM functions in the radio network controllers that coordinates. Table 2 provides a list of pros and cons of integrated and server CRRM approaches.

Even though according to [5][6] both the two mentioned approaches are considered as theoretically possible, no real implementations of new equipments acting as CRRM Server have been realized by vendors, due to the cons summarized in Table 2. For this reason, all the CRRM solutions identified by the AROMA project have been assessed mainly taking as reference the integrated implementation, which has been considered as the most suitable solution to put in practice in the future.

Table 1: Support for the different CRRM building blocks in UTRAN/GERAN.

CRRM Building Blocks	Mechanisms
Inter-RAT measurements from UEs	<p>UTRAN/GERAN multi-mode terminals can measure and report information of cells belonging to the system they are not connected.</p> <ul style="list-style-type: none"> • When the UE is connected to UTRAN [31]: <ul style="list-style-type: none"> ◦ GSM carrier RSSI (with or without BSIC verification) • When the UE is connected to GERAN [32]: <ul style="list-style-type: none"> ◦ CPICH Ec/No or CPICH RSCP. ◦ UTRAN carrier RSSI (This parameter is reported when a frequency without scrambling code is included in the neighbour cell list)
Inter-RAT system information distribution to UEs.	<p>Both UTRAN and GERAN support the dissemination of inter-RAT cell information that is mainly used to control the cell selection and reselection process for idle mode terminals and to define the set of potential cells involved in the measurements reporting process while in connected mode.</p> <ul style="list-style-type: none"> • In UTRAN, the information element "Inter-RAT cell info list" is used to convey information related to GSM cells. • In GERAN, the information field used to convey UTRAN cells information is named "3G Neighbour Cell Description".
Inter-RAT procedures relevant for traffic steering	<p>3GPP specifications foresee several radio procedures that can be exploited by a CRRM algorithm to implement inter-RAT traffic steering:</p> <ul style="list-style-type: none"> • Inter-system Directed Retry (DR) and Inter-system handover (HO) for connected terminals • Cell Reselection for idle mode terminals • Cell Redirection (CR) for terminals switching to connected mode and inter-RAT redirection upon RRC connection release
Inter-RAT information and measurements from radio resource managers.	<p>There are basically two main mechanisms to exchange specific information and measurements among a RNC and a BSC:</p> <ul style="list-style-type: none"> • "Measurements on Common Resources" and "Information Exchange" functions included within the Iur-g interface. • Transparent Containers within inter-systems procedures. <p>The information and measurements exchanged are limited so far to:</p> <ul style="list-style-type: none"> ◦ Cell Capacity Class Value ◦ Load Value ◦ RT Load Value ◦ NRT Load Information Value <p>Additionally, a generic mechanism referred to as RAN Information Management (RIM) has been specified for the exchange of arbitrary information between applications belonging to the RAN nodes.</p>

Table 2: Pros and cons of integrated and server CRRM approaches.

	Integrated CRRM	CRRM Server
Pros	<p>It is not necessary to fully specify the functional interface between RRM and CRRM as a new open interface in this solution.</p> <p>Only "Reporting Information" is exchanged over existing open interfaces.</p> <p>No need to specify additional interfaces for "RRM decision support" functions</p>	<p>Coordination of competing decision points for CRRM algorithms with multi-RAT scope (e.g. load based handover), thus avoiding possible ping pong effects resulting from contradictory decisions of different nodes in a distributed approach.</p> <p>In a multi-vendor environment, the definition of a new open interface to support RRM/CRRM coordination will force to open to some extent RRM "internals" in radio network controllers so that the specification of CRRM algorithms can be facilitated.</p>
Cons	<p>Resolution of conflicts in distributed decisions is more difficult, even though possible. Notice as an example that because of the distributed nature of the VHO decisions and the different measure availability in local and remote nodes it could happen that one entity decides to trigger a VHO towards another RAT and the corresponding entity in the other RAT rejects it, leading to some ping-pong effects.</p> <p>Need to guarantee consistency in the specification of CRRM algorithms in a multi-vendor environment (i.e. RNC and BSC are from a different vendor)</p>	<p>Requires specification of new open interfaces between RRM and CRRM.</p> <p>Increase of the level of signalling in the mobile network in case of measurements per terminal having to be transferred to the CRRM server.</p> <p>Need to guarantee consistency in case that the CRRM server belongs to a different vendor than RNC and BSC.</p> <p>In case of failure of CRRM server, all the CRRM functionalities cannot operate, and a single CRRM server is supposed to serve a very large area of the network.</p> <p>Costs (CAPEX and OPEX) for access network deployment increase due to additional network equipment.</p>

3 CRRM BASED ON RADIO QUALITY AND RADIO COVERAGE

A first set of strategies explored the inter-working mechanisms between GERAN and UTRAN specified by 3GPP, in order to identify useful CRRM strategies exclusively based on radio quality perceived by the users. More in detail, mechanisms related to segregating traffic between these two RATs by means of the parameters affecting inter-RAT cell reselection in idle mode and handover in connected mode were analyzed.

More specifically, focusing on the idle mode, the main parameters governing the inter-RAT cell reselection from GERAN to UTRAN are the Q_{search_I} and the $FDD_Qoffset$ [34]. Particularly, Q_{search_I} controls the instant when a dual mode mobile terminal camped on a GERAN cell starts to measure UTRAN cells. Then, for each cell having a CPICH E_c/N_0 above the minimum quality threshold the measured value of CPICH RSCP is compared with the received level average value (RLA_C) of the GERAN serving cell and neighbour cells. If the measured CPICH RSCP is greater than all the RLA_C values plus the $FDD_Qoffset$, then the UTRAN cell is considered as a suitable target cell for the cell-reselection procedure. By repeating the steps mentioned above, a list of target UTRAN cells is derived by the mobile terminal; among them, the cell with the greatest CPICH RSCP is selected.

From the achieved results in the GERAN to UTRAN case, as shown in Figure 4 it is possible to derive that when FDD_Qoffset decreases, UTRAN RAT is favoured with respect to GERAN and when the lowest value for the FDD_Qoffset is assumed (-28 dB), the maximum UTRAN usage level is achieved. Similarly, when Qsearch_I increases, GERAN is more and more favoured with respect to the UTRAN. For details on the results the reader is referred to [30].

Concerning the cell reselection from UTRAN to GERAN, it can be controlled by regulating the SsearchRAT_GSM and Qoffset1_sn parameters [34]. Specifically, the mobile terminal starts to perform inter-RAT measurements whenever the quality the serving UTRAN cell is below the SsearchRAT_GSM threshold, and Qoffset1_sn is an offset applied to GSM cell measurements and derived from System Information. For this situation, simulation results shown in Figure 5 reveal that, when SsearchRAT_GSM increases, GERAN is more and more favoured with respect to the UTRAN, while, on the contrary, increases in the value of Qoffset1_sn lead to more users connected to UTRAN.

In turn, focusing on the terminals in connected mode, simulation results dealing with UTRAN to GERAN handover highlight that the handover procedure can be effectively exploited in order to take advantage of GERAN as a back-up system when the radio quality of UTRAN cells is not able to support user's service, which occurs, for instance, in case of indoor users [30].

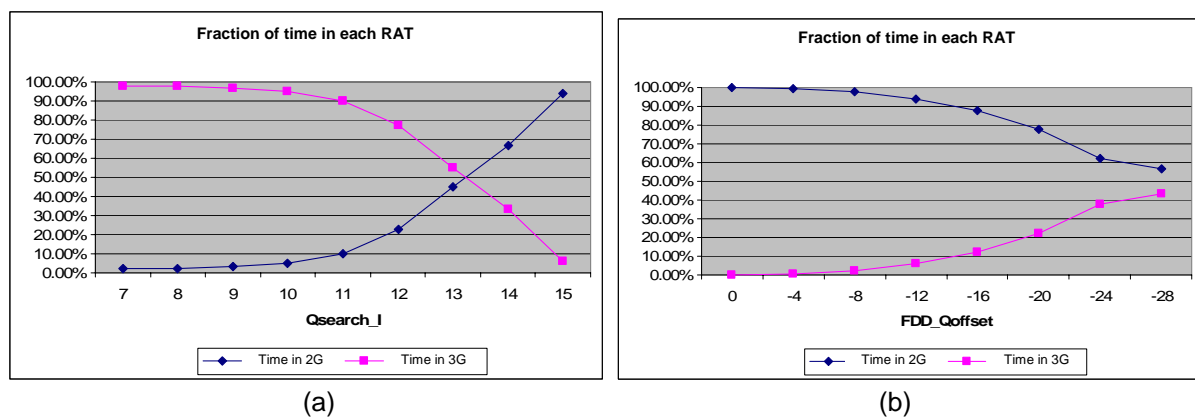


Figure 4: Time spent by users in GERAN and UTRAN versus (a) Qsearch_I and (b) FDD_Qoffset.

