Università degli Studi di Bologna
Dipartimento di Elettronica Informatica e Sistemistica, DEIS

Dottorato di Ricerca in Ingegneria Elettronica,
Informatica e delle Telecomunicazioni
XIX Ciclo

Wireless Heterogeneous Networks

Tesi di:
Ing. Claudio Gambetti

Coordinatore:
Chiar.mo Prof. Ing. Paolo Bassi

Relatore:
Chiar.mo Prof. Ing. Oreste Andrisano

Settore scientifico-disciplicare:
ING-INF/03 TELECOMUNICAZIONI

Contents

List of Tables

List of Figures

Introduction

In the last years, mobile communications have become pervasive to all activities of society; the number of mobile phones and wireless Internet users has increased significantly. Changing private and professional lifestyles has created a surging demand for *communications on the move*, *reachability* and *wireless broadband*.

In fact, latest industrial surveys[1] reveal 2.255 billion subscriptions in the family of all fully open standard cellular technologies GSM, GPRS, EDGE and WCDMA-HSDPA (31.12.2006); 511 million of these subscriptions were just added in 2006, corresponding to a growth of 29% in 2006. The third generation networks (WCDMA-HSDPA) now count almost 100 million subscriptions (31.12.06), with an average monthly growth in 2006 of over 4 millions and an annual growth of more than 100% in 2006; the forecast is that by end-2009, WCDMA-HSxPA subscriptions will be half billion.

Mobile networks evolution

Among all these different mobile technologies, traditionally, first generation (in Italy it was TACS - *Total Access Communication System*) and second generation (GSM - *Global System for Mobile Communications*) wireless networks were primarily targeted at voice communications and finally at data communications occurring at low data rates. The first-phase GSM specifications provided only basic transmission capabilities for the support of

[1]GSA - Global mobile Suppliers Association - representing GSM/EDGE/WCDMA suppliers globally; mobile subscribers data source: Informa Telecoms & Media.

data services, with the maximum data rate in these early networks being limited to 9.6 Kbps on one timeslot in each radio frame.

About 5 years ago, the most advanced cellular technology for mobile internet access became the GSM implementing the *High-Speed Circuit-Switched Data* (HSCSD) evolution, specified in ETSI Rel'96; it was the first GSM Phase 2+ work item that clearly increased the achievable data rates in the GSM system. The maximum radio interface bit rate of an HSCSD configuration with 14.4 Kbps channel coding (corresponding to the best radio conditions) multiplying 4 timeslots is 57.6 Kbps: this was broadly equivalent to providing the same transmission rate as that available over one ISDN B-Channel. It seems prehistory, but just 5 years ago this was a great achievement!

Quickly, the GSM networks were upgraded to 2.5G by introducing the *General Packet Radio System* (GPRS) technology. GPRS provides GSM with a packet data air interface and an IP-based core network, with bit rates varying from 9 Kbps to more than 150 Kbps per user when all eight timeslots of a GSM carrier are assigned to a single GPRS Mobile Station for exclusive use.

The *Enhanced Data Rates for Global Evolution* (EDGE) was a further innovation step of GSM packet data and now EDGE is widely deployed on global GSM networks. Thanks to the 8 phase shift keying (8PSK) modulation scheme, EDGE can handle about three times more data subscribers than GPRS, or triple the data rate for one end-user. EDGE is specified in a way to enhance the throughput per timeslot for both HSCSD and GPRS. The enhancement of HSCSD is called ECSD (*Enhanced Circuit Switched Data*), whereas the enhancement of GPRS is called EGPRS (*Enhanced General Packet Radio System*). In ECSD, the maximum data rate does not exceed 64 Kbps because of the restrictions in the A-interface, but the data rate per timeslot triples. Similarly, in EGPRS, the data rate per timeslot triples and the peak throughput, with all eight timeslots in the radio interface, can reach 473 Kbps.

On the other hand, in the last couple of years, we have also seen a strong development of third-generation (3G) wireless systems that incorporate the features provided by broadband. In addition to support mobility, broadband aims at supporting multimedia traffic, with quality of service (QoS) assurance; four class of services are considered: conversational (both speech and video calls), streaming, interactive and background. In Europe, the 3G standard has been initially developed by ETSI (European Telecommunication Standard Institute), then the work has been continued by Third Generation Partnership Project (3GPP) under the designation of UMTS (*Universal Mobile Telecommunications System*).

The radio access interface of UMTS comprises two standards for operation adopting the *Frequency Division Duplex* (FDD) and *Time Division Duplex* (TDD) modes. The present UMTS FDD networks, based on wideband CDMA (WCDMA), adopt new spectrum and new radio network configurations while using the same GSM core infrastructure. The maximum data rate in the first WCDMA release (Rel'99) is 2 Mbps, but in practice the widely used maximum downlink rate (i.e. direction from NodeB to User Equipment) is equal to 384 Kbps.

With the really recent addition of the *High Speed Downlink Packet Access* (HSDPA) specified in 3GPP Rel'5, that is a sort of 3.5G technology, WCDMA network operators aim at providing extremely high data rate multimedia services and to improve spectral efficiency by higher order modulation using 16-QAM: in March 2007, some of the deployed WCDMA-HSDPA networks are starting to support calls at 7.2 Mbps gross data rate (corresponding to a 6.7 Mbps net data rate) with category 8 mobiles: about 5 years are passed since the HSCSD introduction, and the speed of data transfer over cellular networks has increased of more than 100 times! Moreover, these WCDMA-HSDPA networks will soon achieve data rates in downlink up to 14 Mbps with category 10 mobiles.

Similarly, in the next months in 2007, the *High Speed Uplink Packet Access* (HSUPA), standardized in 3GPP Rel'6, will complement HSDPA by significantly reducing latency on the uplink and offering data speeds up to 5.8 Mbps (peak) on the uplink channel. Together with HSDPA, it means a huge stride in WCDMA-HSxPA[2] network performance.

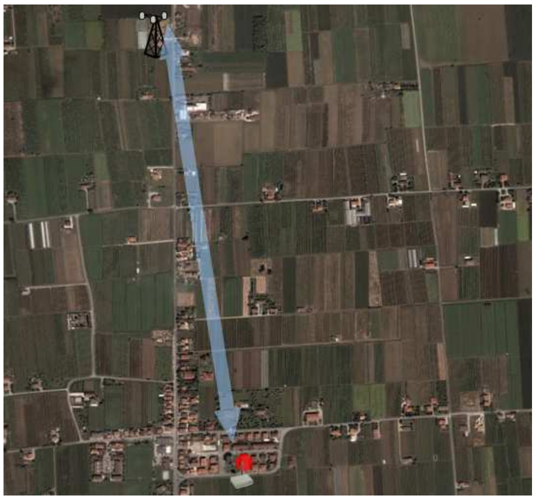
In this extremely fast changing and widened context of mobile networks, my work was initially addressed to evaluate the performance of the innovative 3G networks and to study the impact of physical layer parameters on the network performance. Based on one of the main requirements for 3G systems, that is the ability to support asymmetric uplink/downlink traffic, the choice of the 3G radio interface to be studied has been directed to one of the two TDD modes, the *Time-Division Synchronous Code Division Multiple Access* (TD-SCDMA): thanks to its TDD/TDMA characteristics, the TD-SCDMA network can adapt the uplink/downlink ratio according to the data load within a single unpaired frequency thus utilizing the spectrum more efficiently. This is especially helpful in an environment with increasing data traffic (mobile data), which tends to be asymmetric, often requiring little uplink throughput, but significant bandwidth for downloading information (mobile Internet).

Wireless heterogeneous networks

Coming back to the development of wireless networks, we can observe that some alternative operators are already offering wireless broadband Internet access with WiFi or

[2]Here with the term HS<u>x</u>PA, we mean both the HS<u>D</u>PA and the HS<u>U</u>PA technologies.

(a) WLAN AP in high traffic density area globally covered by 3G



(b) WLAN AP to improve indoor coverage

Figure 1: Scenarios for wireless heterogeneous networks

(pre-)WiMAX 802.16d networks. For example *Wireless Local Area Networks* (WLAN) are achieving a great penetration in the mass market as a really effective solution to provide mobile access to the Internet; companies all over the world are already offering WLAN connections in particular locations, such as airports, hotels or caffés (see figure 1). In these areas, the so-called "hot spots", anyone owning the appropriate technology on his laptop can connect to the Internet at a reasonable price and with a satisfactory connection speed.

Nonetheless, the request for higher bit rates is expected to further increase in the next future and more capacity will be necessary; on these conditions, 3G/WLAN interworking becomes a really significant issue to be investigated: provided that the WLAN hot spot is within the 3G network coverage and that the final user is equipped with a dual mode terminal, integrating the two technologies, thus increasing the "pool" of available resources, would considerably increase both users' satisfaction and networks' utilization efficiency.

In this phase, I therefore extended my analysis from the initial scenario of a 3G "standalone" network to a full 3G/WLAN heterogeneous network: firstly, the feasibility of the integration of these two technologies in a single system has to be evaluated; afterwards, the possible methods for a *Common Radio Resource Management* of the two radio access networks have been studied in depth.

Thesis Outline

This dissertation mainly deals with the study of third-generation and wireless heterogeneous networks.

Chapter 1 presents the design of the third-generation dynamic system simulator I developed during this activity: the standard that has been chosen is the UTRA TDD 1.28 Mcps, also called TD-SCDMA. The main functional blocks composing this tool are the *Link Level simulator*, the *Network Level simulator*, an *Interface module* between the link and the network levels and finally the *Upper Layer simulator*, which receives the input from a *Mobility simulator* and a *User Activity simulator*. An overview of the main changes required to implement UTRA FDD are also shown.

Chapter 2 deals with the analysis of the performance of TD-SCDMA network through system simulation. The influence of packet switched applications on the overall network performance is investigated; in order to balance the combined quality of voice and data services, the tuning of physical layer parameters for the power control algorithm is evaluated.

Chapter 3 describes the method we propose on how to interface link and network level tools through an ”instant value” interface. Although this approach has been adopted in the TD-SCDMA system simulator, it's quite general and can be implemented in arbitrary wireless networks with a time slotted physical frame structure. The proposed approach allows a thorough analysis of performance in networks with mixed circuit and switched services.

Chapter 4 introduces a possible architectural solution for a heterogeneous wireless system exploiting the complementary characteristics of the radio interfaces of a third-generation network and a wireless LAN. Both a logical scheme and a feasible realization are provided; a link between each local UMTS and WLAN *Radio Resource Management* (RRM) entity and a *Common RRM* (CRRM) entity is proposed.

Chapter 5 describes the possible advantages introduced by the CRRM for a heterogeneous integrated and interworking UMTS-WLAN network; the performance evaluation is carried out through full simulations from the physical to the application layer. Firstly, the required interactions and information exchanged among the CRRM entity and the local RRM entities are presented; afterwards a fully configurable CRRM algorithm is proposed: the project is composed of a *Coverage Based*, a *Service Based* and a *Quality of Service (QoS) Based* CRRM algorithm. Finally, the trend in the system capacity provided with the various CRRM options is discussed in a realistic scenario.

Chapter 1

TD-SCDMA System Simulator design

Research activities on current and future communication systems are more and more carried out by means of simulation tools, either developed ad hoc for specific purposes or already existing, such as Opnet [1], ns-2 [2], Glomosim [3].

This trend is mainly due to the increasing complexity of current and forthcoming technologies as well as to the frequent need of complete investigations, embracing the whole protocol stack (from physical to application levels), for which simulation is the only feasible way to get some insights of the performance provided to the final user.

However, as emphasized also in [4], the realization of a reliable simulator of a communication system is quite a tricky task, especially when wireless technologies are concerned which require an accurate modelling of the physical level as well as an adequate characterization of the radio channel behavior.

Off-the-shelf network simulation tools, such as the aforementioned Opnet, ns-2 and Glomosim, generally adopt simplified approaches to model the physical level behavior of the supported wireless technologies. Usually they only account for the path-loss effect or at most a simplified channel model; in all cases, however, accurate bit-level simulations are neglected [4, 5, 6, 7].

At first sight, this choice is acceptable since bit level simulations, which require an accurate implementation of all physical level aspects of the investigated technology (propagation, channel coding and decoding, interleaving, modulation and demodulation,etc.), are very consuming in terms of simulation time. Nonetheless the possible lack of accuracy of physical level characteristics of such tools is felt as problem to be overcome by many researchers [6, 7, 8] and some effort has been made in this direction [9].

Let us emphasize that in this work *Physical Level simulators* are conceived as a part of complete system simulators aimed at reproducing the entire protocol stack (from application to propagation). Within this context, the task of a physical simulator is no more to simply provide curves of bit error rate or packet error rate characterizing the performance of the investigated technology at physical level, but, on the contrary, its task is to interact with the simulation tool which reproduces the upper layers behaviors (from MAC to Application), hereafter denoted as *Network Level simulator*. Let us observe that this task has to be performed for each user within the investigated scenario, that is, for a number of links which could be very relevant.

Moreover, without such an integrated approach from physical to application level, the simulation of advanced wireless heterogeneous networks, which is the main objective of this thesis, would be quite rough: actually, the final direction of our study is to investigate whether a Common Radio Resource Management entity (see chapter 5) could better the system capacity, by exploiting in real time the complementary characteristics offered by the different radio access technologies. In this context, it's therefore strongly required to build for each radio access stratum a System Simulator reproducing with accuracy the main characteristics of the physical layer, as well the main aspects of the related datalink layer, the local Radio Resource Management entity and the upper layer properties.

In this first chapter, the design of the 3G (third generation) system simulator I developed during the thesis is presented.
In section 1.1, an overview of the technology is given, with reference to the selected UMTS standard, the UTRA TDD 1.28 Mcps option, ordinary called TD-SCDMA. Afterwards, in section 1.2 the design of the TD-SCDMA system simulator is described, whose main functional block are the Link Level simulator, the Network Level simulator and the *Upper Layer simulator*. Finally in section 1.3, the main changes required to implement UTRA FDD will be shown.

1.1 TD-SCDMA air interface

TD-SCDMA, which stands for Time Division Synchronous Code Division Multiple Access, is an innovative mobile radio standard for the physical layer of a 3G air interface. It has been adopted by ITU and by 3GPP as part of UMTS release 4, becoming in this way a global standard, which covers all radio deployment scenarios: from rural to dense urban areas, from pico to micro and macrocells, from pedestrian to high mobility.
TD-SCDMA combines an advanced TDMA/TDD system with an adaptive CDMA component operating in a synchronous mode.

TD-SCDMA offers several unique characteristics for 3G services. In particular its TDD nature allows TD-SCDMA to master asymmetric services more efficiently than other 3G standards (for example, the ordinary UTRA FDD, known as W-CDMA from the ITU terminology). Up and downlink resources are flexibly assigned according to traffic needs, and flexible data rate ranging from 1.2 Kbit/s to 2Mbit/s are provided. This is especially helpful in an environment with increasing data traffic (mobile data), which tends to be asymmetric, often requiring little uplink throughput, but significant bandwidth for downloading information (mobile Internet).

Many radio technology, such as GSM, EDGE, W-CDMA or cdma2000, require separate bands for uplink and downlink (paired FDD spectrum). In this case with asymmetric loads, such as Internet access, portions of the spectrum are occupied but not used for data transfer. These idle resources cannot be utilized for any other service, leading to an inefficient use of the spectrum. On the contrary, TD-SCDMA adapts the uplink/downlink ratio according to the data load within a single unpaired frequency band, thus utilizing the spectrum more efficiently.

Highly effective technologies like smart antennas, joint detection and dynamic channel allocation are integral features of the TD-SCDMA radio standard. They contribute to minimize intra-cell interference (typical of every CDMA technology) and inter-cell interference leading to a considerable improvement of the spectrum efficiency. This is especially helpful in high-populated areas, which are capacity driven and require an efficient use of the available spectrum. TD-SCDMA can also cover large areas (up to 40 Km) and supports high mobility. It is therefore well suited to provide mobile services to subscribers driving on motorways or travelling on high-speed trains.

In order to mitigate the effect of interference and improve the coverage at the cells edge, conventional CDMA 3G systems have to use the so-called soft handover when an ongoing call needs to be transferred from one cell to another as a user moves through the coverage area. During soft handover, however, the users terminal has concurrent traffic connections with more than one base station. To handle this increased traffic more

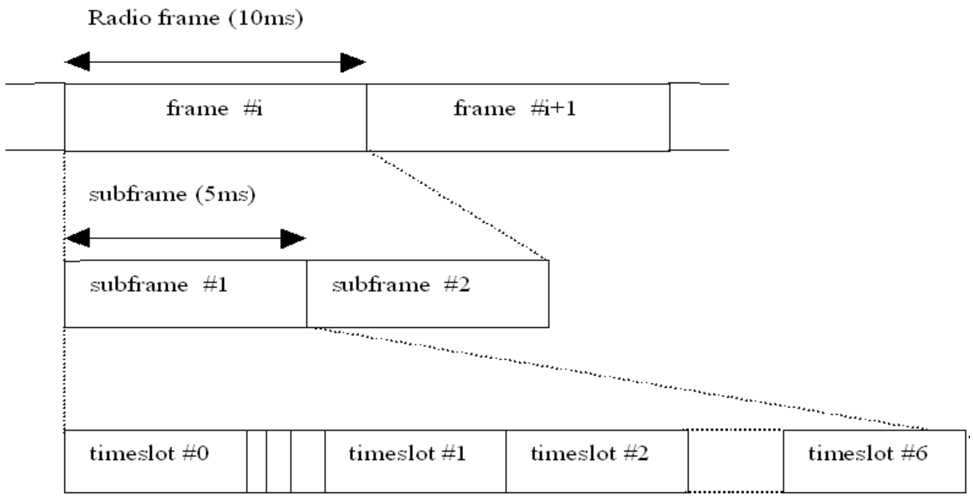


Figure 1.1: TD-SCDMA physical channel signal format

channel units and leased lines are required, resulting in higher operating costs. Thanks to joint detection, smart antennas and an accurate terminal synchronization TD-SCDMA does not need to rely on soft handover.