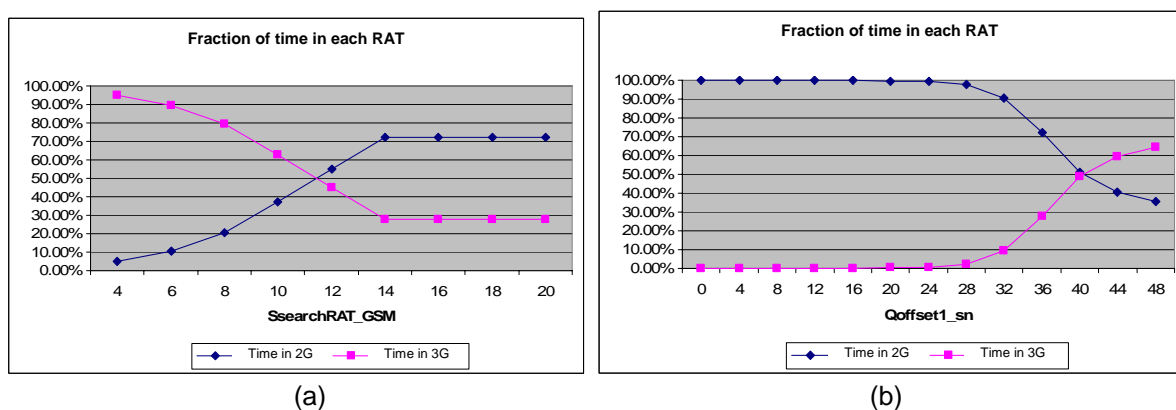


Figure 5: Time spent on GSM and UTRAN cells versus (a) SsearchRAT_GSM and (b) Qoffset1_sn

4 CRRM BASED ON PERCEIVED THROUGHPUT

Another possible guiding principle to develop CRRM strategies consists in focusing on the quality perceived by the user in terms of application throughput. In the framework of AROMA, this has been analysed in a heterogeneous scenario with 2G, 3G and WLAN technologies based on the AROMA target scenario "hotspot in urban area" [27]. The performance in terms of total perceived system throughput, for the streaming, interactive and background services has been obtained, while the conversational services are included only as an additional load factors in the system. A loose coupling with WLAN is assumed, which causes additional signalling delays with respect to the tight coupling approach.

In the considered scenario, different RAT selection policies can be considered depending on a wide variety of both technical and economical aspects. Specifically, the considered aspects are the minimum accepted end user throughput ratio, which captures the user perspective, and the service prioritisation, which enables the operator the possibility to give more importance to some services with respect to others, thus capturing the operator perspective. From these parameters, the following CRRM algorithms are considered in order to decide how to empty the CRRM buffer, which contains all the data to be transmitted in the downlink direction, by allocating the appropriate RAT to each transmission:

- Algorithm 1 - Long term optimisation criteria: Each time the CRRM is activated this algorithm finds the combination of data services allocated to RATs that in the end would lead to the shortest time until the CRRM buffer is empty.
- Algorithm 2 - Short term optimisation criteria: This CRRM algorithm finds the combination of data services allocated to RATs that will lead to the shortest time until one or more radio links is available. Hence, it also minimises the time until a new CRRM allocation will be made.
- Algorithm 3 - Waterfilling RAT prioritisation: The algorithm analyses the expected available data rates for each new transmission from the CRRM buffer. At each new allocation of a transmission request to a radio link the network with highest available data rate is utilised. This is repeated until all radio access networks are fully loaded up to the maximum number of allowed radio links (due to operator policy) or until the CRRM buffer does not contain any more non-allocated transmission requests. In practice, in a scenario with WLAN, R99 WCDMA and GERAN, this results in a waterfilling algorithm where the WLAN (which usually has the highest data rates available) is filled until being fully occupied. Then 3G network is filled, and last 2G network is filled.
- Algorithm 4 - Reference case without CRRM: This algorithm simply makes a fixed service mapping to the different RATs. Particularly, it is assumed that www and streaming use WLAN, email users 3G and MMS uses 2G.

The results of the evaluation of these algorithms with different setups according to Table 3 and presented in Figure 6 reveal that the total perceived system throughput can be improved by advanced CRRM algorithms even if the system architecture is based on loose WLAN coupling. Depending on the operator service prioritisation selections simulations have shown a relative system throughput improvement of 10-50% by using advanced CRRM algorithms compared to a manual RAT selection procedure.

Table 3: Different simulation setups based on the hotspot in urban scenario.

	2G system support	3G system support	WLAN system support	Operator service prioritisation policy prioritises
Case a	GPRS	R99 DCH	802.11 a/b/g	Large payloads
Case b	GPRS	R99 DCH	802.11 a/b/g	Small payloads
Case c	GPRS / EGPRS	R99 DCH / R5 HSDPA	802.11 a/b/g	Large payloads
Case d	GPRS / EGPRS	R99 DCH / R5 HSDPA	802.11 a/b/g	Small payloads

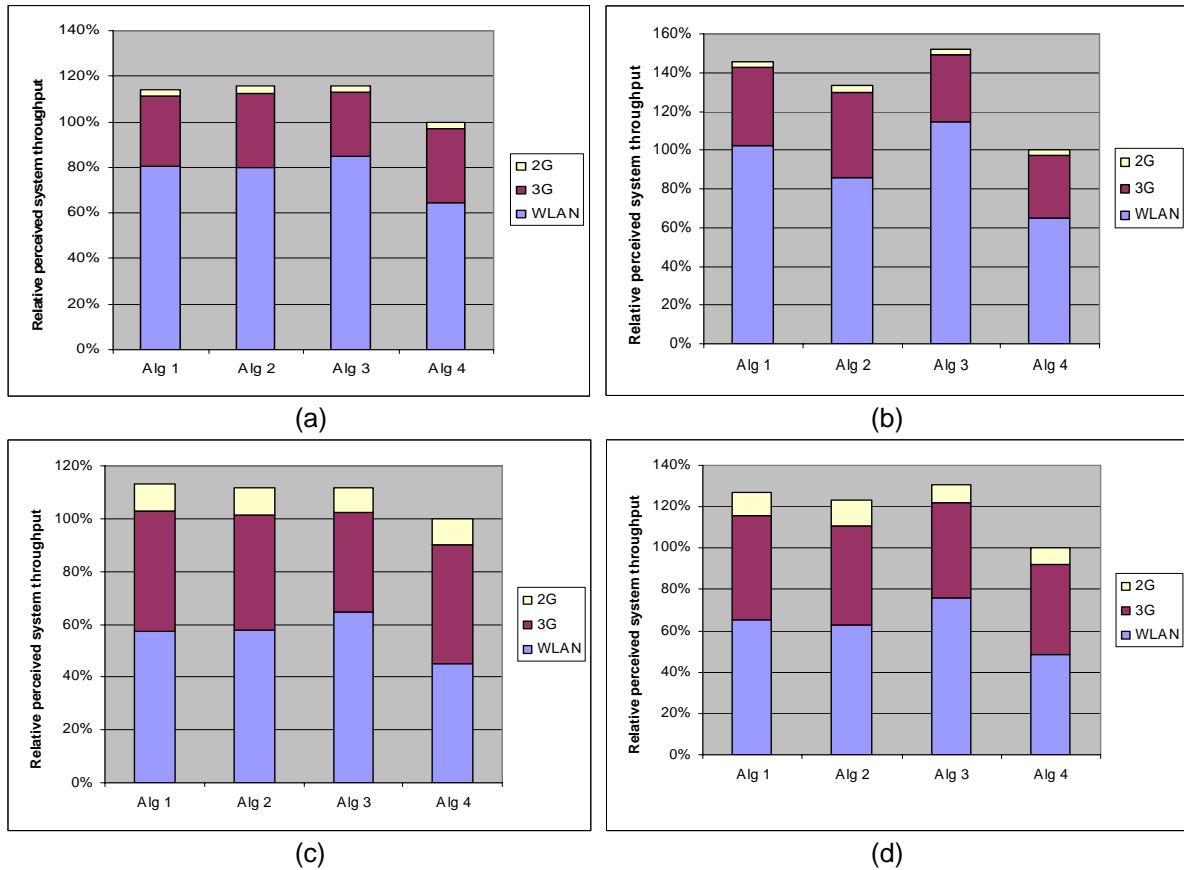


Figure 6: Relative CRRM throughput for the different CRRM algorithms compared to manual RAT selection (algorithm 4). (a) Case a, (b) Case b, (c) Case c, (d) Case d.

5 CRRM BASED ON COST FUNCTIONS

Another model capable of dealing with multiple RATs, which have intrinsically some differences on QoS indicators, is the Cost Function (CF) one. For each RAT, a particular CF definition was identified, by using slightly different KPIs (Key Performance Indicators), i.e., each BS-RAT type has its own CF, supported on different and appropriate KPIs.

Another important issue, related to the computation of the CF model, is the different perspectives that different network players have over the network, which in this model are seen from the operators' and users' viewpoint, as depicted in Table 4.

When each of these groups "looks" to the network, they are sensitive to different parameters: for a user, the operator/network is seen as a service provider/infrastructure, therefore, e.g., service cost (being lower) and quality (being higher) are important; however, for an operator, the same parameter can have an opposite perspective, e.g., service cost should provide good revenue and simultaneously be competitive with other operators. Hence, in order to provide a more realistic balance in the overall network solution, the overall CF should combine both operator's and user's perspectives. Table 4 presents a list of KPIs identified for both perspectives [35]. One should note that not all KPIs have a correspondence to both perspectives, e.g., interference is clearly a very important parameter for an operator, but it does not carry any meaning for a typical user.

Based on the previous concepts, the network total CF is divided into two sub-CFs, one being related to the operator and the other to users. Furthermore, the operator CF is also sub-divided, since different CFs are computed for each different RAT type. Each one of these sub CFs is weighted with different values, enabling the implementation and evaluation of different policies on CRRM and RRM algorithms over each type of RAT. The details on the formulation can be found in [36]-[38].

Table 4: Users and Operators CF Parameters.

KPIs	Perspective	
	User	Operator
Delay	Service	BS Average
Blocking	Service	BS Average
Cost	Service (Free, Flat, Volume or Time dependent)	BS
Throughput	Service	BS Average
Service Availability	Number of RATs available	-
Drop Rate	Service	VHO and HHO
User type	-	Mass Market, Premium
Interference	-	BS Level
Load	-	BS
Channels	-	BS Occupied resources

The CF result applied to all BSs in a heterogeneous network environment offers to RRM and CRRM entities a good way to implement the Always Best Connected (ABC) concept [35], since each BS has a number associated to it, the cost value. Based on these values, the CRRM entity can sort a list of BSs reported/visible by each user, via the RRM entity. On the top of this list, it is expected to have the best BS (the lowest cost one) that potentially offers the best connection to a given user.

Similar to BSs, each user has a cost value attached to him/her. This information is vital to take users' interests into account, in the overall network management.

By using this model, the network heterogeneous environment can be evaluated based on a huge combination of different policies, ranging from only one single KPI, looking just for either the operator's interests or the user's ones, up to the full inclusion of all identified KPIs.

The following figures presents some results based on the CF model and on the AROMA Urban Hotspot scenario [27], by exploring the comparison between operator's and user's KPIs perspectives, previously proposed and identified. Further results can be found in [36] and [37].

Figure 7(a) compares the CRRM delay, when the CF policy is based on independent operator's and user's perspectives. Specifically, the Blocking Only (BO) policy is considered, that aims at decreasing the overall blocking probability at CRRM level. Note that the BO policy is different for both user and operator sides. The users' oriented policy presents worse results, since users are not concerned about the overall network QoS. Another important issue is that users and operators blocking probability do not represent the same quantity, because operators and their BSs QoS counters have all events in memory, but only the ones registered by active users are considered when their perspective is taken into account. Therefore, the CF is more realistic when the operators' perspective is used. Thus, the network QoS is better represented by the operators' perspective. In Figure 7(b), a CRRM blocking comparison is presented, using the same previous policy, illustrating the previous effect.

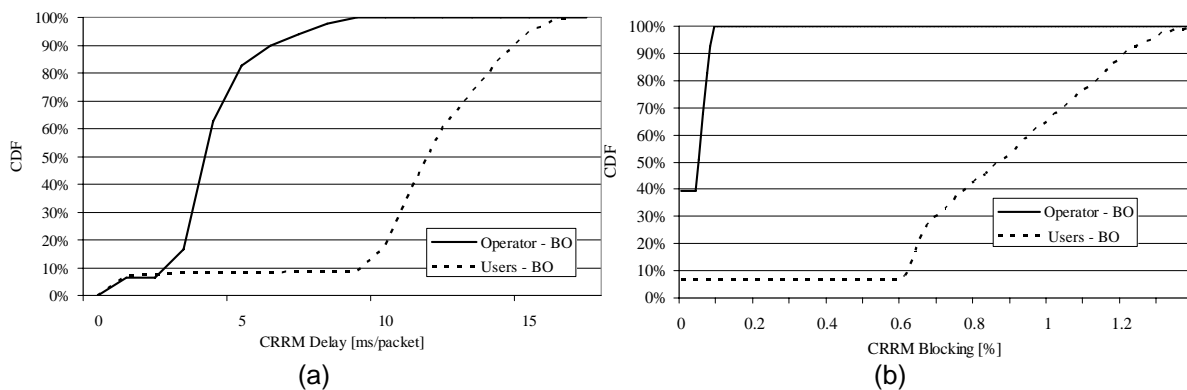


Figure 7: Performance when comparing operators' and users' perspectives using the BO policy in terms of (a) CRRM delay, (b) CRRM blocking probability.

6 A GENERIC CRRM METRIC: THE FITTINGNESS FACTOR

Following the previous studies related with coverage and different operator/user perspectives, and trying to combine the different obtained conclusions, a generic framework trying to capture all the effects influencing on the RAT selection decisions is presented. Specifically, in order to cope with the multi-dimensional heterogeneity reflected in Figure 1, the following main levels are identified in the RAT selection problem:

- 1) Capabilities. A user-to-RAT association may not be possible for limitations in e.g. the user terminal capabilities (single-mode terminal) or the type of services supported by the RAT (e.g. videophone is not supported in 2G).
- 2) Technical suitability at the radio part. A user-to-RAT association may or may not be suitable depending on the matching between the user requirements in terms of QoS and the capabilities offered by the RAT (e.g. a business user may require bit rate capabilities feasible on HSDPA and not on GPRS or these capabilities can be realised in one technology or another depending on the RAT occupancy, etc.). There is a number of considerations, which can be split at two different levels:
 - a) Macroscopic. Radio considerations at cell level such as load level or, equivalently, amount of radio resources available.
 - b) Microscopic. Radio considerations at local level (i.e. user position) such as path loss, intercell interference level. This component will be relevant for the user-to-RAT association when the amount of radio resources required for providing the user with the required QoS significantly depends on the local conditions where the user is (e.g. power level required in WCDMA downlink).
- 3) Technical suitability at the transport part: A user-to-RAT association may or may not be suitable depending on the matching between the user requirements in terms of QoS and the capabilities offered by the transport network, mainly depending on the current load existing in the different links.
- 4) Operator/user preferences: Specific user-to-RAT associations may be preferred without any specific technical criterion but responding to more subjective and economic-related aspects (e.g. due to the investment carried out by an operator to deploy a given technology it can be preferred to serve the traffic through this technology so that investments can be recouped faster, the operator prefers to give some precedence of a service over another one depending on market strategies, etc.).

6.1 Fittingness factor definition

The above concepts can be captured in the so-called *fittingness factor*, which reflects the degree of adequacy of a given RAT to a given service requested by a given user. Although each RAT has its own particularities, it is possible to make a general definition of the fittingness factor by grouping the different terms under some commonalities general to any RAT. Specifically, the fittingness factor is defined with respect to each cell of a given RAT for each user belonging to a certain profile (e.g. business/consumer) and requesting a specific service, as the product of four different terms ranging from 0 to 1, which are:

a) Terminal and RAT capabilities:

The first term reflects the hard constraints posed by the capabilities of either the terminal or the technology, and therefore it simply takes the value 1 if both the terminal and the RAT support the requested service and 0 otherwise.

b) Technical suitability in the radio part:

This factor reflects the user-centric suitability of the RAT to support the service requested by the user. It basically accounts for the bit rate that can be achieved by the corresponding user/service/profile in the RAT normalized with respect to the maximum bit rate that could potentially be achieved by this user/service/profile in any of the existing RATs. It consists mainly of two multiplicative factors, dealing with the microscopic and the macroscopic dimensions. The macroscopic dimension accounts for the load in the corresponding RAT, as well as for the degree of precedence with respect to other services/profiles, which eventually impacts on the bit rate that can be achieved by the user due to user

multiplexing considerations. In turn, the microscopic dimension accounts for the bit rate available per single channel (e.g. one time slot in GERAN, one DCH channel in UTRAN, etc.) depending on the propagation and interference conditions experienced by user as well as on the specific characteristics of the RAT (e.g. in case of UTRAN will account for the load factor and/or power availability, in GERAN and WLAN will depend on the link adaptation mechanisms, etc.). Specific definitions of the technical suitability factor in the radio part for UTRAN/GERAN RATs and different services can be found in [38][39].

c) Network-centric suitability:

This term intends to capture the suitability from an overall RAT perspective, then to provide further flexibility on the fittingness factor definition. For that purpose, this term can include tuneable parameters to allow the enforcement of specific operator policies arising from the trade-off between the degree of QoS to be provided to different types of traffic. Just as an example, let define the non-flexible traffic, which is the traffic that can only be served through one specific RAT and therefore it does not provide flexibility to CRRM (e.g. in UTRAN it could be the videocal users assuming they cannot be served through other RATs such as GERAN, WLAN, etc., while in GERAN, it could be the mono-mode terminals, which cannot be served through other RATs). In such a situation, the operator may decide to give more precedence to the non-flexible traffic depending on the desired policy. Then, the Network-centric suitability term is a function that reduces the fittingness factor of flexible traffic depending on the amount of non-flexible load. The idea is that if there is a high amount of non-flexible load in a given RAT, this RAT is made less attractive for flexible load, thus leaving room to non-flexible users. It is worth mentioning that the definition of the network-centric suitability can include tuneable parameters to allow the enforcement of specific operator policies arising from the trade-off between the degree of QoS to be provided to the non-flexible traffic with respect to the flexible traffic [40]

d) Transport network capabilities:

This term will account for the bit rate available for this user/service/profile in the transport network of the RAT in accordance with the bottleneck link utilisation in the lub interface for a given path between a NodeB and its controller RNC. Consequently, this term reflects the amount of load existing in the transport network, in the sense that a value close to 1 means that the transport network has very low load and therefore it does not introduce limitations in the service bit rate, while a value close to 0 means that the transport network is overloaded and as a result it can limit the achievable service bit rate.