Here the basic technological principles on which the TD-SCDMA technology is based are summarized:

- **TDD (Time Division Duplex)** allows uplink and downlink on the same frequency band and does not require paired bands. In TDD, uplink and downlink are transmitted in the same frequency channel but at different times. It is possible to change the duplex switching point and move capacity from uplink to downlink or vice versa, thus utilizing spectrum optimally. It allows for symmetric and asymmetric data services.
  For symmetric services used during telephone and video calls, where the same amount of data is transmitted in both directions, the time slots are split equally between the downlink and uplink.
  For asymmetric services used with Internet access (download), where high data volumes are transmitted from the base station to the terminal, more time slots are used for the downlink than the uplink.

- **TDMA (Time Division Multiple Access)** is a digital technique that divides each frequency channel into multiple time-slots and thus allows transmission channels to be used by several subscribers at the same time (see figure 1.1).
  TD-SCDMA [56] uses a 5 ms sub-frame subdivided into 7 time slots of 675 μs duration each, which can be flexibly assigned to either several users or to a single user

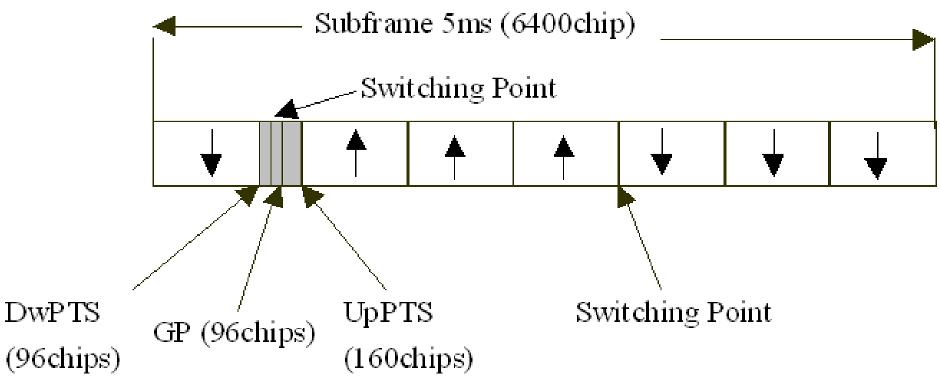


Figure 1.2: Structure of the TD-SCDMA sub-frame

who may require multiple time slots.
Each TD-SCDMA sub-frame contains 3 special time slots DwPTS (Downlink Pilot), GP (Guard Period) and UpPTS (Uplink Pilot) lasting for 75 μs (96 chips), 75 μs (96 chips) and 125 μs (160 chips) respectively (see figure 1.2). The total duration of these 3 slots is 275 μs.

- **CDMA (Code Division Multiple Access)** increases the traffic density in each cell by enabling simultaneous multiple-user access on the same radio channel. Yet each user can interfere with another, which leads to multiple access interference (MAI).
  In TD-SCDMA, within each time slot a number of up to 16 CDMA codes may be transmitted (maximum CDMA loading factor). Using a chip rate of 1.28 Mcps allows a carrier bandwidth of 1.6 MHz. According to its operating license, the network operator can deploy multiple TD-SCDMA 1.6 MHz carriers. Each radio resource unit is thus identified by a particular time slot and a particular code on a particular carrier frequency.

- **Joint Detection (JD)** allows the receiver to estimate the radio channel and works for all signals simultaneously. Through parallel processing of individual traffic streams, JD eliminates the multiple access interference (MAI) and minimizes intra-cell interference, thus increasing the transmission capacity.
  The efficiency of the Joint Detection receiver in TD-SCDMA technology is based on the TDMA/TDD operation and on the limited number of codes employed per time slot: the total number of users per radio carrier is distributed over the different time slots of the basic TDMA frame, so that a maximal number of 16 codes per time slot per radio carrier can be easily processed in parallel and detected.

To compare the two systems, in UTRA FDD 256 CDMA codes might be transmitted: due to the high number of codes, the implementation of an optimal multi-user receiver in FDD is difficult, since the implementation complexity is an exponential function of the numbers of codes. In order to combat MAI, UTRA FDD emploies suboptimal detection schemes, such as the Rake receiver, which do not extract all CDMA codes in parallel.

- **Dynamic Channel Allocation (DCA)**: the advanced TD-SCDMA air interface takes advantage of all available Multiple Access techniques (TDMA (Time Division Multiple Access), FDMA (Frequency Division Multiple Access), CDMA (Code Division Multiple Access) and SDMA (Space Division Multiple Access)).
  Making an optimal use of these degrees of freedom, TD-SCDMA provides an adaptive allocation of the radio resources according to the interference scenario, minimizing intercell interference.

- **Mutual Terminal Synchronization**: like all TDMA systems, TD-SCDMA needs an accurate synchronization between mobile terminal and base station. This synchronization becomes more complex through the mobility of the subscribers, because they can stay at varying distances from the base station and their signal presents varying propagation delays.
  Thanks to synchronization, TD-SCDMA does not need soft handover, which leads to a better cell coverage, reduced inter-cell interference and low infrastructure and operating costs.

- **Smart Antennas** are beam steering antennas which track mobile usage through the cell and distribute the power only to cell areas with mobile subscribers. Without them, power would be distributed over the whole cell. Smart antennas reduce multiuser interference; increase system capacity by minimizing intra-cell interference, increase reception sensitivity and lower transmission power while increasing cell range.

1.2 TD-SCDMA System Simulator

The TD-SCDMA system simulator is a complete home-made test-bed reproducing the main characteristics of a UMTS network, compliant to the standard UTRA TDD 1.28 Mcps, from the physical to the application layer.

Due to the complexity and the vastness of the behaviors to be reproduced for such wireless system, the overall problem has been subdivided in smaller logical blocks (see

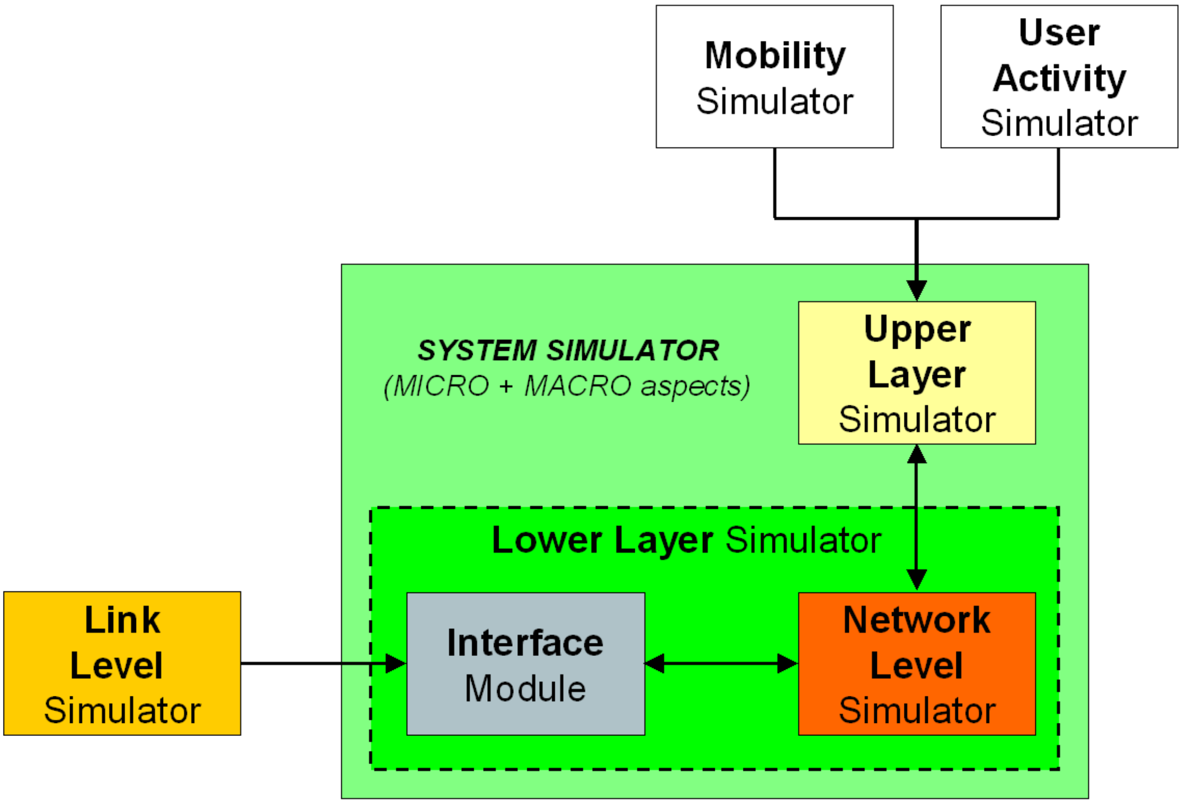


Figure 1.3: TD-SCDMA System Simulator: block diagram

figure 1.3) which can be developed in different software programs, thanks to appropriate communication interfaces.

This planning choice has many advantages: first of all, during the initial phase of software implementation, the effort of the developer is focused on smaller procedures, thus allowing a tidier work. Secondly, some parts can be easily re-used for system simulators of other wireless technologies (for example, the mobility simulator doesn't strictly depend on the radio access technology). Finally, this choice permits to easily project each logical block with a different time scale (for example, the link level simulator works with the bit, whereas the user activity simulator setups voice calls with a much larger time step), provided that there is a clear interface definition between the various blocks.

In the next subsections, a description of each logical block in figure 1.3 is given.

1.2.1 Link Level simulator

The block diagram of the implemented TD-SCDMA Link Level simulator is presented in figure 1.4.

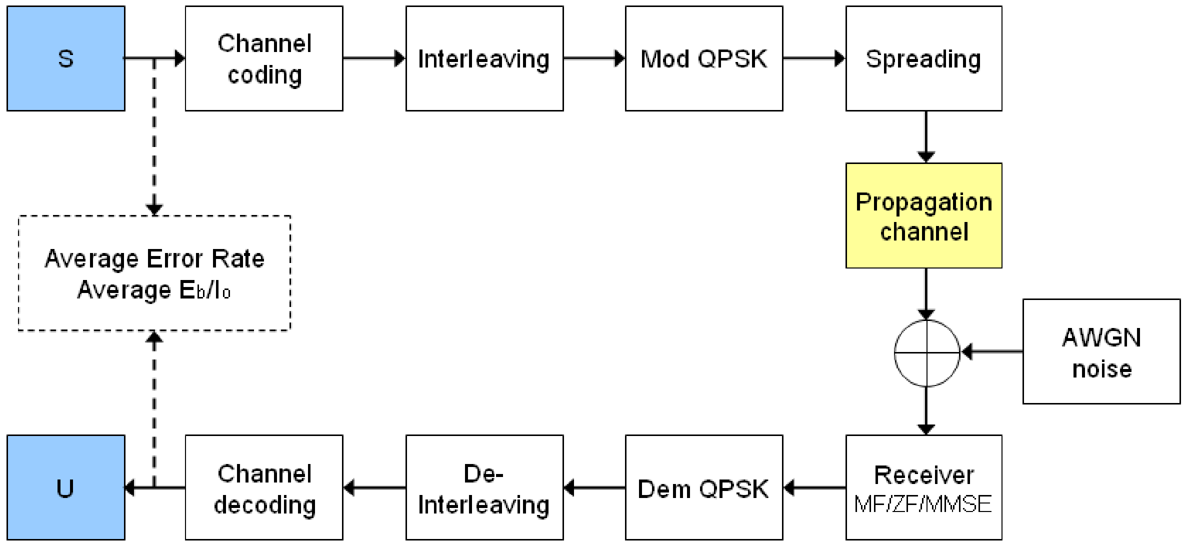


Figure 1.4: TD-SCDMA Link Level simulator: block diagram

- The block *"S"* is the source of the information to be transmitted.
  Every timeslot, the information bits are randomly generated. Each active user may transmit with different bit rates, depending on variable Spreading Factor codes (1, 2, 4, 8, 16) in uplink and depending on multi-coding till a maximum of 16 parallel codes with fixed spreading factor equal 16 in downlink.

- The *channel coding* block performs the coding operation through convolutional codes with variable code rate and constraint length [49], depending on the settings of the configuration file (in the next future, turbo-codes will be implemented).

- The *interleaving* block [49] operates at different periods: 10, 20, 40 or 80 ms.
  Let us recall that the TD-SCDMA radio frame duration is 10 ms, that is the same duration of UTRA TDD 3.84 Mcps and UTRA FDD; nevertheless, the TD-SCDMA peculiarity is that the radio frame is split in two sub-frames of duration 5 ms. In this way the interleaving procedure allocates the bits by spanning the various timeslots over at least two sub-frames.

- The *QPSK modulator* generates the QPSK symbols.
  The number of symbols per burst depends on the type of used burst [48] and on the selected spreading factor.

- The *spreading* procedure performs the "Direct Sequence Spread Spectrum" technique by multiplying the input signal with an *Orthogonal Variable Spreading Factor*

(OVSF) code [50].

- The *propagation channel* block generates the channel impulse response for AWGN channels or non-AWGN channels with frequency-selective fading and intersymbol interference (ISI). The latter uses Jake's doppler channel model for pedestrian and vehicular channels, or a FIR filter for indoor channels.

- Samples of *AWGN noise* are then summed to the signal previously generated with the channel impulse response.

- Several *receivers* have been implemented: the Matched Filter (MF), the Zero Forcing (ZF) and the Minimum Mean Square Error (MMSE) one. These receivers perform the operation of *despreading*.

- The *QPSK demodulator* carries out the de-mapping of the QPSK symbols, that is the association between each symbol of the constellation with the related couple of encoded bits.

- The *de-interleaving* block executes the dual operation of the interleaver, returning the ordered sequence of encoded bits that were previously mixed in various timeslots and subframes.

- The *channel decoding* block performs the decoding through Viterbi algorithm. The decoding can be "hard" or "soft" (that is, by addition of reliability information of the decoding).

- The last block records the result of each transmission and produce the desired output: as explained in chapter 3, different methods shall be used to generate the appropriate information for the rest of the system simulator.

1.2.2 Link-to-Network level Interface module

Unlike systems such as pure GSM (i.e. with no GPRS services) in which link and system level issues can be investigated separately, in case of third generation cellular systems, the interaction between link and system level is much more involved and a large number of link level parameters have to be considered. Such parameters are: the class of service (speech or data CS, data PS), type of ITU propagation channel, direction of the link (DL/UL), number of timeslots and codes occupied in every subframe, spreading factor (SF), code rate, interleaving period, type of receiver.

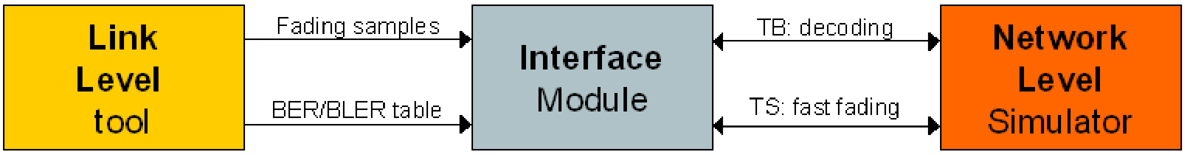


Figure 1.5: Link-to-Network level Interface module: block diagram

As already explained at the beginning of this chapter, a thorough performance investigation of the TD-SCDMA system requires the analysis of both link and network level aspects. Owing to the complexity of this approach, a unique simulation program addressing both these aspects at the same time is expected to requires an inconceivable amount of time to obtain some results.

This suggests to develop two different simulators: one for link level and another one for network level issues (see figure 1.3). While the minimum time step of the TD-SCDMA link level tool described in subsection 1.2.1 is equal to the chip duration, at network level the selected time resolution has a much longer duration, equal to the physical time slot. Therefore, in the design of the TD-SCDMA system simulator, the project of a suitable interface integrating link and network level parts and a careful definition of the parameters exchanged through the interface is required.

As explained in more detail in chapter 3, in this thesis an advanced link-to-network interface module working with "instantaneous values" has been developed; in figure 1.5, a simple diagram of the main interactions between the link and the network level through the interface module.

- First of all, the interface module receives from the link level simulator the BER/BLER look-up tables for each allowed parameter combination and the vector of fading samples for each propagation channel.

- Afterwards, during the dynamic simulation, the interface module communicates real-time with the Network Level Simulator. The exchange time is TS (Time Slot) based for fast fading enforcement or TB (Transport Block) based for BER/BLER evaluation after decoding.
  The transport block represents the elementary data unit managed by encoder/decoder blocks and we assume the TB duration equal to the interleaving period.

Since a large part of the information needed for certain calculations is already included in the network level simulator tool, the interface module could be directly implemented in the same software program, although logically separated (see the dotted rectangular containing both the network level simulator and the interface module in figure 1.3): this choice has been adopted in our project and allows to speed up the overall simulation time.

1.2.3 Network Level simulator

Different simulation approaches for UMTS network performance evaluation can be characterized.

The first one is the **static** approach: in this case, the capacity of a given UMTS network layout can be estimated based on the propagation conditions in a very fast way. This is very useful for a first rough network planning with a few iterations necessary to find a proper network layout that fulfills the needs of the operator.