6.2 Fittingness factor-based RAT selection procedures

Based on the above definition of the fittingness factor, two different RAT selection procedures can be identified depending on whether the selection is done at session set-up or during an on-going connection for a user requesting a given service s , as detailed in the following:

Session set-up case

- Step 1.- Measure the fittingness factor for each candidate cell k_j of the j -th detected RAT. The fittingness factor should be computed separately for uplink and downlink of a given RAT. Then, a weighting between the two values can be carried out depending on the service characteristics
- Step 2.- Select the RAT J having the cell with the highest fittingness factor among all the candidate cells. In case that two or more RATs have the same value of the fittingness factor, then select the less loaded RAT.
- Step 3.- Try admission in the RAT J .
- Step 4.- If admission is not possible, try with the next RAT in decreasing order of fittingness factor, provided that its fittingness factor is higher than 0. If no other RATs with fittingness factor higher than 0 exist, block the call.

On-going connection case

For on-going connections, the proposed criterion to execute a VHO algorithm based on the fittingness factor would be as follows, assuming that the terminal is connected to the RAT denoted as "servingRAT" and cell denoted as "servingCell".

- Step 1.- For each candidate cell and RAT, monitor the corresponding fittingness factor. Measures

should be averaged during a period T .

Step 2.- If the fittingness factor of a given RAT in an specific cell a is greater than the fittingness factor of the current serving RAT and cell for a period T_{VHO} then a vertical handover to the new RAT and/or cell should be triggered, provided that there are available resources for the user in this RAT and cell.

6.3 Evaluation of the fittingness factor-based CRRM algorithm through system-level simulations

The developed framework has been analyzed through exhaustive simulations in a variety of scenarios with different traffic mixes to reveal the ability to adapt to the conditions in each case. In the following some results are given to illustrate the benefits of the proposed framework. They consider a scenario with UTRAN R99 and GERAN cells with EDGE capabilities. Voice at 12.2 kb/s and videocall at 64 kb/s services are considered as representative of the conversational traffic class while a www browsing service with two different profiles, namely consumer (with bit rate up to 128 kb/s in UTRAN and low priority in GERAN) and business (with bit rate up to 384 kb/s and high priority in GERAN), have been selected as representative of the interactive traffic class.

In order to illustrate how the fittingness factor algorithm affects the traffic splitting among the two RATs, Figure 8(a) plots the fraction of traffic served through GERAN for the voice, interactive consumer and business profiles when increasing the total load coming from videocall users (which are always served through UTRAN). For comparison purposes, the distribution according to the load balancing case (LB) is also shown, in which the less loaded RAT is selected at session set-up. It can be observed how LB does not make significant distinctions among the considered services, with the general trend that, by increasing the load of videocall users, more traffic of the other services should be derived to GERAN in order to keep similar load levels in the two RATs. On the contrary, the fittingness factor based algorithm is able to split the traffic according to the peculiarities of each service. In particular, most of the interactive business traffic is served through UTRAN, where this type of traffic can achieve a higher bit rate. Only in case that the videocall load is very high there is a certain interactive business traffic that should be moved to GERAN. In turn, when looking at the voice and interactive consumer users, as a result of the increase in videocall load, the algorithm tends to move to GERAN mainly the voice traffic, while it keeps a significant fraction of interactive consumer traffic still in UTRAN.

The different traffic split impacts over the QoS observed by each service, as it is reflected in Figure 8(b), which compares the packet delay of the interactive consumer and business users with the fittingness factor based algorithm and with LB when increasing the load of voice users in the scenario. It is observed that the performance from a user point of view is better with the fittingness factor-based algorithm than with LB for the two user profiles. Although it is not plot here for the sake of brevity, the total throughput achieved in the scenario in this case is approximately the same for the two approaches, which reflects that the fittingness factor is able to improve the user QoS perception without degrading the overall capacity.

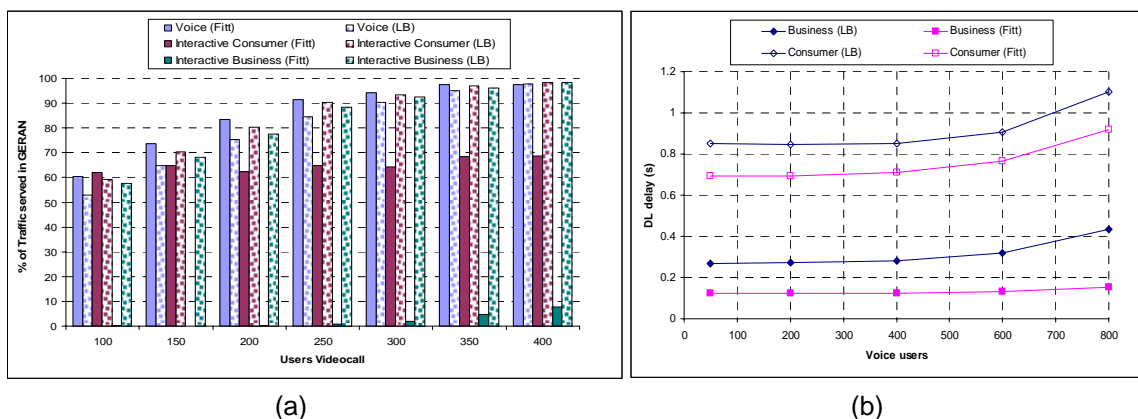


Figure 8: (a) Fraction of traffic served through GERAN as a function of the videocall users for the fittingness factor and the load balancing strategies. (b) DL packet delay for interactive business and consumer users according to the fittingness factor and the load balancing case.

The impact of the videocall users, which are a non-flexible type of traffic, is plot in Figure 9(a) and (b) in terms of packet delay and total throughput in the scenario. To illustrate the ability of the algorithm to reflect different operator criteria, two different settings have been considered in the network-centric suitability component in the fitness factor definition. Setting 1 reflects the case in which the operator aims at improving the QoS of the interactive business users. In this case, as reflected in Figure 9(a), the delay for this traffic is the smallest one among the considered approaches, which is achieved by keeping as much as possible this traffic in UTRAN, even if videocall load is high. However, this improvement occurs at the expense of a reduction in the throughput of non-flexible traffic, because the interactive users leave less room in UTRAN. As a result, there is some reduction in the total throughput in the scenario, as shown in Figure 9(b). In turn, setting 2 reflects the situation in which the operator prefers to keep more capacity for the videocall traffic in UTRAN, which is ensured by allowing that some interactive business users are served through GERAN. Notice in Figure 9(a) and (b) that, with this setting, the total throughput in the scenario is increased with respect to setting 1 at the expense of some delay degradation. Nevertheless, the delay is still much better than that achieved with LB.

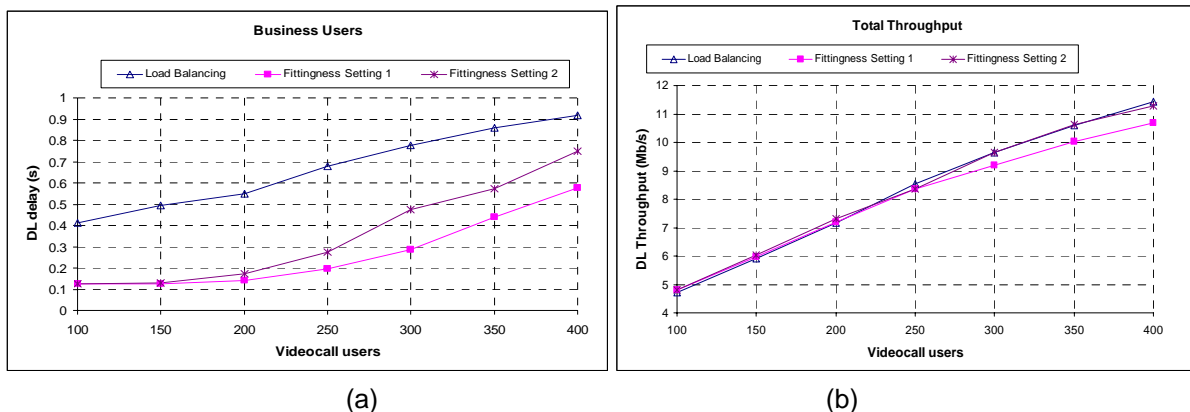


Figure 9: (a) DL packet delay of business interactive users, (b) Total scenario DL throughput when increasing the videocall load.

Finally, in order to illustrate how the fitness factor based algorithm can take into account transport network limitations, results have been obtained in a UTRAN/GERAN co-site scenario in which there exists a bottleneck in the Iub interface of one UTRAN cell, namely BS3. Figure 10 plots the traffic distribution in the UTRAN and GERAN cells BS0 and BS3 for the cases in which no transport network considerations are included in the fitness-factor based algorithm and in which transport network considerations are taken into account in the fitness-factor based algorithm. It can be observed that due to the transport limitations the algorithm is able to reduce the overall traffic in BS3 moving part of this traffic to other UTRAN cells (i.e. BS0) or GERAN cells (i.e. BS0 or BS3). As a result of this traffic splitting, Figure 11 plots the corresponding PDU loss ratio and delay in the Iub interface of BS3 and it can be observed that, when both RAT selection (RS) and cell selection (CS) include transport network considerations in the fitness factor definition performance can be significantly improved. Further results can be found in [38].

6.4 Evaluation of the fitness factor-based CRRM algorithm through real-time emulation

In addition to the performance evaluation of the fitness factor-based CRRM algorithm carried out through system level simulations, also the real time emulator of the AROMA project [28] has been used to test the performance of the algorithm. Specifically, considering a multi-service scenario similar to those in the previous sub-section but without transport network limitations, Figure 12 shows the evolution of some illustrative parameters for an interactive consumer terminal. For the two first columns of graphs, upper graphs correspond to UTRAN while lower graphs are related to GERAN. The left graphs of the figures show the measured path loss between the user under test (UUT) and the two base stations located at the end-points of the trajectory. At time instant 0 seconds the UUT is near base station 2 (white line). Base station 3 (dark blue line) is reached at time instant around 130 seconds. Then, the UUT turns back to base station 2, which is reached again at time instant around 260 seconds. The middle graphs of the figures show the value of the technical suitability at the radio

part for both directions in each RAT (uplink in orange and downlink in blue). The values of these parameters for both directions are averaged with the same weight to obtain the final value of the fittingness factor for each RAT, which is shown in the top right graph (UTRAN in yellow and GERAN in red). The bottom right graph shows the current RAT the UUT is connected to. Values 0 and 1 correspond to UTRAN and GERAN, respectively. When the UUT is not connected to any RAT (during the first 5 seconds) the value -1 is shown. As it can be appreciated, the algorithm always allocates the user to the RAT offering the highest fittingness factor, triggering VHOs whenever they are required.

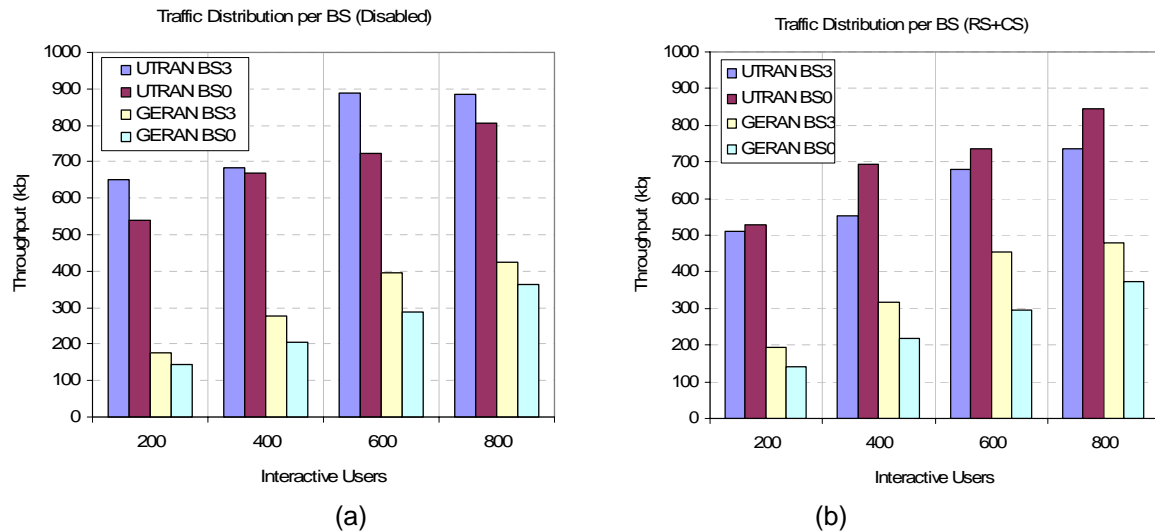


Figure 10: Traffic distribution among UTRAN and GERAN cells for the cases (a) No transport network considerations in RAT selection and (b) Transport network considerations in both RAT and cell selection

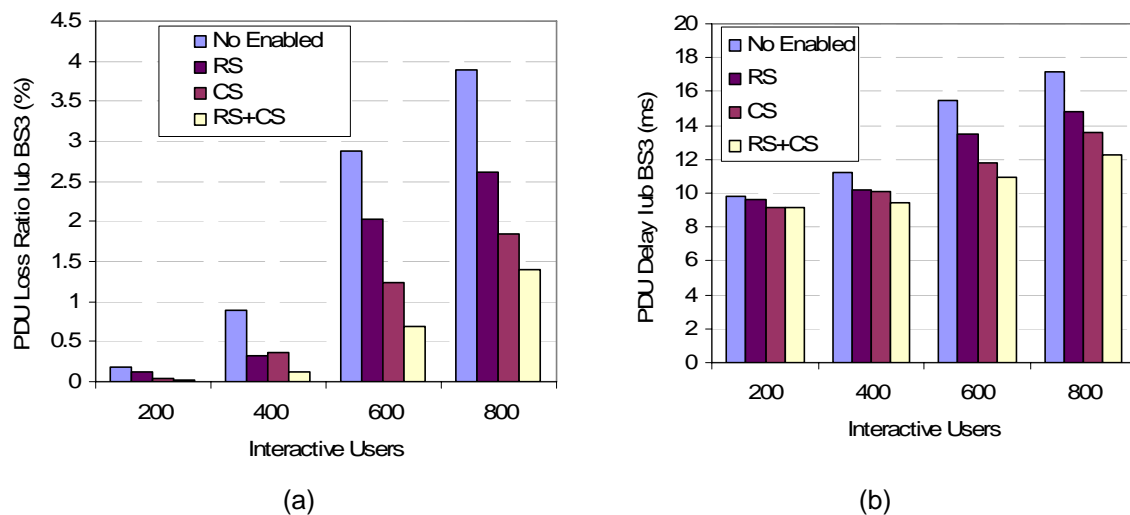


Figure 11: Performance in the BS3 lub interface with different strategies including both RAT selection (RS) and cell selection (CS) with transport network considerations. (a) PDU loss ratio. (b) PDU delay.



Figure 12: Evaluation of the fittingness factor based RAT selection for an interactive service with 500 emulated users

To analyse the behaviour of the fittingness factor algorithm, Figure 13(a) shows the average fittingness factor for interactive consumer users in GERAN and UTRAN, while Figure 13(b) shows the distribution of users among the available RATs. The presented results correspond to two different cases. The first case (depicted with black lines) considers the availability of 3 carrier frequencies in each GERAN cell. Since one time-slot must be reserved for signalling, 23 time-slots are available for Traffic Channels (TCHs). The second case (depicted with pink lines) considers 2 carrier frequencies in each GERAN cell (15 TCHs). For the first case, Figure 13(a) shows that GERAN, in general, offers a higher average fittingness factor than UTRAN. This is due to the fact that the bit-rates obtained in GERAN with 2 slots in uplink (up to 118.4 kbit/s) and 4 slots in downlink (up to 236.8 kbit/s) are considerably higher than those of UTRAN interactive bearers with 64 kbit/s in uplink and 128 kbit/s in downlink. In addition, UTRAN is already loaded with other types of traffic, as shown in Figure 13(b). As a result, the fittingness factor is higher for GERAN. One interesting point in Figure 13(a) is the higher sensitivity of the UTRAN fittingness factor to an increment in the number of users. While the value for UTRAN rapidly decreases as the number of users increases in the first case, the value for GERAN experiences a small variation. This is due to the low number of users admitted in GERAN, see Figure 13(b), which forces UTRAN to absorb the increment of users. If the number of admitted users in a given RAT increases, or alternatively the amount of available resources decreases, the value of the multiplexing factor will decrease, and therefore a reduction of the fittingness factor is expected. In effect, this behaviour is observed for GERAN in Figure 13(a) when the number of carrier frequencies per GERAN cell is reduced from 3 to 2. In this second case, the fittingness factor for GERAN exhibits a higher sensitivity to the number of users than in the first case due to the smaller amount of available resources. As a result, the number of users admitted to GERAN decreases with respect to the first case, and some users are moved to UTRAN as it can be appreciated in Figure 13(b). The higher load level supported by UTRAN in this second case is at the origin of the reduction in the UTRAN fittingness factor shown in Figure 13(a) (the amount of UTRAN resources is the same in both cases).

In general the same performance trends have been obtained by means of the testbed than by means of the simulations, as presented in [41].