Starting with a fixed downlink load the coverage area is evaluated as well as the serving area for each cell. For the obtained coverage areas the cell capacity is evaluated using a distributed UE method, which means that for each prediction pixel a fractional UE is considered. The fractional coefficient is determined by assigning the total available downlink power resources to all pixels according to their radio channel conditions. The resulting cell capacities are valid for the given downlink load and the procedure can be repeated for different loads to evaluate the dependencies between coverage and capacity.

The traffic density can be defined either homogeneous or location dependent. During the static capacity prediction the distributed UE per pixel (fractional UE) are weighted according to the traffic distribution. This enables for example that areas with poor radio channel conditions having a low expected traffic density do not decrease the cell capacity very much, despite these areas require considerable resources to be connected.

A second common possibility for TD-SCDMA network simulation is based on the well known **Monte Carlo** approach. Compared to the static approach, the Monte Carlo simulation tool provides more detailed outputs and gives a better understanding of the network behavior. Compared to the full dynamic simulations, the Monte Carlo approach is much faster, thus enabling the simulation of much larger areas.

The Monte Carlo (MC) method consists in repeating an experience many times with different randomly determined data in order to draw statistical conclusions; in mobile network case, the users are deployed in the network with random positions. The results given by a high number of tests are considered to be representative for all the possible states of the network.

The third main method for UMTS network study is the **dynamic** simulation: this is the approach that has been chosen in this work. With this methodology, time variant effects like the fast power control, the Radio Resource Management algorithms and the UE mobility are taken into account and their impact on the overall system performance

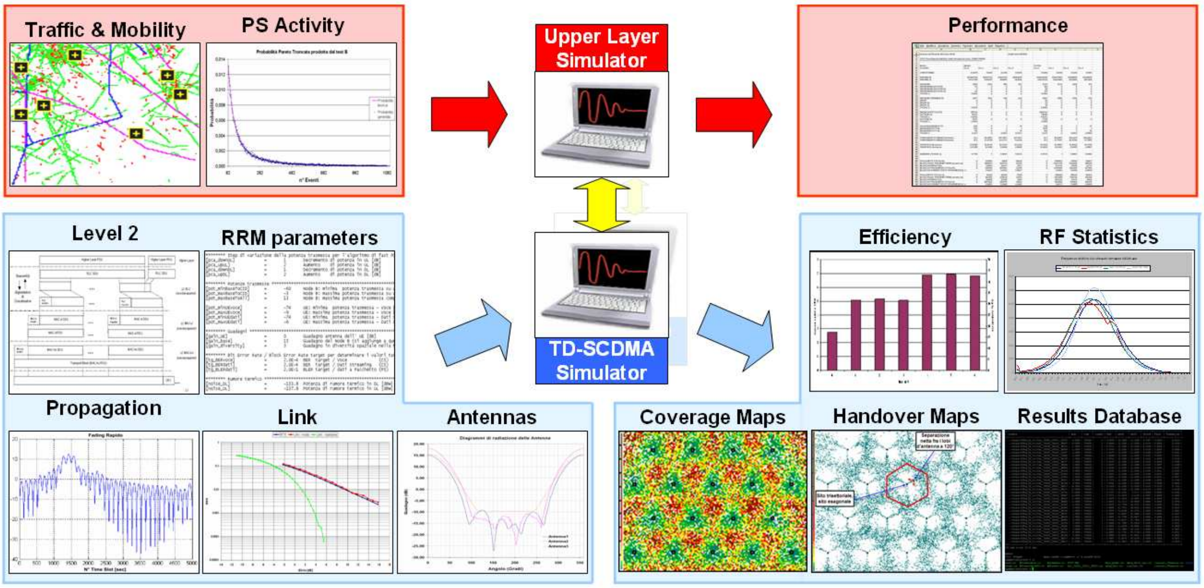


Figure 1.6: Main functional blocks of the TD-SCDMA simulator

can be studied in a very detailed manner.

For example, in this approach, a new radio connection is reproduced according to arrival and ending statistics. If there are enough radio resources available to serve the new mobile station, the initial transmission power is determined and the closed loop fast power control process is started. Considering all transmission links from all mobiles in uplink and downlink directions with their corresponding transmission powers, the interference levels can be determined for each mobile in uplink and downlink, respectively. The resulting signal-to-noise and interference ratio levels can be compared with the target levels specified within the service definition, and the required power control command is sent back to update the transmission power in the next active slot.

The main characteristics of the TD-SCDMA dynamic network level simulator developed in this study are summarized in the following (see also figure 1.6):

- The simulation step is equal to the time slot duration (675 μs): all the power measurements refers to this time resolution. Other system operations occur with larger time intervals: frame (10 ms), transport block duration (10, 20, 40 or 80 ms) or longer periods for the averaging of the measurements (0.5-1.0 s).

- Channel propagation model: Pathloss, Walfish Ikegami model: $K_1 + K_2 \cdot \log_{10}(d)$ with arbitrary values of K_1 and K_2; shadowing, modelled by means of log normal random variables with zero mean and exponential correlation function. Fast Fading: different ITU channels are simulated (pedestrian, vehicular, indoor, both A and B)

[54].

- Power Control: both fast closed loop power control (rate 200 Hz) and slow outer loop power control are implemented. Each service is characterized by a proper value of the $(E_b/I_0)_{target}$. Different TPC(Transmit Power Control) step sizes can be selected.
- Antennas: either omnidirectional or sectorized antennas with arbitrary (2D) antenna pattern.
- Links: both uplink and downlink BER/BLER are considered to estimate the quality of radio links. This information is obtained through communication with the link-to-network level interface module (see subsection 1.2.2).
- At datalink layer (level 2), Automatic Repeat reQuest (ARQ) is implemented for packet switched services [19].
- Radio Resource Management (RRM) - Call Admission Control: a new user is accepted provided that codes are available, the estimated interference is less than a given threshold and the power required in downlink is available. Other specifications might be the maximum number of RUs occupied per timeslot by all CS users and allocation strategies that subdivide various class of service users (i.e. speech and data users) on different timeslots.
- RRM - Dynamic channel allocation (DCA): based on code-timeslot-frequency allocation. Three carriers with a separation of Δf=1.6 MHz are considered. A scheduler, which combines throughput maximization and fairness guaranteeing, is proposed for multi-service TD-SCDMA system.
- RRM - Handover Control: in TD-SCDMA, only hard-handover is defined. In the simulator, the event 1G for Change of Best Cell in intra-frequency scenario has been implemented [51]; the condition is that "a P-CCPCH RSCP becomes better than the previous best P-CCPCH RSCP". In similar way, the event 2A applies for inter-frequency scenario.
- RRM - intersystem Handover: due to the Common Radio Resource Management (see chapter 5) for heterogeneous networks, calls can be moved from/to TD-SCDMA network from/to an alternative radio access technology.
- Both conversational (speech and video) and interactive/background services are simulated. Every class of traffic has the following characteristics: number of RUs

(number of timeslots, number of codes used in downlink and uplink), number of information bits carried in a transport block.
The mobility characteristics and the type of user activity (i.e. voice call arrival rate, packet distribution statistics, etc.) are defined offline by other simulators, respectively the Mobility simulator and the User Activity simulator; these data are communicated to the network simulator by the Upper Layer simulator.

The TD-SCDMA dynamic network simulator provides an extensive set of radio performance measures (on the contrary, the overall system performance in terms of call block, call drop, system throughput, etc. is evaluated by the Upper Layer simulator):

- System efficiency: in TD-SCDMA network, it's essential to exploit the oscillating switching point between downlink and uplink timeslots in order to fulfill asymmetric traffic request.
- Coverage, capacity and handover maps.
- Radio statistics: the TD-SCDMA network simulator provides statistics of transmit power (both in the NodeB and in the UE), received power, interference, signal-to-interference ratio, etc.

The union of the TD-SCDMA network level simulator with the link-to-network level interface module constitutes a **Lower Layer Simulator** (see figure 1.3), according to the definitions and the terminology introduced in our home-made platform, SHINE, *Simulation platform for Heterogeneous Interworking Networks*, see appendix A for more details. In practice, all aspects related to the access technology adopted, hence related to the physical and data-link layers as well as the local Radio Resource Management, are managed by the lower layer simulator.

1.2.4 Upper Layer simulator

The Upper Layer simulator manages the end-to-end aspects of each connection, no matter the supported access technology at the physical and data-link levels. This means the upper layer simulator design is independent from the specific wireless network implemented in the lower layer simulator, therefore the same upper layer simulator program can be connected to different lower layer simulators (i.e. TD-SCDMA, W-CDMA, WLAN, etc.).

In SHINE (see appendix A), a further step has been done: a unique upper layer simulator may be connected to several lower layer simulators at the same time, in order to reproduce the conditions of a wireless heterogeneous network. Actually, the end-to-end connection control is terminated at the upper layer simulator side, whereas peculiar radio access aspects are reproduced by each different lower layer simulator.

The tasks of the upper layer simulator are mainly concerned with user activity management and its corresponding mobility in the simulated scenario. For these purposes, the upper layer simulator receives as input the results of a proper **User Activity simulator** and a **Mobility simulator** (see figure 1.3); aiming at the reproducibility of the system simulation results, the mobility and the user activity simulators provide their results to the upper layer simulator in offline mode.

The user activity simulator, according to customized call arrival statistics and IP traffic models, generates the profile of the service usage of each user: the instant of call initiation and termination for both voice and data users are defined as well as the instant of generation and the dimension of each IP packet based on FTP download and web browsing application characteristics.

On the other hand, the mobility simulator reproduces, based on the specific characteristics of each traffic class (for example, vehicular voice users as well as static email service users), the movements of each mobile station within a realistic scenario. The movements can be completely randomly generated or forced to follow certain rules (for example, users moving along the streets).

The main tasks of the upper layer simulator are therefore defined:

- communication to the TD-SCDMA lower layer simulator of the starting instant of each new traffic session according to the user activity simulator as well as the time-varying user position within the investigated scenario according to the mobility simulator;

- generation of the bit-flows up(down)loaded by users in each session according to the user activity simulator: each packet transmission shall be simulated by the lower layer simulator, but the transport protocol is transparent to the lower layer simulator.

  The upper layer simulator implements the most important transport level protocols (TCP, UDP, etc.);

- execution of all *Common Radio Resource Management* (CRRM) functions (see chapter 5): the upper layer simulator selects through which technology should each user be connected on the basis of customized rules and the available networks' information; it can also decide to move a connection from a lower layer simulator to another (that is, from a given technology to another) at any time, thus simulating the interworking;

- collection of all simulation results in order to provide application level performance, that is from an end-to-end point of view (i.e. call setup success rate, call drop rate, user throughput, etc.).

1.3 How to convert the TD-SCDMA system simulator to the UTRA FDD option

In the previous section 1.2, the characteristics of the TD-SCDMA system simulator developed during this work are illustrated. In this section, an overview of the required changes to convert this environment to the UTRA FDD system (called W-CDMA in the terminology of ITU) is given.

Let us recall that in Europe, the 3G standard has been initially developed by ETSI (European Telecommunication Standard Institute) under the designation of UMTS (Universal Mobile Telecommunications System). The radio access interface of the UMTS (UTRA) comprises two standards for operation in the FDD and TDD modes. Both interfaces have been accepted by ITU and are designated IMT-DS (Direct Spread) and IMT-TD (Time Division) respectively.

The UMTS standard is being currently defined by Third Generation Partnership Project (3GPP): a joint venture of industry organizations and of several Standards Developing Organizations from Europe (ETSI), US (T1), Japan (ARIB), Korea (TTA), and China (CWTS).

According to the second 3GPP release (called Release 4) the UMTS terrestrial radio access standard includes the following modes:

- UTRA FDD (W-CDMA)
- UTRA TDD_{HCR} (3.84 Mcps, 5 MHz bandwidth, TD-CDMA air interface)
- UTRA TDD_{LCR} (1.28 Mcps, 1.6 MHz bandwidth, TD-SCDMA air interface)

where HCR stands for High Chip Rate and LCR stands for Low Chip Rate (in the first release, called Release'99, TD-SCDMA wasn't yet defined).

These systems share the same Core Network and a common set of features within the UTRAN (UMTS Terrestrial Radio Access Network). On the other side, the Radio Access Technology (RAT) (collecting with this term the radio frame format, channel coding procedures, definition of transport channels and so on) is the distinguishing aspect among them.

	FDD	TDD	
Air Interface	**W-CDMA**	**TD-CDMA**	**TD-SCDMA**
Bandwidth	2 * 5 MHz paired	1 * 5 MHz unpaired	1 * 1.6 MHz unpaired
Frequency re-use	1	1	1 or 3
Handover	soft, softer (interfreq:hard)	hard	hard
Receiver	Rake	Joint Detection/Rake	Joint Detection/Rake
Chip rate	3.84 Mcps	3.84 Mcps	1.28 Mcps
Spreading factor	4-256	1,2,4,8,16	1,2,4,8,16
Power control	fast: every 667 μs	slow: 100 cycles/s	slow: 200 cycles/s
	closed loop	UL:open, DL:closed loop	closed loop
Frame duration	10 ms	10 ms	5ms (2 subframes in 10ms)
Timeslot duration	0.667 μs	0.667 μs	0.675 μs
Time slot/frame	15	15	7

Table 1.1: 3GPP modes: FDD, TDD 3.84 Mcps, TDD 1.28 Mcps

In this thesis, the main studies have been focused on TD-SCDMA because of its compelling characteristics, one above all the possibility to manage asymmetric traffic thanks to its TDD nature. Nevertheless, here a comparison with W-CDMA is given and the main changes are described.

In table 1.1, a summary of the main differences between TD-SCDMA and W-CDMA (for the sake of completeness, also the characteristics of TD-CDMA, that is the TDD high chip rate option, are included).

It's evident that the main differences between TD-SCDMA and W-CDMA are in the physical signal (i.e. bandwidth, frame structure, etc.) and in the physical procedures (power control, receiver characteristics, handover type, etc), but also Radio Resource Management procedures like Dynamic Channel Allocation have a different (reduced) impact in the W-CDMA simulator.

The simulators (see figure 1.3) that are impacted by the implementation of W-CDMA are the TD-SCDMA link level simulator and the TD-SCDMA network level simulator. On the contrary, the link-to-network level interface module is not impacted, since the "instant values" approach (see chapter 3) has a quite general validity. Also the current upper layer simulator doesn't need any change to communicate with a W-CDMA network simulator.

The main modifications in the existing TD-SCDMA link and network simulators are here described.

W-CDMA is not TDMA

First of all, the Time Division Multiple Access characteristic of the TD-SCDMA system implies that in the time domain different connections can be multiplied (for example, two users can use the same spreading factor 16 code, but in different timeslots). This access technique is not present in the W-CDMA mode, therefore a W-CDMA simulator can simplify the structure of the so called *Resource Units* (RU), by eliminating one dimension:

$$RU_{TD-SCDMA}[code, timeslot] \longrightarrow RU_{W-CDMA}[code]$$

Both the W-CDMA link and the W-CDMA network level simulators, shall assign a code for each connection, and this code is used continuously by the same user, until the call release (or a RAB modification, still not implemented in the current TD-SCDMA network simulator).

Note that in a further step of the development of the W-CDMA system simulator to comply to 3GPP Rel5 specifications, the introduction of HSDPA (High Speed Downlink Packet Access) channels will require a structure in $[code, subframe]$, since in every HSDPA sub-frame (2 ms), the HS-DSCH channels (1 to 15, with fixed spreading factor 16) can be flexibly assigned by the MAC-hs scheduler to different Rel5 users.

W-CDMA is FDD

Since W-CDMA uses paired bands, the resources are doubled for downlink and uplink usage. This mainly impacts the current TD-SCDMA network simulator, because the TD-SCDMA link level simulator is already split in two part, one for downlink simulation, the other one for uplink simulation.

Frame organization

Given the differences of timeslot and frame duration specified in table 1.1, the major impacts in the W-CDMA simulators are that it is not anymore requires to cycle each frame two times to reproduce the behavior of the two sub-frames of 5 ms, typical of TD-SCDMA.

On the other hand, the FDD characteristic involves a doubled effort for the calculation of the transmitted powers and the received signal-to-interference plus noise ratio (SINR) of each connection, because these calculations shall be repeated each timeslot both in downlink and in uplink (on the contrary, in the TD-SCDMA system, each timeslot is exclusively used in downlink or in uplink).

Power Control

Both the TD-SCDMA and the W-CDMA modes define a closed loop power control. However, the mechanism of W-CDMA is slightly simpler to be implemented, because in each timeslot, both in downlink and in uplink, the measured SINR is immediately compared to the $SINR_{target}$ and a consequent Transmit Power Control (TPC) command is sent to the transmitting entity (UE or NodeB) in order to control its transmitted power level and therefore satisfy the quality criterion.

On the contrary, in TD-SCDMA, the receiving entity measures the SINR in each active timeslot (i.e. the timeslots in which there are data for that user), and waits the next switching point (i.e. the point in which the direction of transmission change from/to downlink/uplink) to evaluate the average value of the accumulated SINR measurements and finally send the related TPC command.

Receiver

Because of the higher number of active codes at the same time in W-CDMA compared to TD-SCDMA, the joint detection is not feasible in W-CDMA. The reason of this large difference in the number of codes is that W-CDMA user codes at higher spreading factor: for example, for a speech AMR 12.2 Kbps connection, in W-CDMA a code with spreading factor equal 128 in downlink is used; on the contrary, TD-SCDMA for the AMR call uses 2 parallel codes at spreading factor 16 in one timeslot every 5 ms sub-frame.

Handover

Many differences distinguish W-CDMA and TD-SCDMA regarding the intrafrequency handover procedure (here in this work, the interfrequency handover is neglected): since TD-SCDMA implements the hard handover only, whereas W-CDMA uses soft and softer handover, different measurements shall be defined in the W-CDMA network simulator.