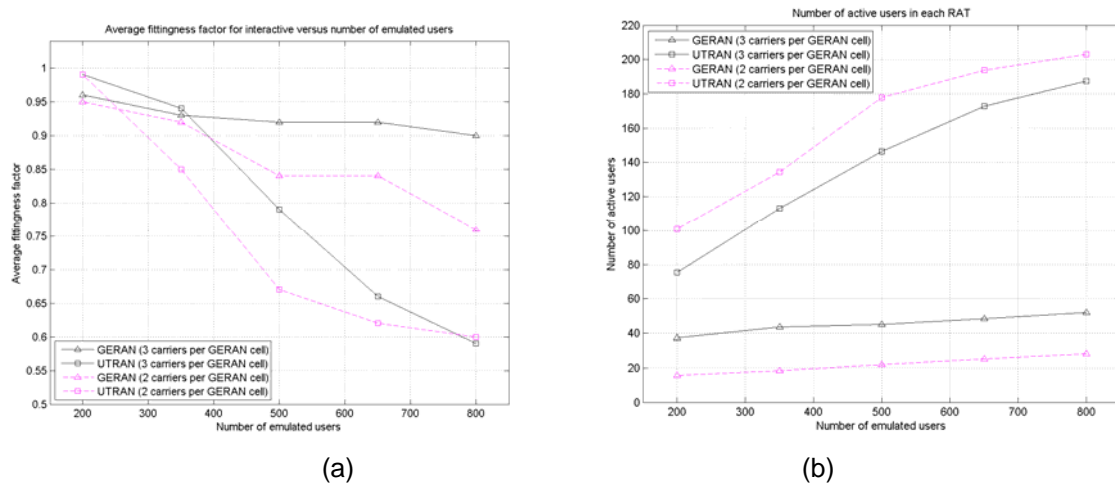


Figure 13: (a) Average fittingness factor in each RAT (b) Number of active users in each RAT.

6.5 Techno-economic evaluation of the fittingness factor-based CRRM algorithm

In order to complete the analysis of the proposed fittingness factor-based CRRM framework, a techno-economic analysis has been carried out trying to identify the potential economic benefits of applying the algorithm. The techno-economic evaluation has been carried out by assuming an increase of the data traffic demands for the next 5 years which requires new investments on network resources and by analyzing the potential savings offered by the CRRM algorithm with respect to the case when it is not used. In both cases, the specific investments considered in order to increase the capacity of the network to fulfil the traffic increase consist in the upgrade of the already present UTRAN sites by activating the second UTRAN carrier.

The considered scenario consists in an already deployed 2G/3G heterogeneous network which offers the radio coverage in a specific area by means of GSM and UMTS co-site cells supporting HSDPA. Within this scenario, the following three main categories of mobile terminals diffused in the market nowadays are supposed:

- Category 1: single-mode "GSM-only" mobile terminals (i.e. 2G terminals)
- Category 2: dual-mode "GSM/UMTS-R99" mobile terminals (i.e. 3G terminals, not HSDPA capable)
- Category 3: multi-mode "GSM/UMTS-R99/HSDPA" mobile terminals (i.e. 3G terminals, HSDPA capable)

In the considered case study, it was assumed that subscribers having GSM terminals request only the voice service (neglecting the case of requests of data services over GPRS or EDGE, due to the fact that packet data services are provided in a best effort way on these technologies and this deals with no impacts on the voice capacity). On the other hand, 3G users are supposed to be able to request voice, video call as well as web browsing services.

When the CRRM algorithm is absent (reference case), the prearranged camping strategy assumes that 2G terminals camp on GSM, and that 3G terminals always camp on the UMTS system when a suitable UTRAN cell is available from a radio quality point of view, so that they camp on GSM only in lack of 3G coverage. This strategy corresponds to what most often happen nowadays in real network scenarios, where both 2G and 3G systems coexist in the same area. Moreover, in the reference case it was assumed that users camped on UMTS which request the www services are always allocated on HSDPA, on condition that they own a HSDPA capable terminal. Instead, video call service is always allocated on UMTS R99, since this real-time service requires a fixed amount of bandwidth both in uplink and downlink.

In the second case, when the CRRM algorithm is present, the selection of GSM, UTRAN dedicated transport channels or HSDPA is always performed in accordance of the fittingness factor values evaluated by the CRRM algorithm for each RAT, on the basis of the cell load of the cells.

According with the fittingness factor framework, the behaviour of the CRRM algorithms depends (also) on the definition of the network-centric suitability associated to each RAT, which represents a function that reduces the fittingness factor of the RAT depending on the amount of non-flexible load. The specific functions chosen for GSM, UMTS-R99 and HSDPA have been identified with the aim of implementing the following high-level load balancing strategy:

1. Voice calls are preferably allocated to GSM, when possible. Only when the GSM cell has no more radio resource available the user requesting voice is allocated in the UTRAN co-located cell, if available.
2. WWW connections are preferably allocated to HSDPA. Only when a high number of contemporary HSDPA users are present within a cell, the CRRM algorithm may select dedicated channels for allocating the new www request.

A Markov based analytical model has been developed and exploited to derive technical indicators concerning the performance of the heterogeneous network with and without the presence of the fittingness factor based CRRM algorithm. The blocking probability experienced by the users as well as the mean per user perceived throughput for the data service have been considered as key performance indicators (KPIs) of the network's performance. Details on the technical approach can be found in [29]. The above mentioned KPIs have been evaluated for all the 72 couples of 2G and 3G cells located in the area of interest, and per each of the 5 years taken into account.

The KPIs mentioned above have been used to estimate how many cells cannot respect the QoS constraints specified in Table 5, due to an excessive amount of traffic offered by the users. In this way we derived the number of UTRAN second carriers that has to be considered in order to support the increase of data traffic during the five years.

Table 5: QoS constraints for a pair of co-located cells.

Performance	Threshold value
Voice loss	2%
Video loss	5%
Data loss	5%
HSDPA throughput	400 kbps
Data throughput	350 kbps

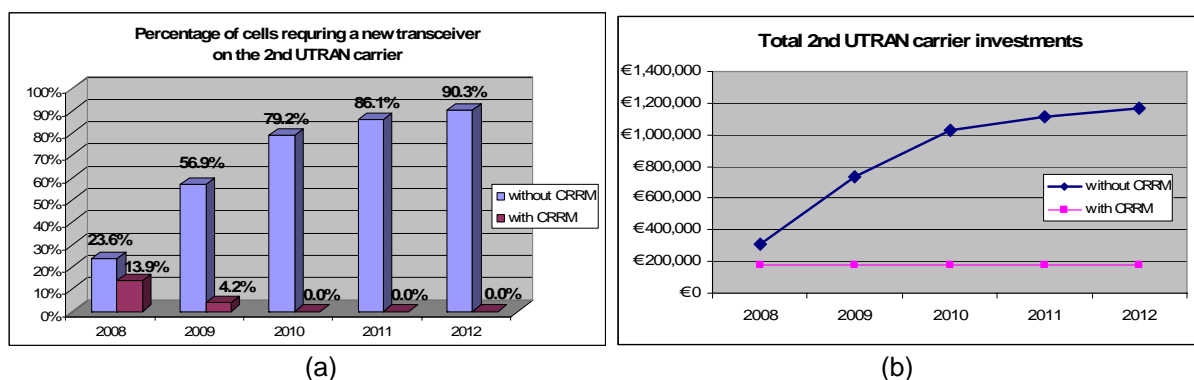


Figure 14: (a) Percentage of UTRAN cells which require the introduction of new transceivers on the 2nd carrier. (b) Total investments needed to support the offered traffic with and without the CRRM algorithm.

Results related to the number of UTRAN cells that should be upgraded by introducing the second UTRAN carrier, shown in Figure 14(a), demonstrate clearly the benefits of using the CRRM algorithm. It is worth noting that, when the CRRM algorithm is used, the number of activations estimated by considering the QoS constraints decrease after the first year. This issue can be understood by

considering that the penetration of 3G terminals (as well as the penetration of HSDPA ones) increases during the time. For this reason, also the overall network performance increases due to the higher degrees of freedom of the CRRM algorithm. Even though this phenomenon certainly happens, it should be considered that in practice the investments related to the upgrade of UTRAN cells to introduce new transceivers using the second carrier are not alienable. This practical consideration has been taken into account in the estimation of the total investments, which has been done by considering that the investments related to the UTRAN cells required in the first year of the analysis cannot be recovered in the following years. In this sense, the number of second UTRAN carrier activations for the estimation of the investments is determined by the first year. Hence, the total investments (CAPEX + OPEX) needed to introduce the additional spectrum with and without the CRRM algorithm are depicted in Figure 14(b).

Figure 14(b) clearly demonstrates the positive economic impacts offered by the CRRM algorithm in terms of investments savings. In any case, concerning this aspect, it is worth noting that the cost of the introduction of the algorithm has not been considered, since it is very difficult to estimate and it strictly depends on the specific implementations and technological choices which can be very different case by case. In any case, also by considering that an extra cost should be included for the introduction of the CRRM algorithm, the economic benefits for an operator in terms of investments savings is not jeopardized.

Finally, Table 6 reports how the investments should be spread over the considered five years (note that with the CRRM algorithm, investments are needed only in the first year, in order to upgrade the ten most critical cells with the second UTRAN carrier).

Table 6: Investments needed to introduce new transceivers on the 2nd UTRAN carrier (k€).