Currently, in the TD-SCDMA network simulator, the event 1G for Change of Best Cell (TDD) has been implemented [51]; the condition is that a P-CCPCH RSCP becomes better than the previous best P-CCPCH RSCP (see figure 1.7); this measurement shall trigger the replacement of the previously best cell with the current evaluated cell.

On the contrary the W-CDMA simulator shall implement at least the following three different events (see figure 1.8):

- event 1A: a Primary CPICH enters the reporting range; this measurement shall trigger a cell addition to the current cell active set.

- event 1B: a primary CPICH leaves the reporting range; this measurement shall trigger a cell deletion form the current cell active set.

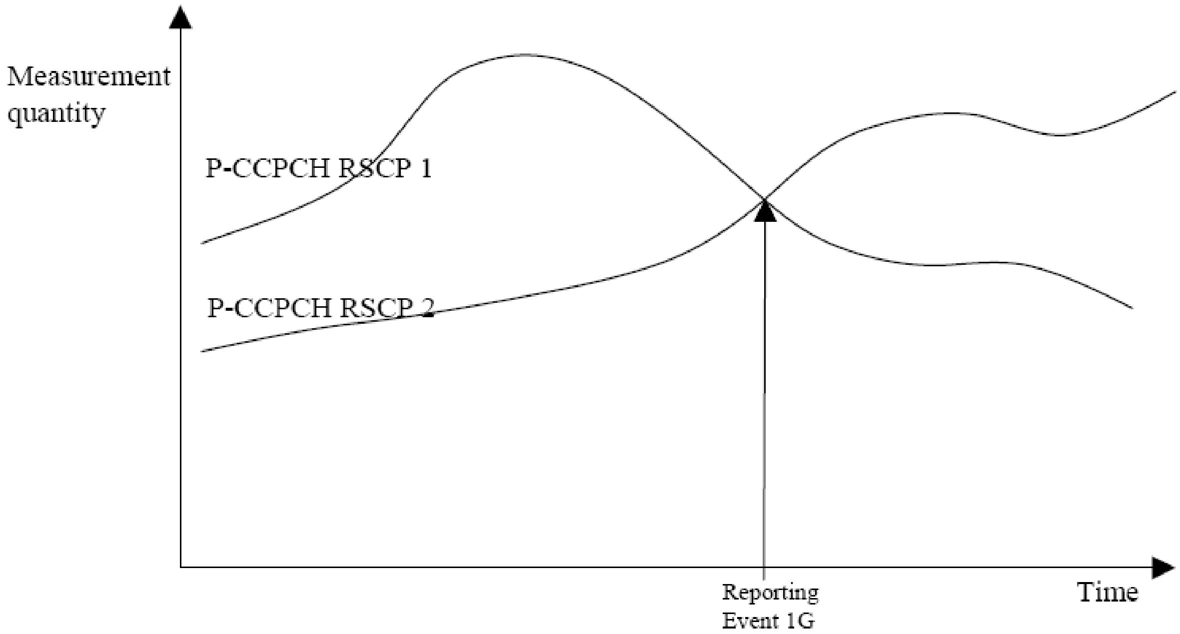


Figure 1.7: event 1G for TD-SCDMA: a P-CCPCH RSCP becomes better than the previous best P-CCPCH RSCP

- event 1C: a primary CPICH that is not included in the active set becomes better than a primary CPICH that is in the active set; this measurement shall trigger a cell replacement within the current active set.

Apart from the different measurement triggers, two new issues shall be solved by the W-CDMA network simulator in order to correctly implement the FDD soft/softer handover: first of all, the data structure of each radio connection in the W-CDMA simulator shall implement the possibility to manage more than one radio link at the same time (typically a maximum of 3 radio links in active set). Whereas the uplink transmit power in the mobile is obviously the same for all radio links, in downlink each cell shall transmit with a proper value defined by the closed loop power control. In order to avoid drift effects among the different active NodeBs, the implementation of a Downlink Power Balancing algorithm should be considered.

Secondly, in the TD-SCDMA network simulator, since the P-CCPCH is transmitted only on the timeslot 0 of each subframe without other dedicated physical channels, the only appropriate quantity used for handover measurement is the Received Signal Code Power (RSCP) from the P-CCPCH. On the other hand, in the W-CDMA mode, the reference for handover measurements is the Common Pilot Channel (CPICH), which is transmitted in all timeslots, and multiplied with all the other common and dedicated

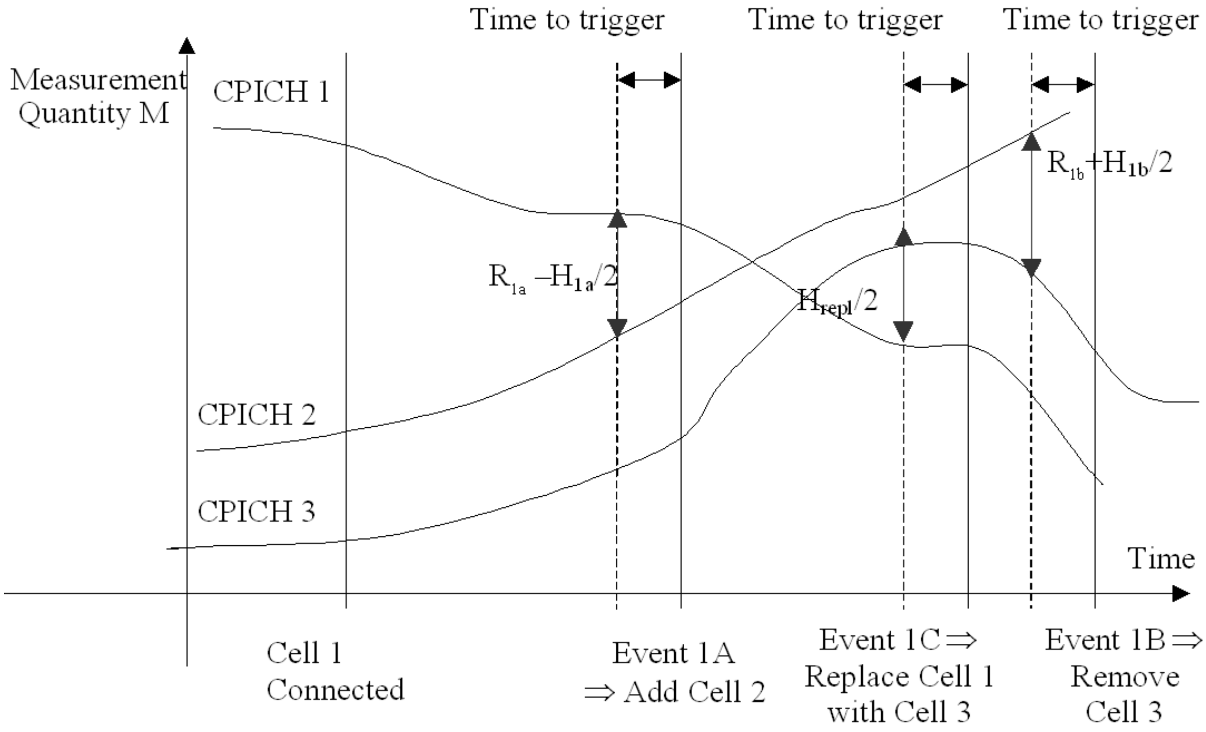


Figure 1.8: W-CDMA basic handover algorithm

physical channels, therefore for handover control it's appropriate to measure also the downlink E_c/N_0, that is the received energy per chip from CPICH divided by the power density in the band.

Chapter 2

Performance of TD-SCDMA in mixed CS/PS traffic scenarios

In this chapter, the performance of a time-division synchronous code-division multiple access (TD-SCDMA) system is analyzed. In particular, we investigate the impact of packet switched applications (for instance web browsing sessions) on the overall performance of the network. Here, we quantify the degradation of the voice users quality in the presence of packet data services and viceversa. Finally, we investigate the impact of some physical layer parameters for the power control algorithm (PCA) on the overall quality of service and we show that these parameters should be carefully chosen in order to balance the quality of voice and data users. The analysis of the influence of physical layer on network performance is carried out through the method based on an ”instant value” interface depicted in chapter 3.

The numerical results have been obtained through the system simulator developed during this research which takes both link and network level issues into account; to the author’s knowledge only few papers in literature provide a thorough investigation of TD-SCDMA networks since the majority of them deals with either link or network level analysis, separately (see for instance [15, 16]).

The main characteristics of the used TD-SCDMA system simulator are described in chapter 1.

In the next sections 2.1 and 2.2, we present in more detail two of the main algorithms directly impacting the performance of TD-SCDMA in mixed CS/PS traffic scenarios, respectively the packet scheduler and the power control. Afterwards in section 2.3, the simulated scenario is described and in section 2.4 the merit figures for performance evaluation are defined; finally, in section 2.5 the results obtained with the TD-SCDMA system simulator are discussed.

2.1 Packet scheduler

The main goal of a scheduling algorithm should be the achievement of as many satisfied data packet users as possible at a given system load. It turns out therefore necessary to define a criterion based on which a user is considered satisfied; in fact, various criteria of evaluation carry to different choices for the allocation algorithms [19].

For a *minimum delay* criterium a first-in-first-out (FIFO) scheduling strategy would be a good choice: the data of all users are appended to a single queue in the order they are requested and they are read out on a FIFO basis. An other choice is a *throughput based* criterium: a user is defined to be satisfied if the average data rate during the entire connection, the active session throughput (AST), does not remain under a certain threshold.

In this work, we adopted the second criterium: in particular, ETSI [18] specifies that the AST threshold value should be 10% of the nominal bit rate, Bit_{RX}. AST represents a performance figure reported for every single user, and is defined as the ratio between the total number of bits at level of application received during the whole session (block retransmissions and overhead due to lower protocol layers are not considered) and the duration of the session, $T_{session}$, excluded the time intervals in which no information from upper layers request to be transmitted, T_{inact} (that is the transmission queues do not contain packets from that user).

$$AST = \frac{Bit_{RX}}{T_{session} - T_{inact}} \geq 10\% \cdot Br_{nom} \tag{2.1}$$

In this work we decided that, for Non-Real Time (NRT) packet data services, uplink and downlink shared channels (respectively, USCH and DSCH [48]) can be used to allow efficient allocations for a short period of time: these transport channels exist only in TDD mode. The MAC-c/sh packet scheduler, localized in the Radio Network Controller

(RNC)[1], redistributes dynamically the RUs between all active users, that is the users with data in the buffers, waiting to be transmitted.

We considered three Quality of Service packet bearers, with different requests of maximum rate, at 64 Kbps, 144 Kbps and 384 Kbps. In downlink, for every cell, three buffers are used (one for every class of service): they hold the packets addressed to mobile terminals connected to that cell.

The packet scheduler algorithm proposed here considers two level of priority to establish a data packet transmission. First of all, transmission requests are ordered based on the service class to which they belong; level of priority, from highest to lowest, is following: 384 kbps, 144 kbps, 64 kbps. Second level of priority orders the requests of users who belong to the same class of service: a fair round robin (RR) algorithm is used.

When a data packet transmission request is finally admitted to air interface by the scheduler, the packet is fragmented into transport blocks: the number of TBs necessary for the complete transmission of the considered packet is obtained by dividing the number of bits at level of application contained in the packet by the number of information bits carried by a TB (every bearer service has its own characteristics, in terms of number of bits, timeslots and codes per TB [52, 53]). During a packet transmission, if the link-to-system level interface (see chapter 3) evaluates that an error has been occurred in the reception of a transport block, only that TB is retransmitted.

2.2 Power control

Power control is the mechanism which takes care of maintaining the received signal-to-interference ratio (or the received power level) at a constant value. In modern cellular network this is commonly carried out in both uplink and downlink. In our TD-SCDMA system simulator three different kinds of the power control are implemented. [21]:

- *Open-loop power control:* owing to the correlation among the average path loss of downlink and uplink, the user equipment (UE) can estimate the initial power needed in uplink and downlink based on the path loss calculations in the downlink direction. This mechanism is commonly used to set the initial value of received value of signal-to-interference ratio.

- *Closed-loop power control:* this PCA uses the feedback information from the opposite end of the radio link. This allows the considered terminal (UE in uplink and base

[1] On the other hand, the MAC-hs scheduler using the latest Rel5 HSDPA transport channels is located in the NodeB. This function is not yet implemented in our system simulator.

station in downlink) to adjust the value of the transmitted power. Let us consider for instance the uplink power control, user equipment transmits using a given value of power, the base station measures the value of received power and compared it with its target value of signal-to-interference ratio, if this value is smaller than the threshold, the base feeds back, through the TPC control bits, this information and the UE increases its value of transmitted power. As few bits are carried by TPC, the step between the old and the new value of transmitted power is quantized (PC step).

If the speed of the power update is sufficiently high, terminal (or base) can compensate for the fast fading contributions. In W-CDMA network, as terminal transmits in each timeslot, the value of transmitted power is changed with a frequency of 1500 Hz (15 slot/10 ms). In case of TD-SCDMA, owing to time division multiple access nature, transmitted power is updated one time each sub-frame, so the frequency is 200 Hz.

- *Outer loop power control:* The aim of outer-loop power control is that of maintaining the quality of the communication at the level defined by the quality requirements of the bearer service. This is carried out by changing the target value of the received E_b/I_0; this is necessary when propagation conditions change, i.e. a change in the mobile speed, and therefore the previous value of received E_b/I_0 does not guarantee an acceptable value of bit error rate.

2.3 Scenario and propagation environment

As far as channel model is considered, in this chapter we assume that the received power (in uplink and downlink) is calculated according to the following expression:

$$P_r[dBW] = P_t[dBW] - K_1 - K_2 \cdot log_{10}(d) + S[dB] + F[dB] \tag{2.2}$$

where P_t represents the power transmitted by the mobile user (in uplink) or the base station (downlink), pathloss is modelled by means of the coefficients K_1 and K_2 (expressed in dB), and d is the distance between transmitter and receiver. Several other pathloss models can be considered: Walfish Ikegami, taking both line-of-sight (LOS) and non-LOS conditions into account; $K_1 + K_2 \cdot \log_{10}(d)$ with arbitrary values of K_1 and arbitrary coverage maps.
S stands for log-normal Shadowing, modelled by means of Gaussian random variable (expressed in dB) with zero mean and exponential correlation function. Finally fast

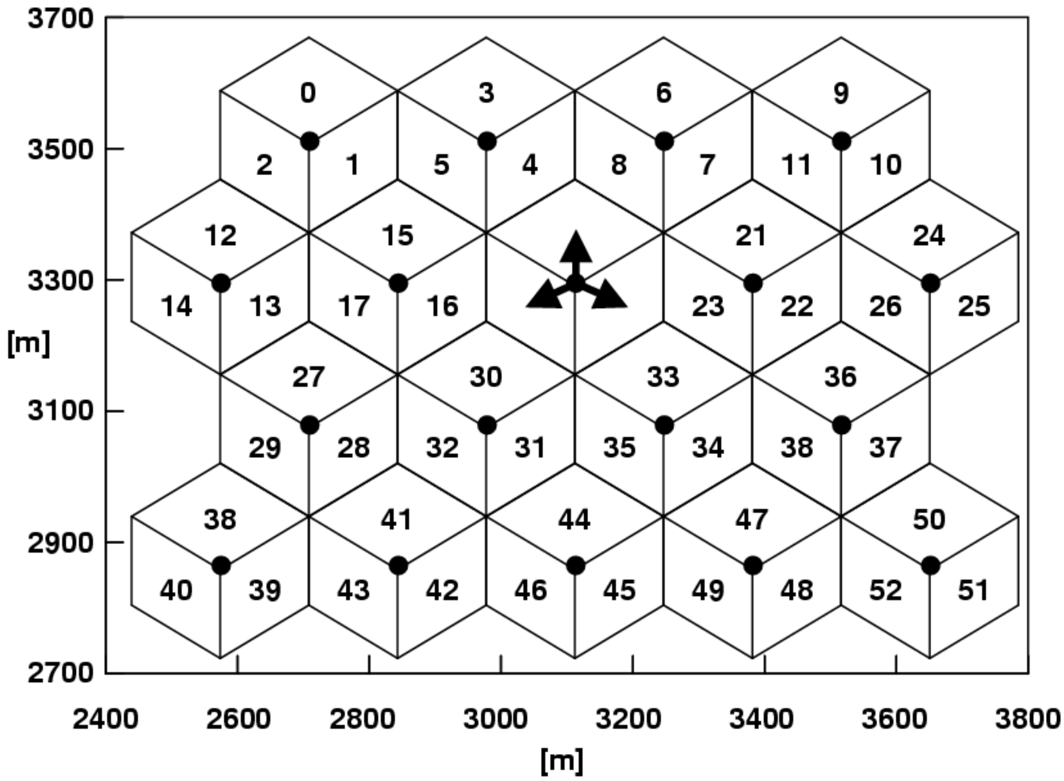


Figure 2.1: The considered network scenario

fading is modelled using the time-correlated random variable F: different ITU channels are simulated (pedestrian,vehicular,indoor, both A and B).

The simulated scenario considered in this chapter is depicted in figure 2.1. Here a regular lay-out with 18 sites is designed. However, the simulation tool is quite general and allows the evaluation of a larger number of sites located in arbitrary positions.

As far as antenna pattern is considered, a two-dimensional antenna pattern is considered. Both omnidirectional and sectorized antennas can be considered by the simulation tool. Here, we assume that each site uses three sectorized antennas. The total number of sectors considered in this chapter is 54 (see figure 2.1).

The total bandwidth assigned to each site is 5 MHz; we have assumed that every sector uses a different 1.6 MHz carrier (1/3 frequency reuse).

Main simulation parameters are summarized in table 2.1.

Parameter	Default value	Description
K_1	15.3 dB	Pathloss model
K_2	37.6 dB	Pathloss model
σ	5 dB	dB spread of the log normal shadowing process
d_o	20 m	Correlation distance of the log normal shadowing process
N_{o_DL}	-169 dBm/Hz	downlink noise power spectral density
N_{o_UL}	-174 dBm/Hz	uplink noise power spectral density
α	0.2	downlink orthogonality factor
β	50%	MUD efficiency
BER_{out}	10^{-2}	Threshold on averaged BER for outage
BER_{drop}	$2 \cdot 10^{-2}$	Threshold on averaged BER for the evaluation of the dropout

Table 2.1: System parameters fixed in the numerical results

Traffic characteristics

Both voice and data bearer services are considered in this chapter in order to achieve the main target of this study that is to quantify the degradation of the voice users quality in the presence of packet data services and viceversa.
Every class of traffic is characterized by the following parameters: the number of Resource units, classified in number of time slots required, number of codes used in both downlink and uplink, the number of information bits carried by the transport block [52][53].

The main characteristics of traffic generated in the simulations of this chapter are listed as follows for every class of service:

- Voice services (CS): call arrival and departure processes are Poisson distributed; average call duration is 120 s.
  T_{CS} denotes the voice offered traffic per site (in Erlang).

- Data users (PS): the data packet application we have simulated is Web browsing. Three layered stochastic processes are considered: the session arrival process, the packet call arrival process and the packet arrival process. Session arrival process is Poisson distributed. Packet size is Pareto distributed; all the parameters have been fixed according to [18].
  Mean traffic parameters are specified in table 2.2, where N_{pc} is the average number of packet calls per session, RT_{pc} is the average reading time between consecutive packet calls, N_p is the average number of packets per packet call, IAT is the average packet inter-arrival time. Number of packet calls, reading time, number of

QoS service	N_{pc}	RT_{pc}	N_p	IAT
64 Kbps DL	5	120 s	25	60 ms
144 Kbps DL	5	120 s	25	20 ms
384 Kbps DL	5	120 s	25	10 ms
64 Kbps UL	5	200 s	12	10 ms
144 Kbps UL	5	180 s	12	10 ms
384 Kbps UL	5	180 s	12	10 ms

Table 2.2: Packet switched session parameters for web browsing services

packets per packet call and inter-arrival time are all modelled by means geometric distributions.

Because of bursty and asymmetric nature of data packet sources, we do not characterize the data traffic by means of the average number of data user in the system. To measure the data packet traffic per site (using separate metrics for downlink and uplink) we have defined T_{PS}, which represents the ratio between the total number of bit at application level received (if we consider the downlink) or transmitted (uplink) by all users connected to the site and the total time of simulation (i.e. 100 minutes). Block retransmissions and overhead due to lower protocol layer are not considered.

2.4 Performance Metrics

Several metrics can be evaluated by the simulation tool for circuit-switched applications:

- *Blocking rate (T_b):* defined as the ratio between the number of voice connections blocked by the call admission control algorithm and the total number of voice connections generated.

- *Dropping rate (T_d):* defined as the ratio between the number of voice connections dropped owing to quality problems and the total number of terminated connections. The simulator decides to drop a call using a "leaky bucket" algorithm, which compares the averaged measured BER with a given threshold BER_{drop}. If the BER is larger than the threshold the call gains two points, otherwise it looses one point. The call is dropped when the counter reaches a given value.

- *Outage rate (T_{out}):* an event occurs when the averaged value of the BER exceed a threshold - BER_{out}.

- Satisfaction rate (T_{sat}): a user is said to be satisfied if his call is neither blocked nor dropped, and during the call the outage events have a duration which is smaller than 10% of the total duration of the call.

In case of packet switched applications, the simulation tool evaluates the following performance metric:

- *Transport block error rate (BLER):* defined as the rate of transport blocks erroneously detected by the receiver owing to interference.
- *Active session throughput (AST)*: defined by (2.1).
- *Distribution of packet arrival delays*: it gives the distribution in terms of probability density function (pdf) and cumulative density function (cdf) of the delays in the delivery of the packets.

2.5 Simulation results

The results of three different simulation sessions are here presented. Firstly, in subsection 2.5.1 we show the results in a basic scenario with only speech users; afterwards, in subsection 2.5.2, we analyze the performance in a mixed traffic scenario with both voice and data user, when varying the network load; finally, in subsection 2.5.3, we discuss how to balance the performance of voice and data users through tuning of physical layer parameters.

2.5.1 Circuit switched services

In the first simulation session, we have considered three values of offered traffic for voice users (using a AMR codec at bit rate of 12.2 Kbps): $T_{CS} = 21.6$, 36 and 43,2 Erlang per site; we repeated each simulation for different cell radius values ranging from 150 m to 3600 m.

Figure 2.2 shows that for a fixed value of offered traffic, when the cell radius is smaller than 450 meters, the performance in terms of voice satisfaction rate T_{sat} remains practically constant; for radius up to 900 meters it starts to decreases, and it is remarkably reduced for larger values. The behavior can be explained by remembering that, for large radius, the system is limited by coverage and the limiting factor is the maximum power transmitted by the user's terminals.

On the other hand, fixed the cell radius, due to capacity limitations, the less loaded configuration ($T_{CS} = 21.6$ Erlang/site) provides the best performance.

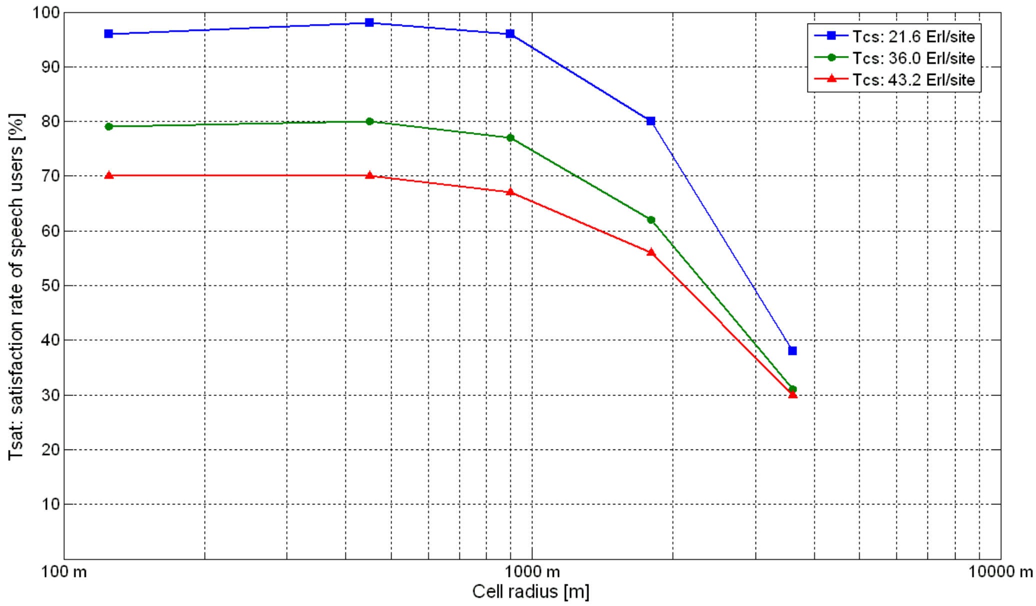


Figure 2.2: Voice satisfaction rate (T_{sat}) vs. cell radius

These results are in agreement with those obtained in literature for W-CDMA [20, 21]. The clear dependence of T_{sat} from the network load could be minimized with the introduction in the simulation tool of smart antennas which would reduce the levels of interference (for future implementation in the TD-SCDMA system simulator).

The geographical distribution of unsatisfied users for the three values of network load, at $R = 450m$ is shown in figure 2.3: the pixels in green represent the areas in which speech users are satisfied, the pixels in yellow and red represent different ascending levels of users unsatisfaction, whereas the black squares are located in the positions of the NodeBs. It's clear from the figure that, for higher network loads, the users at border cell suffer with

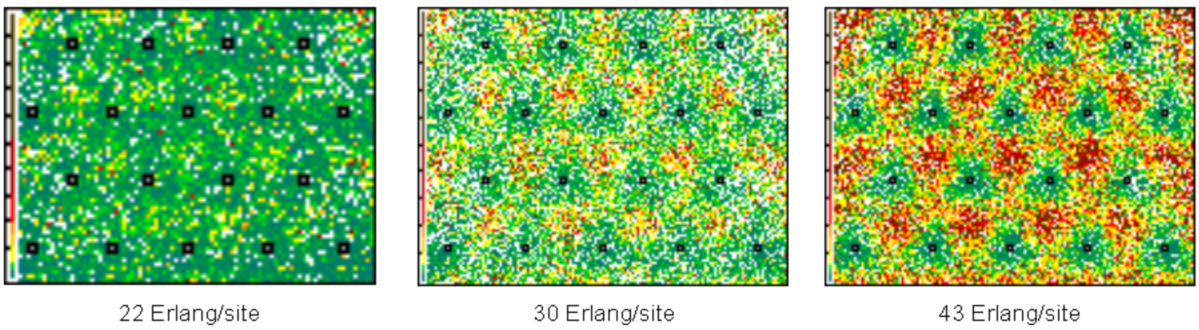


Figure 2.3: Areas of user satisfaction in the scenario

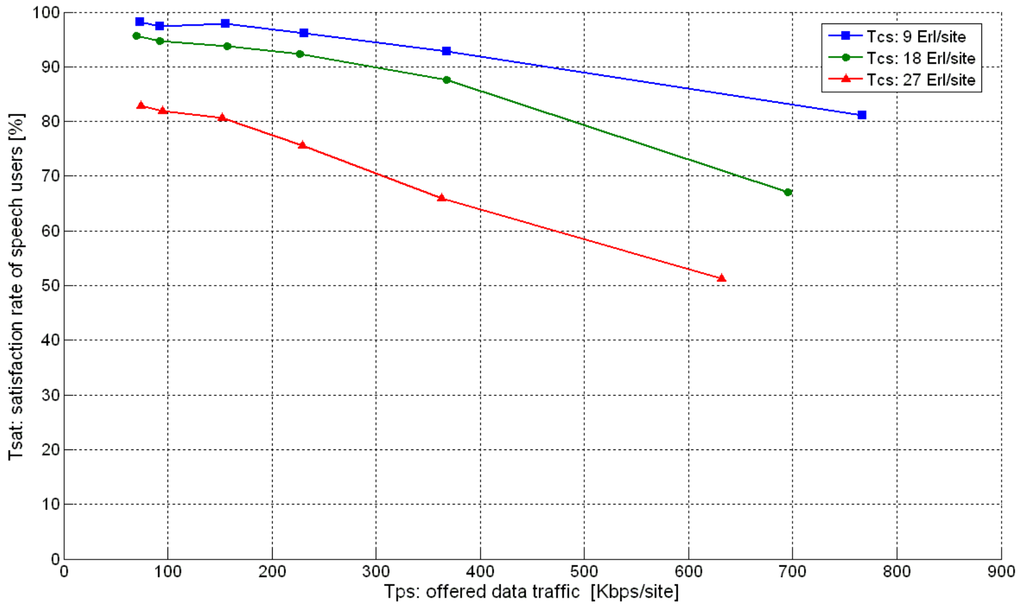


Figure 2.4: Voice satisfaction rate (T_{sat}) vs. downlink data packet traffic (T_{PS}), for different values of voice offered traffic (T_{CS})

larger probability outage situations or call drops.

2.5.2 Packet and circuit switched services

In this second simulation session, mixed scenarios with both circuit and packet switched services are considered: voice at 12.2 Kbps, and data packet services at 64 Kbps, 144 Kbps and 384 Kbps; in each simulation the offered traffic of data users T_{PS} is varied.

For what concerns the performance in terms of active session throughput for packet switched classes, we focus on downlink; however, for web browsing sessions, our program simulates also uplink traffic: in this scenario the asymmetry between uplink/downlink traffic is about 1:3. On the other hand, the metric for voice users (T_{sat}) considers both uplink and downlink.

Figure 2.4 shows the voice satisfaction rate versus the offered downlink data packet traffic: T_{sat} decrease with the increasing amount of downlink data traffic due to the increasing level of interference. In this simulation session, to augment the downlink data packet traffic, the average arrival frequency of data sessions has been increased; this have an impact also on the data traffic generated in uplink: the degradation of the quality of voice users is due to the combined effect of data packet traffic in both downlink and uplink.

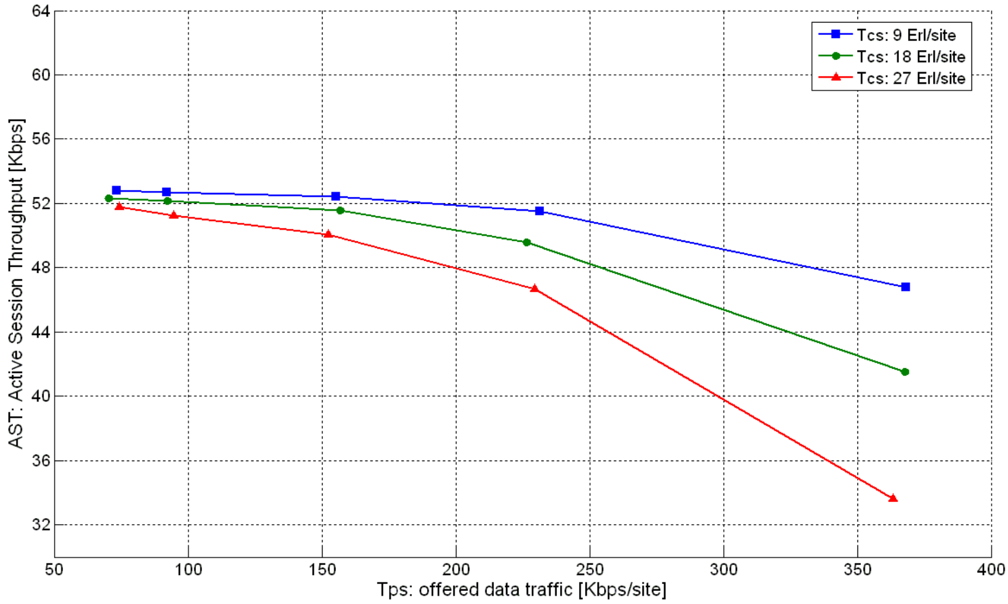


Figure 2.5: Downlink active session throughput (AST) vs. downlink data packet traffic (T_{PS}), for different values of voice offered traffic (T_{CS})

Figure 2.5 shows the downlink active session throughput for data packet users with 64 Kbps QoS versus the downlink data packet traffic; AST[2], that is the average throughput per session, decreases firstly with the increasing amount of downlink data traffic because of higher latency time in buffers and secondly it decreases when a higher number of voice users is served in the network.

2.5.3 CS/PS performance enhancements

In this final simulation session, we analyze a more specific optimization problem. In fact, here we investigate the parameterization of the fast power control algorithm (PCA) of TD-SCDMA (see section 2.2): we evaluate the impact of the target signal-to-interference and noise ratio (E_b/I_o) of the PCA on the combined performance of voice and data users.

In all the next simulations, we have considered a value of T_{CS} of 18 Erlang/site and a value of T_{PS} of 161 Kbs/site in the downlink and 71 kbs/site in the uplink.

The default values of the target (E_b/I_0) for voice and data users, in uplink and downlink, are shown in table 2.3. Only one target (E_b/I_0) has been changed in each of the

[2]The maximum value of AST for 64 Kbps bearer service should be 64 Kbps, but because of fragmentation of packets from the application level in the transport blocks of fixed dimension, there is a reduction on maximum AST.

Service	$(E_b/I_0)_{DL}[dB]$	$(E_b/I_0)_{UL}[dB]$
Voce (CS)	5	7
Data (PS)	-0.5	-1.5

Table 2.3: Default values for (E_b/I_0)

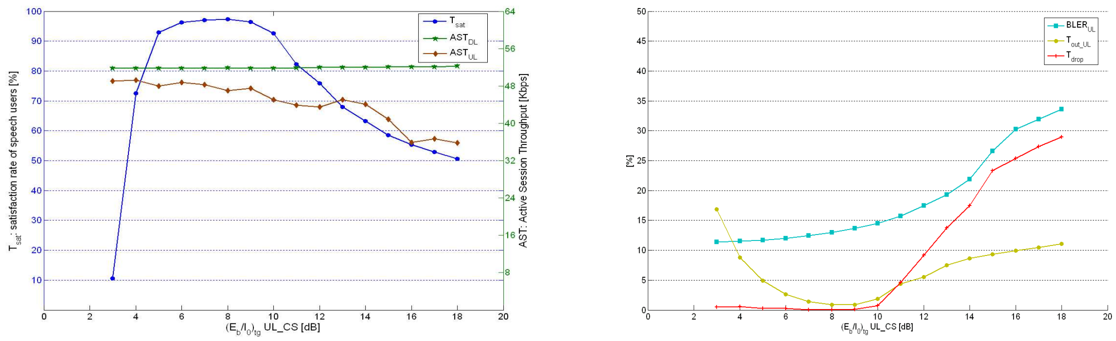


Figure 2.6: Network Performance as a function of $(E_b/I_o)_{UL-CS}$

following figures, the remaining parameters were fixed to the default values indicated in table 2.3.

In this session, we've disabled the outer-loop power control feature, so the target value of (E_b/I_0) is not changed dynamically during the simulation of each connection.

Impact of $(E_b/I_0)_{UL-CS}$ and $(E_b/I_0)_{DL-CS}$

We start the discussion by considering the impact of the target values of (E_b/I_0) of voice users on the overall performance of the network.