2nd UTRAN carrier investments (k€)	2008	2009	2010	2011	2012
without CRRM	306	432	288	90	54
with CRRM	180	0	0	0	0

By means of the figures reported in table above, it is possible also to estimate the Net Present Value (NPV)² for the two considered cases, which represents the actual value of the entire investments made over the five years. NPV of investments without and with CRRM is € 1,118,669 and € 176,471 respectively, so that the difference in terms of NPV between the two considered cases (delta NPV) is equal to € 942,199. This figure can be considered as the last key indicator of the positive economic impacts offered by the CRRM algorithm.

6.6 Implementation of the fittingness factor-based CRRM algorithm

Based on the implementation methodology described in section 2.5, in the following the implementation considerations of the fittingness factor-based CRRM algorithm are presented. Specifically, the development of an algorithm such as the fittingness factor-based would require the modification of the specifications in order to incorporate new metrics required by the algorithm. In this way, a potential implementation of the fittingness factor-based CRRM algorithm according to the example expressions given in [18] for the technical suitability function may rely on the following aspects:

- Both directed retry and inter-system handover procedures would be used for distributing connected terminals between RATs according to the computation of the fittingness factor metrics for UTRAN and GERAN cells.
- The fittingness factor metrics would be computed for each terminal in its serving radio network controller (i.e. the radio network controller ending the RRC connection).
- The computation of the fittingness factor metric of candidate cells from other RATs requires the knowledge of hard constraints (i.e. capabilities) and the availability of both inter-RAT macroscopic and microscopic measurements in the serving radio network controller.

² A WACC of 2% is considered.

- Hard constraints considered in the fittingness factor, which are posed by the characteristics of either the terminal or the technology, can be known in the serving radio network controller from information provided by the terminal and O&M for the network side.
- Microscopic measurements currently considered (i.e. path loss) can be obtained in the serving radio network controller from the terminal itself through available inter-RAT measurement reporting mechanisms.
- Exchange of macroscopic measurements among radio network controllers should be incorporated in the specifications. In particular, and attending to the definition of the macroscopic components given in [18], the set of metrics needed for the fittingness factor-based algorithm are summarised in Table 7. It is worth mentioning that for the local computations (e.g. UTRAN fittingness factor computed at the RNC or GERAN fittingness factor computed at the BSC) the required parameters and measurements would be the same included in the table.

Table 7: Locally available measurements and measurements to be exchanged among controllers for the computation of the fittingness factor-based CRRM algorithm.

	Service	Configuration parameters (e.g. setting up via O&M)	Locally available Measurements	Inter-RAT measurements exchanged among controllers
Fittingness factor for UTRAN cells computed in BSC	Voice	<ul style="list-style-type: none"> • Ratio between maximum power and noise power in UL • Orthogonality factor • Maximum power per connection in DL • Maximum DL cell power • Service bit rate in UL/DL • Eb/No target in UL/DL 	<ul style="list-style-type: none"> • Path Loss to UTRAN cells • Ec/No measurements to UTRAN cells 	<ul style="list-style-type: none"> • UTRAN UL cell load factor • UTRAN DL total transmitted power
	Interactive	<ul style="list-style-type: none"> • Ratio between maximum power and noise power in UL • Orthogonality factor • Maximum power per connection in DL • Maximum DL cell power • Service bit rate in UL/DL • Eb/No target in UL/DL 	<ul style="list-style-type: none"> • Path Loss to UTRAN cells • Ec/No measurements to UTRAN cells 	<ul style="list-style-type: none"> • Multiplexing factor for UTRAN traffic depending on service and profile.
Fittingness factor for GERAN cells computed in RNC	Voice	<ul style="list-style-type: none"> • Maximum path loss allowed in a GERAN cell. • Transmitted power in BCCH channel 	<ul style="list-style-type: none"> • Path Loss to GERAN cells 	<ul style="list-style-type: none"> • None
	Interactive	<ul style="list-style-type: none"> • Maximum path loss allowed in a GERAN cell. • Transmitted power in BCCH channel 	<ul style="list-style-type: none"> • Path Loss to GERAN cells 	<ul style="list-style-type: none"> • Multiplexing factor for GPRS traffic.

7 CONCLUSIONS

This White Paper has presented the main results and conclusions obtained in the AROMA project regarding the RAT selection problem in heterogeneous wireless networks. Probably one of the main conclusions is that the RAT selection problem is in general very complex and accounts for many variables depending on all the possible heterogeneities arising in each particular scenario. Consequently, CRRM solutions should try to capture all these variables in a general framework

flexible enough to accommodate the different operator criteria and to cope with the particular considerations of each situation. In this way, enhanced service provision and improved network capacity can be achieved, which can eventually turn into investment savings for operators.

According to the above considerations, this document has presented the different methodological approaches followed in the project to deal with the different solutions. As a starting point, analytical studies were performed to identify those parameters relevant at the radio level to optimise the distribution of users in different RATs, accounting for path loss and service considerations. Then, after these analytical studies, specific algorithms covering individually some important variables in the RAT selection were analysed making use of system level simulations. In particular, CRRM strategies exclusively based on radio quality and radio coverage in accordance with the inter-RAT cell re-selection and inter-RAT handover procedures and the corresponding parameters governing these procedures were presented. It was obtained on the one hand, that, by modifying these parameters different traffic splitting between RATs were achieved and, on the other hand, that the vertical handover procedure can be effectively exploited in order to take advantage of GERAN as a back-up system when the radio quality of UTRAN cell is not able to support user's service, which occurs, for instance, in case of indoor users.

Continuing with the system level simulations, further studies have considered a heterogeneous system with 2G, 3G and WLAN radio access networks, exploiting the perceived user throughput as a driver to perform RAT selection. It has been obtained that, by using proper CRRM mechanisms, system capacity improvements in the order of 10-40% compared to a manual RAT selection reference case were achieved.

Furthermore, and in order to introduce the different views that the different players, i.e. operators and users, may have in the RAT allocation process, together with specific radio network considerations, a Cost Function has been developed to be used by RRM and CRRM entities on cellular heterogeneous networks. It includes a wide range of KPIs, which take both user's and operator's perspectives into account, and can serve as the basis to implement different CRRM policies, since all BSs and terminals will be marketed by their own cost on the network. Thus, it is easy to compare and classify the most relevant nodes in the radio network, enabling the creation of candidate lists for a given criterion. The service priority scheme has an important impact on results, since this mechanism switches terminals services to a given RAN, therefore, being responsible for the load distribution factor within the CRRM domain.

From all the previous results, and trying to combine all the relevant conclusions obtained, a general framework capturing all the different aspects involved in the RAT selection process has been presented based on the so-called fittingness-factor, which is a new metric introduced in AROMA. As a result, a generic CRRM framework that comes up with suitable RAT selection principles under any possible circumstance is obtained. The fittingness-factor definition is split in different terms reflecting the main levels in the RAT selection, namely the terminal and network capabilities, the technical suitability at the radio part, which considers both a macroscopic and a microscopic component, the technical suitability at the transport part, and the operator/user preferences, which allows enforcing different operator policies in the decision. Using this new metric, which can be regularly updated with the user and network measurements, a specific algorithm for RAT selection has been proposed. It can be executed both at session initiation as well as during the session lifetime, so that vertical handover procedures and even horizontal handover procedures can be triggered.

The evaluation of the proposed strategy, both by means of system level simulations and by using the AROMA real time emulator, has revealed its ability to split the traffic in the considered RATs reflecting the variations in the propagation, interference and network load, and accounting for different operator preferences. As a result of this, by taking the appropriate decisions in each case, the performance of the different services can be improved. For example, delay reductions of more than 50% for interactive business users can be obtained when compared to a simple load balancing case. Similarly, through different settings of the fittingness factor it has been shown that the operator can give more or less precedence to e.g. services that can be only supported in one technology, such as the videocall traffic. Furthermore, by including transport network considerations in the fittingness factor definition, combined transport-aware RAT selection and cell selection schemes can be deployed, which turn to be beneficial in scenarios where a given BS or a limited group of BSs can experience transport

limitations, thus reducing the packet loss ratio and corresponding delay in around 70% and 30%, respectively, with respect to the case in which these strategies are not enabled.

On the other hand, the evaluation methodology of the proposed approach has also included techno-economical studies to identify the potential saving investments achieved through the application of the algorithm. This has been carried out by assuming an increase of the data traffic demands for the next 5 years which requires new investments on network resources and by analyzing the potential savings offered by the CRRM algorithm with respect to the case when it is not used. It has been obtained that the application of the CRRM algorithm can significantly reduce the required investments in a five year period, with reductions of around 84% with respect to the case without CRRM.

Finally, and after the evaluation through different means of the proposed algorithms, the targeted nature of the AROMA project is reflected on the interest to deal with algorithm implementation aspects. In this respect, the basic methodology to be followed in such implementation feasibility studies has been identified, together with the corresponding considerations related with the fittingness factor based framework.

8 ACKNOWLEDGEMENTS

This work has been performed in the framework of the IST AROMA project (<http://www.aroma-ist.upc.edu>), which is funded by the European Commission. The authors would like to acknowledge the contributions of their colleagues from the AROMA project.