Figure 2.6 shows all the main performance metrics considered in this work (see section 2.4). In this first scenario, we have considered *AST* in uplink and downlink, BLER of data connection in uplink, T_{sat}, T_d and T_{out} in the uplink as a function of $(E_b/I_0)_{UL-CS}$.

Let us consider the curve corresponding to the satisfaction of voice users T_{sat}: as expected for small values of $(E_b/I_0)_{UL-CS}$ the quality of service for voice users is very low, since the associated average BER values are rather high. Let us highlight that this kind of analysis is allowed thanks to the tight integration between link and network level simulators (see chapter 3).
If we increase $(E_b/I_0)_{UL-CS}$, T_{sat} increases and it reaches a maximum around 5-9 dB; the use of larger values of $(E_b/I_0)_{UL-CS}$ forces the PCA to increase unnecessarily the UE transmitted power with a consequent degradation in terms of T_{sat}. In fact, the distribution

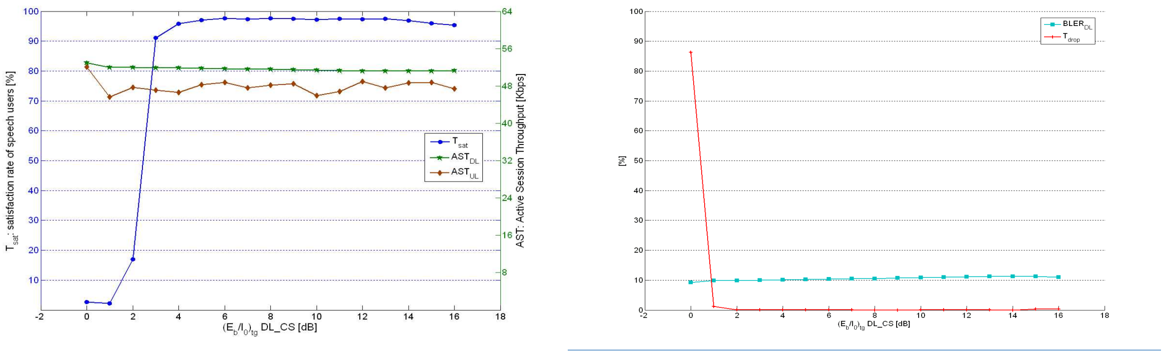


Figure 2.7: Network Performance as a function of $(E_b/I_o)_{DL-CS}$

of the UE transmitted power (which has not been shown here for the sake of conciseness) shows that many terminals are transmitting the maximum allowed power in order to achieve the target $(E_b/I_0)_{UL-CS}$.
Note also that both T_{out} and T_d increases for large values of $(E_b/I_0)_{UL-CS}$.

Now, if we consider the quality of packet data users in this simulation, we observe that AST in uplink is almost constant for small values of $(E_b/I_0)_{UL-CS}$, with a value around 48 kb/s, and decreases for large values of $(E_b/I_0)_{UL-CS}$, as BLER in uplink of data connections increases because of the high level of interference introduced by the voice connections. As expected AST in downlink is not influenced by $(E_b/I_0)_{UL-CS}$.

Figure 2.7 shows the same performance metrics of figure 2.6 as a function of $(E_b/I_0)_{DL-CS}$. Here we can observe that T_{sat} does not have the same trend: this depends on the fact that in the downlink the intercell interference does not play the same role owing to the multi-user detection, here considered through the parameter β. Large values of $(E_b/I_0)_{DL-CS}$ do not increase significantly the level of interference in the network. However, it can be expected that for a higher network load, a too high target for $(E_b/I_0)_{DL-CS}$ will impact significantly the system capacity. As far as the performance of packet-switched services is concerned, here the curves corresponding to AST in uplink and downlink are almost constant.

Comparing the results of figures 2.6 and 2.7, we can observe that the impact of the target value of (E_b/I_0) is much more emphasized in the uplink owing to the strongest contribution of the uplink intercell interference.

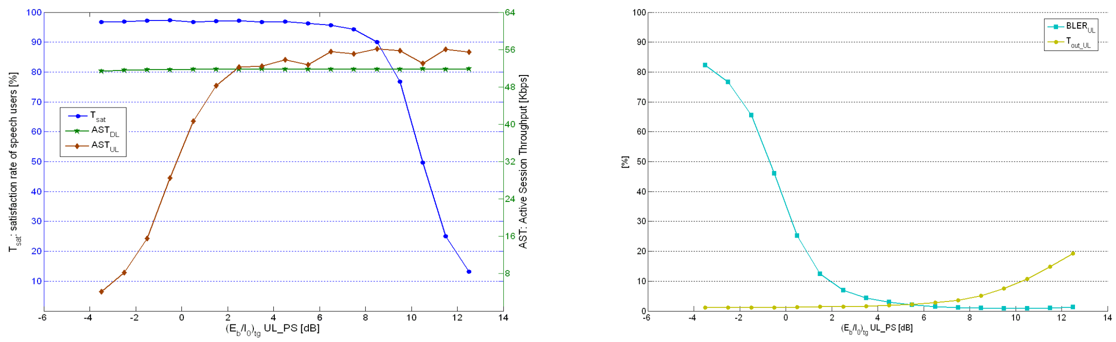


Figure 2.8: Network Performance as a function of $(E_b/I_o)_{UL-PS}$

Impact of $(E_b/I_0)_{UL-PS}$ and $(E_b/I_0)_{DL-PS}$

In this part, we evaluate the performance of the network by varying the target values of $(E_b/I_0)_{UL-PS}$ and $(E_b/I_0)_{DL-PS}$.

Figure 2.8 shows the main performance metrics as a function of $(E_b/I_0)_{UL-PS}$. Small values of $(E_b/I_0)_{UL-PS}$ provide large values of T_{sat}; however, T_{sat} remains almost constant until a value of $(E_b/I_0)_{UL-PS}$ of 8 dB. This behavior can be explained by looking at the curve corresponding to T_{out} in uplink: this value is constant for small values of $(E_b/I_0)_{UL-PS}$ but becomes significant for $(E_b/I_0)_{UL-PS} > 4dB$ because of the increase of interference introduced by data users during uplink transmissions.

Concerning the performance of data services, small values of $(E_b/I_0)_{UL-PS}$ cause a large number of re-transmissions (see the curve corresponding to the uplink BLER) with a consequent reduction in terms of AST in uplink. Larger values of $(E_b/I_0)_{UL-PS}$ increase the uplink throughput until a maximum of about 56 kb/s. Again AST in downlink is not influenced by the variation of $(E_b/I_0)_{UL-PS}$.

Finally, figure 2.9 shows the main performance metrics as a function of $(E_b/I_0)_{DL-PS}$. Here, AST in downlink reaches a value of about 58 kb/s for values of $(E_b/I_0)_{DL-PS}$ larger than 2 dB; this values corresponds to a BLER of about 0.01. As expected T_{sat} decreases when $(E_b/I_0)_{DL-PS}$ is configured to an unnecessary high value; however, it maintains an acceptable value for $(E_b/I_0)_{DL-PS}$ smaller than 4 dB.

To summarize: in this chapter we have evaluated the performance of a TD-SCDMA cellular network in the presence of both circuit and packet switched services. The analysis has been carried out using our home-made simulation environment in which both link and network level issues are taken into account. Results show that the presence of data services

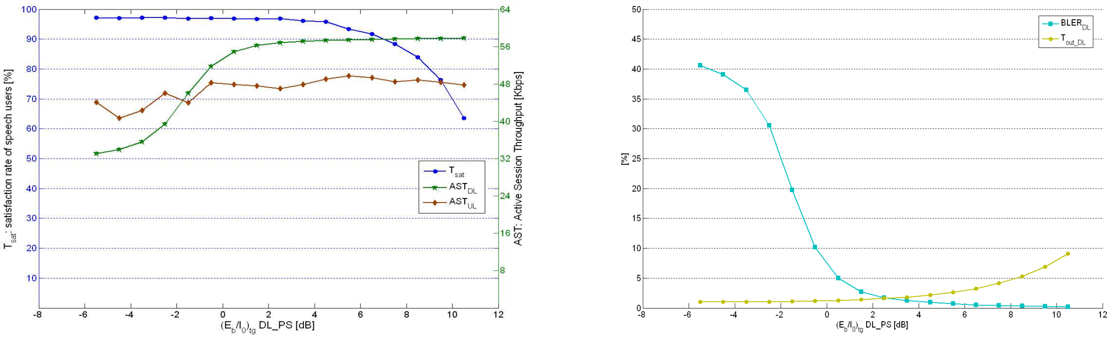


Figure 2.9: Network Performance as a function of $(E_b/I_o)_{DL-PS}$

can reduce the quality perceived by voice users and viceversa. Moreover, the target values of (E_b/I_o) in the Power Control Alogorithm play a crucial role in the balancing of the performance of voice and data services. Owing to the latter considerations, the different values of target (E_b/I_o) should be carefully chosen: the optimum value can be found as a trade-off between the obtainable quality of packet and circuit-switched applications.

Chapter 3

Link Level aspects modelling in the simulation of packet switched wireless networks

The investigation of both link and network level aspects is fundamental when a thorough analysis of mobile radio systems is required. The conventional approach, based on a complete separation between the two levels, adopted in the past for the performance evaluation of second generation systems, is not enough accurate for third generation ones, characterized by the presence of both voice and data services as well as circuit and packet switched applications.

Here we describe the methodology we followed on how to interface link and network level tools; this approach has been adopted in the TD-SCDMA system simulator developed during this thesis and whose results are described in the numerical sections of chapters 2 and 5. Nevertheless, the proposed method is quite general and can be adopted for arbitrary wireless networks with a time slotted physical frame structure and packed-switched applications.

3.1 Wireless system simulations

Third generation systems such as UMTS [21], that will make available high data rate transmission are expected to facilitate the diffusion of a large class of new services that should increase the number of subscribers of wireless applications. Moreover, the presence of services characterized by different quality requirements, are making the design and the management of such networks very challenging.

In the past, system performance evaluation was carried out by separating the aspects related to the radio transmission chain, i.e. coding, modulation, fading mitigation, etc., from those related to the multiple access, radio resource management (RRM) and so on. This division was possible because second generation systems (i.e. GSM [22]) were originally designed to provide mainly voice services, i.e. at constant bit rate.

With the advent of third generation systems, the focus moves from voice to data services and from circuit to packed switched networks. This latter aspect requires the use of performance evaluation tools characterized by the ability to follow all the rapid changes (power control, packet scheduling,..) of the network. The design of such a platform is extremely challenging as link level aspects cannot be separated by those of network level and viceversa [17, 23, 24]. As a matter of fact, since the Call Processing part of UMTS takes decisions every 10 ms (duration of a radio frame) or 2 ms (duration of a sub-frame when supporting HSDPA physical channels in FDD mode), and since fast power control at rate of 1500 Hz shall be implemented and bit rate shall be adapted to actual radio frequency condition, a different design of network level simulators is required.

In the next sections, we address the problem of how to combine link and network level analysis in order to realize a thorough simulation tool able to investigate a modern mobile radio network from a system ("system" is defined in this chapter as the union of link and network level issues) point of view. Here, we show that the investigation of both link and network level issues requires the design of a suitable interface integrating link and network level parts and a careful definition of the parameters exchanged through the interface.

Since our work is focused on the analysis of this interface and that of the quantities to be taken in account, a performance comparison of different simulation results is out of the scope of this chapter. However, we demonstrate that, in a typical 2G scenario with speech service, the proposed method provides the same results of the conventional approach. On the other hand, our approach becomes the unique adoptable when simulating complex scenarios with both circuit (CS) and packet switched (PS) bearers and mixed (voice and data) applications (i.e. refer to to the simulated traffic scenarios in chapter 2).

A similar approach has been proposed in [57] however this proposal is simpler to implement as it takes into account only the most relevant aspects of the link-level.

3.2 Link and Network Level Analysis

A thorough performance investigation of mobile radio systems requires the analysis of both link and network level aspects. Owing to the complexity of this approach, an integrated simulation tool addressing both these aspects is expected to requires hundred of CPU time to obtain results. This suggests to separate the two aspects, and exchanging data through a suitable interface [23] -[24].

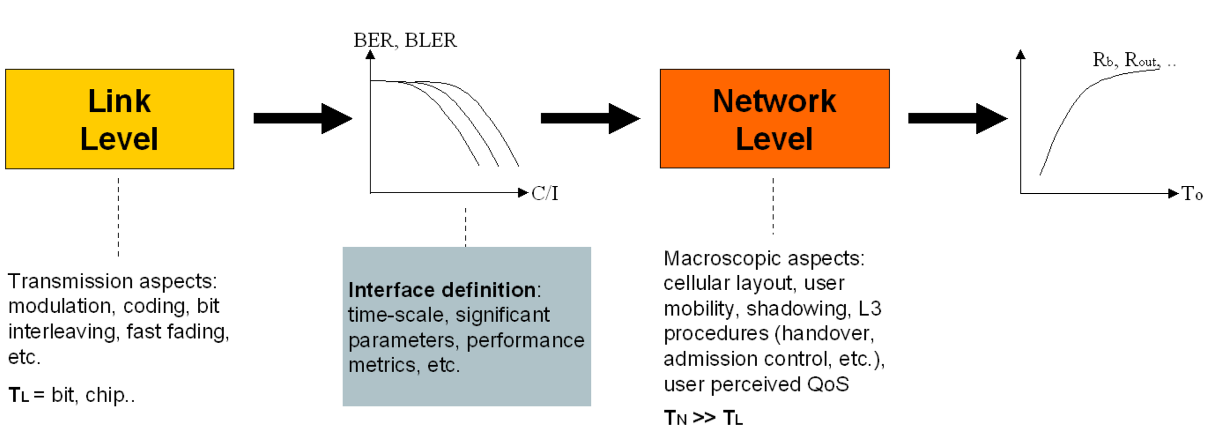


Figure 3.1: Link and Network levels in operation

Referring to figure 3.1, at link level, all aspects related to the transmission chain (modulation, coding, bit interleaving, etc) have to be taken into account in order to evaluate the performance metrics, typically expressed in terms of bit error rate (BER), frame error rate (FER) and block error rate (BLER) as a function of a suitable definition of the carrier to interference ratio.

At network level, the macroscopic aspects of the scenario (user mobility, cellular layout, shadowing, handover, power control and so on) have to be considered to evaluate the system performance (i.e. the merit figure R_b of rate of blocked users as a function of the offered traffic T_o). In order to measure the system performance, the network level simulator exploits the information provided by the link level (BER/FER/BLER) in many RRM procedures.

Obviously, in such a integrated approach, the interface between the two simulation tools has to be carefully designed since the time simulation step at link and at network level are usually different. While the simulation step T_L of link level tools is typically equal to the bit or chip duration, at network level the time resolution T_N has usually a much longer duration, equal to the physical time slot (see section 3.2.2) or to a time interval in the order of the second (see section 3.2.1).

3.2.1 Network level simulator with a large simulation step

Network level simulators with large value of simulation step (that is a step with a duration of hundreds of transmission frames) are usually used to investigate scenarios with CS services only. In such a network, the conventional approach [23] consists in considering that all the RRM procedures rely on averaged channel measurement reports provided by the mobile stations and the base stations, over the fast fluctuations due to fading to provide stable information about the current electromagnetic status of the network.

Assume this averaging is performed over an interval of T_a seconds. For instance in GSM, T_a could be the duration of a slow associated common channel multiframe (SACCH) equal to about 0.48 s [22]. So, the network level evaluation could be performed by choosing this interval as the time resolution (network level simulation step). Therefore, suitable definition of the measurements performed by mobiles and bases stations at physical level has to be given, according to the network level simulation step.

All changes in the scenario occurring with a rate larger than $1/T_a$ (the effect of fast fading, fast power control and so on) cannot be explicitly modelled, they have to be taken into account at link level and reported in averaged terms at network level. Then, by means of this approach the network level simulation tool is a computer program which is driven by an internal clock equivalent to T_a seconds. All the measurements at this level have to be considered averaged over a period of time equal to T_a.

At each simulation step, the network level simulator evaluates the carrier to interference ratio (i.e. C/I, E_b/I_0 or similar) averaged over a T_a period, and takes information about the quality of radio links by means of look-up tables containing the values of BER/FER/BLER computed at link level. It should be noticed that owing to the size of T_a, the fast fading has a mean value equal to zero within this period, and therefore the BER as a function of E_b/I_0 averaged over a period T_a can be considered equivalent to that of the average over infinitive period of time.

The corresponding link level simulator generates the look-up tables of BER/FER/BLER following these steps: a given reference value of E_b/I_o is selected and the whole transmission chain (modulation, coding, interleaving, etc.) is simulated by considering a huge number of transmitted bit. At the end of simulation, it evaluates the number of wrong detected bits and the corresponding BER. Look-up tables will contain on the y-axes the values of BER and on the x-axes the corresponding reference E_b/I_o value obtained by averaging the E_b/I_o over the whole simulation time. Note that, by means of this approach, the presence of fast fading generates, in each (link level) simulation step, instantaneous values of E_b/I_o different from the reference one.

3.2.2 Network level simulator with small value of simulation step

In the presence of systems where RRM procedures have a decision rate in the order of several hundreds of Hz (this is the case of WCDMA for 3G networks [21]), the size of the simulation step of the network level should be considerably reduced, by considering observation intervals corresponding to one time slot or frame, according to the system considered [25, 24].