9 REFERENCES

- [1] Advanced Resource Management Solutions for Future All IP Heterogeneous Mobile Radio Environments (AROMA) Project, <http://www.aroma-ist.upc.edu>, IST-4-027567
- [2] A.K. Salkintzis, "Interworking techniques and architectures for WLAN/3G integration toward 4G mobile data networks", *IEEE Wireless Communications*, Vol.11, No.3, pp. 50-61, June, 2004.
- [3] E. Gustafsson, and A. Jonsson, "Always Best Connected", *IEEE Wireless Communications*, Vol. 10, No. 1, pp. 49-55, Feb. 2003.
- [4] J. Pérez-Romero, O. Sallent, R. Agustí, and M. Díaz-Guerra, *Radio Resource Management Strategies in UMTS*, John Wiley & Sons, 2005.
- [5] 3GPP TR 25.891 v0.3.0, "Improvement of RRM across RNS and RNS/BSS (Post Rel-5) (Release 6)".
- [6] 3GPP TR 25.881 v5.0.0 "Improvement of RRM across RNS and RNS/BSS".
- [7] A. Tölli, P. Hakalin, H. Holma, "Performance Evaluation of Common Radio Resource Management (CRRM)", *Proc. of ICC'2002 - IEEE International Conference on Communications*, New York, USA, May 2002.
- [8] J. Pérez-Romero, O. Sallent, and R. Agustí, "Policy-based Initial RAT Selection algorithms in Heterogeneous Networks" *Proc. of MWCN'2005, 7th IFIP International Conference on Mobile and Wireless Communication Networks*, Marrakech, Morocco, September, 2005.
- [9] G. Fodor, A. Furuskär, and J. Lundsjo, "On access selection techniques in always best connected networks," *ITC Specialist Seminar on Performance Evaluation of Wireless and Mobile Systems*, August, 2004.
- [10] A. Tölli, P. Hakalin, "Adaptive load balancing between multiple cell layers", *Proc of VTC-Fall 2002, 56th IEEE Vehicular Technology Conference in Fall*, Vancouver, Canada, September, 2002.

- [11] A. Pillekeit, F. Derakhshan, E. Jugl, A. Mitschele-Thiel, "Force-based load balancing in co-located UMTS/GSM networks", *Proc of VTC-Fall 2004 - 60th IEEE Vehicular Technology Conference in Fall*, Los Angeles, USA, September, 2004.
- [12] X. Gelabert, J. Pérez-Romero, O. Sallent, R. Agustí, "On the suitability of Load Balancing Principles in Heterogeneous Wireless Access Networks", *Proc. of WPMC'2005 - Wireless Personal Multimedia Communications Symposium*, Aalborg, Denmark, September, 2005.
- [13] S. J. Lincke "Vertical handover policies for common radio resource management", *International Journal of Communication Systems* 2005; Published Online: 15 Mar 2005. DOI: 10.1002/dac.715.
- [14] S. Lincke-Salecker, "The Benefits of Load Sharing when Dimensioning Networks", *Proc. of ANSS'04 - 37th Annual Simulation Symposium*, April, 2004.
- [15] J. Pérez-Romero, O. Sallent, R. Agustí, "Network Controlled Cell Breathing in Multi-service heterogeneous CDMA/TDMA scenarios", *Proc of VTC-Fall 2006 - 64th IEEE Vehicular Technology Conference in Fall*, Montreal, Canada, September, 2006.
- [16] J. Pérez-Romero, O. Sallent, R. Agustí, P. Karlsson, A. Barbaresi, L. Wang, F. Casadevall, M. Dohler, H. González, F. Cabral-Pinto "Common Radio Resource Management: Functional Models and Implementation Requirements", *Proc. of PIMRC'2005 - 16th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Berlin, Germany, September, 2005.
- [17] W. Zhuang, Y. Gan, K. Loh, K. Chua, "Policy-Based QoS Management Architecture in an Integrated UMTS and WLAN Environment", *IEEE Communications Magazine*, November, 2003, pp. 118-125.
- [18] J. Pérez-Romero, O. Sallent, R. Agustí, L. Wang, H. Aghvami "A Novel Algorithm for Radio Access Technology Selection in Heterogeneous B3G networks", *Proc of VTC-Spring 2006 - 63th IEEE Vehicular Technology Conference in Spring*, Melbourne, Australia, May, 2006.
- [19] S. Lincke-Salecker and C. S. Hood, "Integrated Networks That Overflow Speech And Data Between Compnets", *International Journal of Network Management*, John Wiley & Sons, vol 12. issue 4, pp. 235-257, July-Aug. 2002.
- [20] A. Furuskär and J. Zander, "Multiservice allocation for multiaccess wireless systems," *IEEE Transactions on Wireless Communications*, vol.4, no.1, pp. 174- 184, January, 2005
- [21] I. Koo, A. Furuskär, J. Zander and Kiseon Kim, "Erlang capacity of multiaccess systems with service-based access selection," *IEEE Communications Letters*, vol.8, no.11, pp. 662- 664, November, 2004.
- [22] G. Bolch, S. Greiner, H. Meer, and K.S. Trivedi, *Queueing networks and Markov chains: modelling and performance evaluation with computer science applications*, John Wiley & Sons, 1998.
- [23] J. Pérez-Romero, O. Sallent, R. Agustí, "On the optimum traffic allocation in heterogeneous CDMA/TDMA networks", *IEEE Transactions on Wireless Communications*, Vol. 6, No. 9, pp.3170-3174, September, 2007.
- [24] J. Pérez-Romero, O. Sallent, R. Agustí, "Enhanced Radio Access Technology Selection exploiting Path Loss Information", *Proc. of PIMRC'2006 - 17th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Helsinki, Finland, September, 2006.
- [25] X. Gelabert, J. Pérez-Romero, O. Sallent, R. Agustí "A 4-Dimensional Markov Model for the Evaluation of Radio Access Technology Selection Strategies in Multiservice Scenarios", *Proc of VTC-Fall 2006 - 64th IEEE Vehicular Technology Conference in Fall*, Montreal, Canada, September, 2006.
- [26] A. Serrador (editor) et al. "Simulation tools: final version capabilities and features", IST AROMA project, Deliverable D13, EC-IST Office, Brussels, Belgium, April 2007.

- [27] R. Ljung, (editor) et al., Target Scenarios specification: vision at project stage 1, IST-AROMA Project, Del. D05, EC-IST Office, Brussels, Belgium, Mar. 2006.
- [28] Anna Umbert (editor) et al., "D07 Testbed Specification", IST-AROMA Project, Deliverable D07, EC-IST Office, Brussels, Belgium, July, 2006.
- [29] R. Farotto (editor) et al., "Economic Evaluation of novel AROMA RRM/CRRM Algorithms and Solutions", IST-AROMA Project, Deliverable D17, EC-IST Office, Brussels, Belgium, December, 2007.
- [30] O. Sallent (editor) et al., "First report on AROMA algorithms and simulation results", IST-AROMA Project, Deliverable D09, EC-IST Office, Brussels, Belgium, November, 2006.
- [31] 3GPP TS 25.331 "Radio Resource Control (RRC) protocol specification".
- [32] 3GPP TS 45.008 v7.3.0, "Radio Access Network; Radio subsystem link control (Release 7)"
- [33] 3GPP TS 44.018 v7.5.0 "Mobile radio interface layer 3 specification; Radio Resource Control Protocol (Release 7)".
- [34] 3GPP TS 25.304, "UE Procedures in Idle Mode and Procedures for Cell Reselection in Connected Mode (Release 99)".
- [35] L. Ferreira, A. Serrador and L.M. Correia, "Concepts of Simultaneous Use in Mobile and Wireless Communications", *Wireless Personal Communications, Special Issue on Elective Topics towards Vision to Reality of the Wireless Future*, Vol. 37, No 3, May 2006, pp. 317-328.
- [36] A. Serrador, L.M. Correia, "A Cost Function for Heterogeneous Networks Performance Evaluation Based on Different Perspectives", *Proc. of IST'2007 - 16th IST Mobile & Wireless Communications Summit*, Budapest, Hungary, July, 2007.
- [37] A. Serrador, L.M. Correia, "Policies for a Cost Function For Heterogeneous Networks Performance Evaluation", *Proc. of PIMRC'2007 - 18th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Athens, Greece, September, 2007.
- [38] J. Pérez-Romero (editor) et al., "Final report on AROMA algorithms and simulation results", IST-AROMA Project, Deliverable D18, EC-IST Office, Brussels, Belgium, October, 2007.
- [39] J. Pérez-Romero, O.Sallent, R.Agustí, "A Novel Metric for Context-Aware RAT Selection in Wireless Multi-Access Systems", *Proc. of ICC'2007 - IEEE International Conference on Communications*, Glasgow, Scotland, June, 2007.
- [40] O.Sallent, J. Pérez-Romero, R.Ljung, P.Karlsson, A. Barbaresi, "Operator's RAT Selection Policies Based on the Fittingness Factor Concept", *Proc. of IST'2007 - 16th IST Mobile & Wireless Communications Summit*, Budapest, June, 2007.
- [41] A. Umbert (editor) et al., "Trial results and algorithm validation", IST-AROMA Project, Deliverable D20, EC-IST Office, Brussels, Belgium, December, 2007.