In this case fast fading effects and the sudden changes in the level of radio interference have to be considered both at network and link level, whereas averaging radio measurements over seconds like in section 3.2.1 would lead to a rough approach. This approach, which implies a network level simulation step T_a equal to the time slot duration (typically less than 1 msec), makes network level simulations much more complex, but allows us to extend the investigation to PS services. With this time resolution, it is also possible to evaluate the correct/wrong transmission of each transport block and analyze the impacts of MAC-hs or RLC protocols on the TCP-IP level. Obviously, the whole system analysis is not allowed with the approach depicted in section 3.2.1.

Since we have to consider the effects of fast fading both at link and network level, in order to maintain the synchronization between the two simulators, it is necessary that the link level tool, which generally considers several fast fading contributions according to the selected channel (pedestrian, vehicular, etc.) generates also a suitable look-up table of correlated fast fading samples at a rate equal to the network simulator time step. These values will be used by the network level simulator.

Besides fading samples, the link simulator, that has typically a simulation step equal to bit or chip durations has to simulate the whole transmission chain to obtain the BER/BLER curves. Nevertheless, the number of erroneous bits have to be evaluated at the end of channel decoding phase. This means that, if we consider a mobile radio system with a time structure organized in frame and slots (i.e. GSM and UMTS), the evaluations on the link quality have to be computed with a rate equal to $1/T_B$, where T_B is the duration of a transport block, the elementary data unit managed by the encoder/decoder blocks [21], which corresponds to an interleaving length of K frames of duration T_{frame} ($T_B = K * T_{frame}$).

Using this approach, at network level, for each simulation step T_a, the software obtains the value of "instantaneous" carrier to interference ratio (E_b/I_o) affected by fast fading samples generated at link level and subsequently the network simulator calculates the averaged value over the duration of a transport block ($\langle(E_b/I_o)\rangle_{T_B}$). This value will be provided as input to link level which will return the corresponding BER/BLER related to that particular value of $\langle(E_b/I_o)\rangle_{T_B}$. Note that look-up tables generated following this approach are completely different from those created using the conventional approach (see subsection 3.2.1). Furthermore, coding gain effect is still present, even if (E_b/I_o) is averaged over all the duration of the transport block: the prerequisite is that for every single transport block, both link and network simulator use the same fading channel,

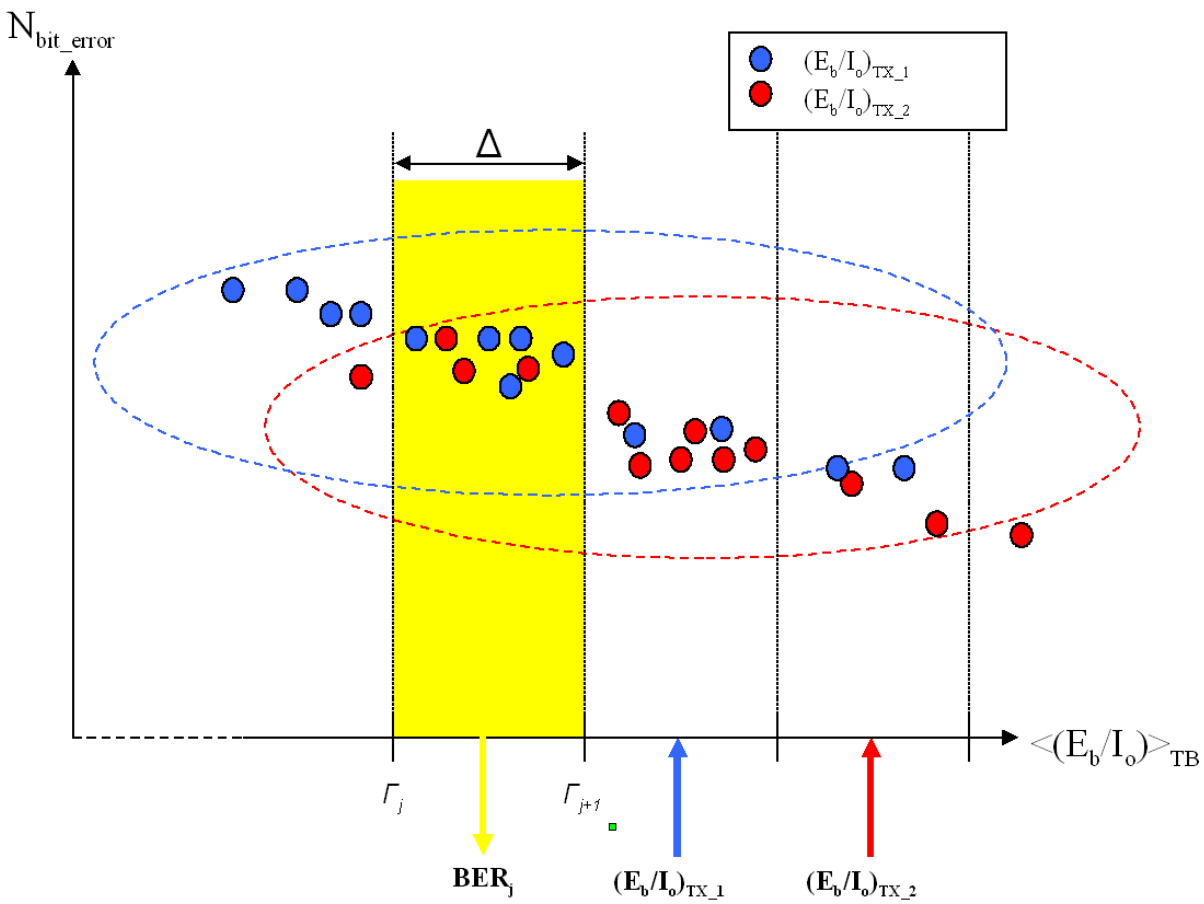


Figure 3.2: Example of data collection of $N_{bit-err}$ of the transmitted blocks

power control settings, code rate, etc.

Now, link level tool has to provide values of BER/BLER as a function of $\langle(E_b/I_o)\rangle_{T_B}$. In fact, in this case the value of $\langle(E_b/I_o)\rangle$ is averaged over a very short (few frames) period of time. To summarize, the link level tool analyzes the transmission of a huge number of bits, and considering the interleaving length, for each received transport block B_i, the algorithm calculates the averaged value (over the duration of the transport block) of the signal to interference ratio, $\langle(E_b/I_o)\rangle_{T_B}(B_i)$ and the number of wrong detected bits after the decoding of the transport block $(N_e(B_i))$.

The overall range of interest of E_b/I_o $((E_b/I_0)_{min}, (E_b/I_0)_{max})$ is subdivided in N intervals $\{(\Gamma_1, \Gamma_2), (\Gamma_2, \Gamma_3), \ldots, (\Gamma_{N-1}, \Gamma_N)\}$, where $\Gamma_1 = (E_b/I_0)_{min}$ and $\Gamma_N = (E_b/I_0)_{max}$. Each interval is identified by its mean value $(E_b/I_0)_j = (\Gamma_j + \Gamma_{j+1})/2$. Each value of $N_e(B_i)$ is assigned to the interval corresponding to the value of $\langle(E_b/I_o)\rangle_{T_B}(B_i)$. The values of $N_e(B_i)$ assigned to the same interval (Γ_j, Γ_{j+1}) are then averaged to provide $(N_e)_j$.

The look-up table contains on the x-axes the values of $(E_b/I_0)_j$ (in the following j will be omitted and the notation $< (E_b/I_o) >_{T_B}$ will be used to emphasize that the average is performed over the transport block) and in the y-axes $(N_e)_j$, that is the expected BER value when $< (E_b/I_o) >_{T_B}$ is measured in the interval (Γ_j, Γ_{j+1}) . We behave in similar way when obtaining the BLER look-up table: in this case the number of wrongly decoded transport blocks is counted, where a TB is considered to be wrong when at least a wrong decoded bit is present.

The example in figure 3.2 summarizes the main algorithm steps. In this case, two different values of reference $\langle (E_b/I_o) \rangle$ (that is averaged over an infinite period of time) have been used to transmit a large amount of Transport Blocks; let us recall that $T_B = K * T_{frame}$ and in each frame a certain number of timeslots will be required for the transmission of each transport block.

The effect of fast fading becomes evident by observing that the values of $N_e(B_i)$ of each transmitted block are associated to different intervals although related to the same reference value $\langle (E_b/I_o) \rangle$. At the end of the transmission of all the different reference $\langle (E_b/I_o) \rangle$, in each interval $(E_b/I_0)_j$, the average value of the contained $N_e(B_i)$ is calculated to derive the value of BER_j. Similarly, to obtain $BLER_j$, the average value of wrongly decoded transport blocks in $(E_b/I_0)_j$ is calculated.

3.3 An Example on how to interface Link and Network Levels

In this section, we present an example of interface between link and network level with reference to the UTRA TDD 1.28 Mcps system, also well-known as TD-SCDMA, which is characterized by high spectrum efficiency and by the capability of managing asymmetric traffic conditions [14, 15, 16]. This design has been implemented in the TD-SCDMA System Simulator developed during this thesis. However, the methodology previously described can be used to investigate mobile radio systems based also on different technologies (i.e. UTRA FDD).

The system is reproduced by means of a time discrete model with a simulation step equal to the time slot (675 μs) of the TD-SCDMA frame. The approach described in subsection 3.2.2 has been used.

The simulation environment we have created is composed of three main blocks: a Link Level tool, a Network Level simulator and an Interface module logically positioned between the two tools.

3.3.1 Link Level to Interface Module communication

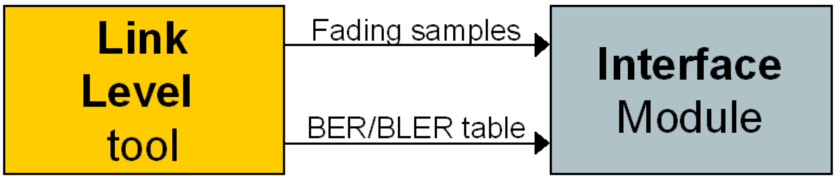


Figure 3.3: Link tool-to-interface module communication: block diagram

As previously described, in case of third generation cellular systems, the interaction between link and network level is much more involved and a large number of link level parameters have to be considered [16, 17]. Such parameters are: type of ITU propagation channel, direction of the link (DL/UL), number of timeslots and codes allocated in every radio frame, Spreading Factor (SF), code rate, interleaving period, receiver characteristics, etc.

The first step for the integration of the link and the network level aspects in the same system simulator is represented in figure 3.3: for each allowed configuration of physical parameters in the scenario that will be later simulated at system level, the link level simulator generates the related BER/BLER look-up tables (as described in section 3.2.2), which are then recorded in a new tool, called *Link-to-Network level Interface module* or more in short *interface module*

In this phase, a lot of curves have to be generated by the link level simulator; however, this effort is required only at the beginning of the study and doesn't depend on the network level configurations, i.e. these tables shall not be created again for different traffic statistics or different RRM algorithms, but only when the configuration of physical parameters changes which isn't expected to happen frequently during the system simulation study.

Moreover, for each propagation channel, the link level simulator provides to the interface module the vector of the time-correlated fading samples of resolution equal to the duration of the timeslot.

3.3.2 Interface Module to Network Level communication

In the system simulation, where "system" is defined as the union of link and network level issues, there is a continuous communication between the network level simulator and the interface module which contains all the link level characteristics in terms of BER/BLER curves and fast fading profiles (see previous subsection). The scheme describing the interactions between these two blocks is depicted in figure 3.4.

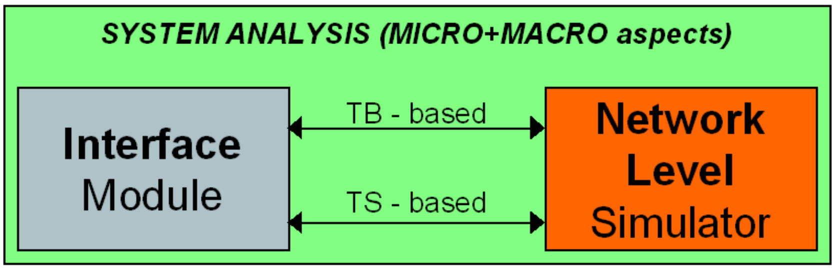


Figure 3.4: Interface module-to-network simulator communication: block diagram

The interface module communicates real-time with the network level simulator; the quantities exchanged by the two blocks can be classified in time slot (TS)-based (here denoted as Mode A) for fast fading enforcement and transport block (TB)-based (Mode B) for BER/BLER evaluation. We assume that the interleaving period is coincident with the duration of a TB.

Mode A - Time Slot Based

The system model considers a time discrete version, with a resolution equal to the timeslot duration, of the correlated fast fading. The link-to-network level interface computes the received carrier, C_{PLSH}, and the contribution of the interference coming from each single source (terminal or base station), I_{K_PLSH}; both terms are affected by pathloss (PL) and log normal shadowing (SH): the scope of this phase is to determine the final signal-to-interference ratio affected by fast fading.

C_{PLSH} and the various terms I_{K_PLSH} represent an input for the interface block which adds a sample of fast fading (ITU channels are considered) to both C_{PLSH} and each I_{K_PLSH} giving:

$$C = C_{PLSH} \cdot A_{fad} \tag{3.1}$$

where A_{fad} is a factor that represents the additional attenuation of the user signal caused by multipath propagation.

An expression similar to (3.1) can be written for each interfering signal I_{K_PLSH}:

$$I_K = I_{K\_PLSH} \cdot A_{K\_fad} \tag{3.2}$$

Indeed no assumptions on gaussian distribution of inter-cell interference is made. This more realistic way of proceeding is necessary in our simulations, since high data rate packet transmissions are characterized by sudden variations of the interfering power that cannot be modelled using the gaussian hypothesis.

Once the fast fading samples have been added, the quantity $(E_{bc}/I_o)_{fading}$ is evaluated as:

$$\left(\frac{E_{bc}}{I_o}\right)_{fading} = \left(\frac{C}{I}\right) \cdot \frac{Q}{log_2 L} = \frac{C}{I_{inter} + \alpha \cdot I_{intra} + N} \cdot \frac{Q}{log_2 L} \tag{3.3}$$

where C is the received power, I_{inter} is the inter-cell interference, I_{intra} is the intra-cell interference, N is the thermal noise power, Q is the Spreading Factor and L denotes the levels of modulation (L=4 in QPSK), and α represents the orthogonality factor. In uplink we substitute α with $(1-\beta)$, where β represents the Multi User Detection (MUD) efficiency. All the terms of power and interference are affected by fast fading as explained in equations 3.1 and 3.2.

E_{bc} denotes the encoded bit energy in the section before the decoder. Since the interleaving duration is equal to 10, 20, 40 or 80ms, corresponding to 2, 4, 8 or 16 subframes, the interface module through the link simulator results couldn't evaluate the signal-to-interference ratio after decoding with a resolution equal to the time-slot.

Mode B - Transport Block Based

At the end of the reception of a transport block, the network level simulator evaluates (E_b/I_o) averaged over the duration of the transport block (where E_b is the information bit energy after decoding):

$$\left\langle \left(\frac{E_b}{I_o}\right) \right\rangle_{T_B} = \frac{\sum_{i=1}^{M} \left(\frac{E_{bc}}{I_o}\right)_{fading}}{M} \cdot \frac{1}{r} \tag{3.4}$$

where r denotes the code rate and M is the number of timeslot required for the transmission of a single transport block (in mode A, link level has previously provided these M values of E_{bc}/I_o to the system simulator).

E_b/I_o represents an input for the link-to-network level interface module which returns, through look up tables, the BER and the BLER referred to the considered block.

In our network simulator, during a CS connection (RLC protocol in transparent mode), long term averaged BER values determine the quality of the connection. Viceversa, in case of PS sessions, through the value of the measured BLER, the system evaluates whether a transport block belonging to a data packet is correctly received. For PS services we use a pair of RLC protocol istances in acknowledge mode, providing a reliable radio bearer service, including error correction by automatic retransmission, thus when the transmission fails, the transport block is re-transmitted.

Thanks to the proposed approach, described in section 3.2.2, it is actually possible to calculate the Block Error Rate related to each transport block transmission; on the contrary, the more conventional approach in section 3.2.1 doesn't allow such a possibility: a BLER averaged over a long period wouldn't be useful when evaluating the data link performance of a PS call.

Direction	**Time interval**	**Quantity**	**Definition**
From NETWORK Level *To* INTERFACE module	TS	C_{PLSH}	Received user code power affected by path loss and shadowing
	TS	I_{k_PLSH}	Received power coming from the generic $k-nth$ interference source (UE or BS), affected by path loss and shadowing
	TB	$\left\langle\left(\frac{E_b}{I_o}\right)\right\rangle_{T_B}$	Signal-to-Interference ratio, averaged over the duration of the Transport Block, after despreading, demodulation and decoding
From INTERFACE module *To* NETWORK Level	TS	$\left(\frac{E_{bc}}{I_o}\right)_{fading}$	Signal-to-Interference ratio affected by fast fading, before decoding
	TB	BER	Information bit error probability in a Transport Block (CS)
	TB	$BLER$	Transport Block error probability (PS)

Table 3.1: Definition of quantities exchanged between the interface module and the network simulator

In table 3.1, the definition of the relevant quantities exchanged between the interface module and the network simulator is summarized; the time interval in which they are calculated and the direction of the information are reported.

3.4 An experiment for the validation of the proposed link-to-network interface method

In this final section we will show the results of a simulative session drawn to compare in a basic situation the BER calculated in the two approaches explained in sections 3.2.1 and 3.2.2. For more complex scenarios with both CS and PS users, refer to the simulation results in chapter 2 which were obtained with the described link-to-network level interface module implementation.

As the aim of this investigation is the comparison between the two approaches, a simple case is here considered: the simulation was performed in the absence of convolutional coding and fast power control; all the results are related to only circuit switching sessions, because the approach in section 3.2.1 isn't suitable for packet data analysis.

In figure 3.5, the logical steps adopted during this test are described.

Firstly, for each (E_b/I_o) reference, in the range of interest, through the link level simulator implemented with the conventional approach described in section 3.2.1, it has been measured the BER as a function of $\langle(E_b/I_o)\rangle$, that is as a function of the the Signal-to-Interference ratio averaged over a long time period - see the red curve in figure 3.6.

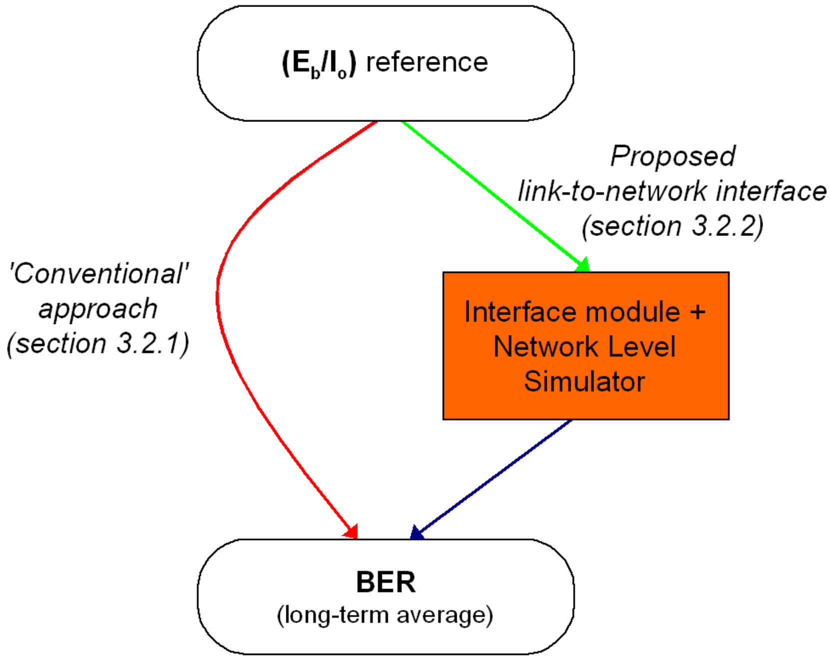


Figure 3.5: Validation experiment: block diagram

Secondly, by using the link-to-network level interface approach proposed in this work and described in section 3.2.2, it has been generated the curve of BER as a function of $\langle (E_b/I_o) \rangle_{T_B}$, that is as a function of the Signal-to-Interference ratio averaged over the duration of the Transport Block - see the green curve in figure 3.6 - that is over a duration much shorter compared to the first method.

As it can be observed in figure 3.6, the BER curves generated by the link level simulator based on the two approaches are significantly different (see the red and the green curve). In particular the curve based on the averaged values of (E_b/I_o) over a period T_B shows smaller values of BER for the same nominal value of (E_b/I_o). The behavior can be explained if we consider for instance a value of (E_b/I_o) of 5 dB. In case of conventional approach, the (E_b/I_o) on the x-axes is averaged over a large time-interval, this means that, owing to fading, in some cases the "instantaneous" value of (E_b/I_o) maybe significantly smaller than 5 dB, with a consequent increase of the "instantaneous" BER. As the BER plotted in figure 3.6 is averaged over a large time interval, the values of BER corresponding to small (E_b/I_o) have a weight within the average larger than those corresponding to large values of "instantaneous" (E_b/I_o).

Afterwards, for this specific validation activity, it has been added to the network simulator a function for the calculation of the average of all $\langle (E_b/I_o) \rangle_{T_B}$ samples received during the system simulation and for the calculation of the related averaged BER. For

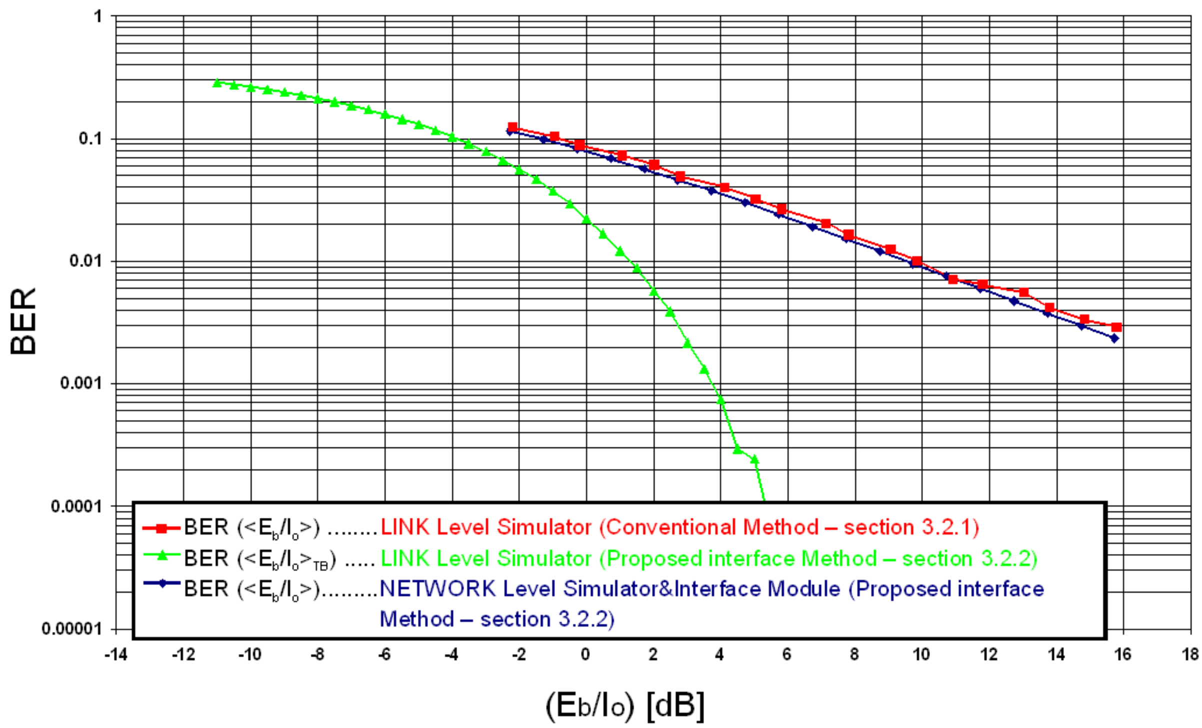


Figure 3.6: BER as a function of $\langle(E_b/I_o)\rangle$ and $\langle(E_b/I_o)\rangle_{TB}$.

each (E_b/I_o) reference, a new system simulation has been performed, and the blue curve in figure 3.6 of BER as a function of $\langle(E_b/I_o)\rangle$ as been obtained by the network simulator. In this case $\langle(E_b/I_o)\rangle$ has been calculated as an average over a huge number of transport blocks; in figure 3.6, it can be observed that this method provides therefore similar results compared to the conventional approach when focusing the attention on merit figures averaged over long term period (see the red and the blue curve).

Although in this validation experiment the agreements between the conventional approach performance and the "instantaneous" one with post-averaging at network level is very good, we have to highlight that the proposed interface is the only suitable approach when a more accurate analysis is needed and PS traffic is simulated.

To conclude, in this chapter we have described a methodology to interface link and network level tools. Performance evaluation of modern mobile radio systems requires the joint consideration of link and network level aspects; this requires the use of an approach which is different from that used in the past for 2G systems when the two aspects were completely separated and few look-up tables, with values of BER averaged over fast fading, were sufficient to take link level into account.

We stress that although our approach becomes equivalent to the conventional one when circuit switched services are considered, the conventional approach cannot be used in case of packet switched applications owing to the necessity to evaluate sudden variations of the link quality. In this latter case, only approaches based on the proposed implementation can be applied. The methodology presented here is quite general and is valid for different radio access technologies.

Chapter 4

Architectures for Heterogeneous Wireless Networks

Third generation for cellular telephony is already a reality. A new way to communicate and a growing number of different services will be the challenge for UMTS. High data speed will allow video-communication and mobile Internet on cellular terminals. But the request for bit rate will never stop, and greater bandwidth will be necessary.

At the same time WLAN technology is starting to be an ordinary way to realize mobile connection to the Internet, while companies all over the World realize wireless connections in particular locations (hot spots), like airports or hotels. Anyone, owning the appropriate technology on his laptop, can connect to the Internet in these places, at a reasonable price with satisfactory connection speed.

On these conditions, UMTS and WLAN interworking becomes a really significant issue: combining the two technologies would double the available resources. How to gain this interwork is a large field for researches.

The heterogeneous technologies employed in 3G cellular networks and WLANs bring many challenges to the interworking. Based on different radio access techniques, cellular networks and WLANs present distinct characteristics in terms of mobility management, security support, and quality of service (QoS) provisioning. In order to achieve seamless integration, these issues should be carefully addressed while developing the interworking schemes [28]).

Different aspects have to be considered while discussing about UMTS and WLAN interworking: in this chapter, the state of the art in services, the most important network architecture solutions, comprising an original one, and some functional issues will be discussed; in the next chapter 5, an advanced Common Radio Resource Management algorithm will be presented and the related performance will be simulated through the simulation platform SHINE, developed in our laboratories (see appendix A).

4.1 Service interworking

In [61] 3GPP defines six levels of interworking, focusing on the service provided:

Level 1: "Common billing and customer care". Users pay a common bill to connect both to UMTS and to WLAN. This level needs only a commercial agreement among the different operators, sharing their costumer information; no technologic advance is required. Each system offers the same service, as there was no interworking.

Level 2: "Common access control and charging". WLAN reaches the same level of security of UMTS. Users notice the same interface connecting to both the systems.

Level 3: "Access of all UMTS Packet Switching based services". No handover is provided, but both technologies offer the same services: QoS is managed even by WLAN.

Level 4: "Service continuity". Handover is managed, but the user can experience service interruption or noticeable degradation.

Level 5:" Seamless mobility". Users do not notice any difference using any of the two systems neither during the handover. Circuit Switching services of UMTS are not provided by WLANs.

Level 6: "Access to Circuit Switching services". Even Circuit Switching services are supported by WLANs.

The basic target of our investigations was the interworking level 5. Nevertheless, although the road map of the development of the Core Network foresees in the next future an all-IP Core Network, we propose an integrated UMTS-WLAN architecture which allows the WLAN users to exploit also the same Circuit Switching services provided by the Core Network to the UMTS users, thus achieving the interworking level 6.

4.2 Proposed architectures

Different hypothesis are considered about how should the networks be connected. Three in particular are proposed (for example in [28]): loose coupling, gateway and tight coupling.

Loose coupling (fig.4.1) considers the two networks as separated. To permit mobility, maintaining the session open, Mobile-IP is implemented. A user is referred to a server called home agent, owning the IP address of its connection. When it moves to a different server (called foreign agent), the two servers communicate to each other. Every message to the user is now sent to the home agent (maintaining the same IP address), whose duty is to redirect it to the foreign agent. This solution makes new networks easy to be installed and it does not require a tight cooperation among different operators. Real-Time services hardly will survive during a handover, because of the time latency introduced by mobile IP.

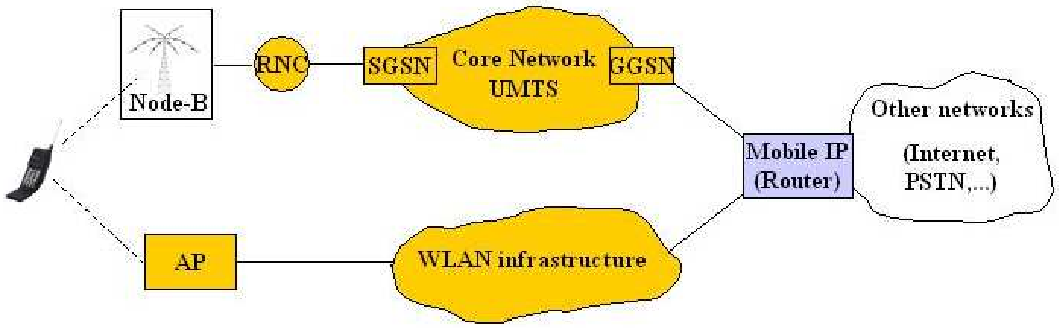


Figure 4.1: Loose coupling

Another solution is to introduce a new node, a *Gateway* (fig.4.2), connecting the WLAN network to SGSN and GGSN in the UMTS core network. The Gateway to SGSN connection is used while a WLAN customer is roaming on UMTS, while the Gateway to GGSN connection is used when a UMTS customer is roaming on WLAN. Handover can so be realized in a quicker way, while maintaining the two networks enough separated. Operators of the two networks must cooperate, while hardware of the two parts needs to be connected together. Introduction of new networks is not that easy.

The third solution is to use the WLAN technology as an access stratum to the UMTS network (*Tight coupling*, fig.4.3). An interworking node is required to connect the WLAN AP to the UMTS network. This connection can be done in the RNC or in the SGSN. In the former case the AP acts as a Node-B. In the latter it acts as an RNC; a connection even to the neighboring RNCs can be thought in this case. This solution reduces the handover latency, making the UMTS Core Network to be the bottle neck for the traffic

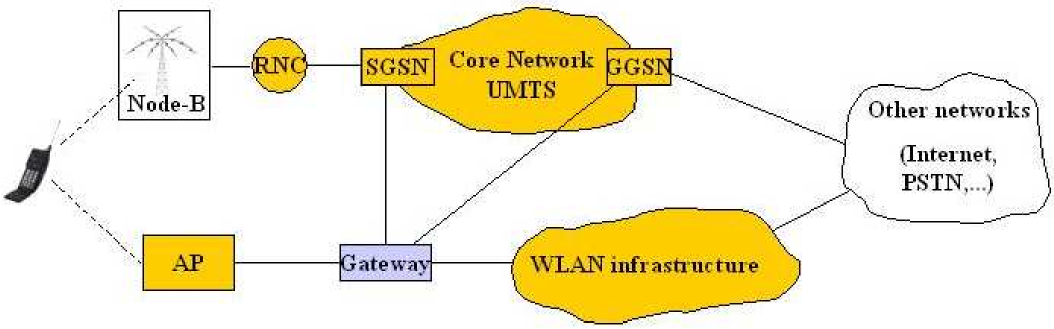


Figure 4.2: Gateway approach

speed; the Interworking Unit must be designed appropriately.

Based on the Tight Coupling concept, a number of leading international companies in the field of wireless communications have jointly developed a set of open specifications (that can be found in [35]) in order to encourage the deployment of interworking networks involving GSM and unlicensed spectrum networks (i.e. WLAN, bluetooth, etc.): the result was the design of a new technology called UMA (Unlicensed Mobile Access).

Although the Loose Coupling approach seems to be the easier and thus the more practicable way for 3G and WLAN interworking, the UMA specifications draft was a clear demonstration that the Tight Coupling approach was much more than a "proposed solution": actually, the 3GPP continued the work based on UMA, and under the designation of *Generic Access Network* (GAN) in 2006 the first specifications for such a system architecture were officially released [60].

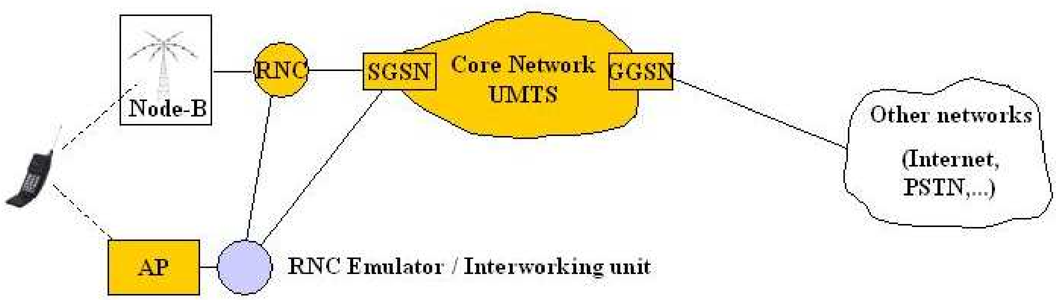


Figure 4.3: Tight coupling

4.3 Architectural and implementation issues

Assuming that UMTS/WLAN dual mode terminals are available, two aspects have to be detailed in order to make UMTS and WLAN interwork: an architectural proposal for the integrated network and a feasible practical realization.

4.3.1 System Architecture

Among the different solutions proposed from an architectural point of view (as detailed in subsection 4.2), in this thesis, the performance of the UMTS-WLAN integrated network is investigated with reference to the tight coupling approach, which is at the same time the most challenging and the most promising architectural solution, as already mentioned.

The basic prerequisite to deploy a tight interworking is that the technology of **dual-mode UMTS&WLAN mobiles** is available: this type of equipment should be able to switch between the two radio access network, i.e. when in UMTS mode, the mobile should periodically sense 802.11 coverage, for example every 2-3 sec.

Different solutions can be envisioned to realize the tight coupling between the WLAN Access Point (AP) and the UMTS Terrestrial Radio Access Network (UTRAN): the AP could be seen, for instance, as an additional cell or as an additional Node-B connected to the same Radio Network Controller (RNC) which controls the adjacent UMTS Nodes B.

Let us observe, however, that the Iub control plane (see left side of figure 4.4) should be adapted to support an "enhanced" cell or Node-B: consider, for instance, that an NBAP (Node-B Application Part) Cell Setup procedure for WLAN AP should be introduced in 3GPP specifications. The user plane implementation is also a limitation, since the frame protocol should be adapted to the higher WLAN rates.

Thus, in our opinion, the best architectural solution is moving to a Radio Network Subsystem (RNS) perspective (see right side of figure 4.4): an RNC emulator is inserted between the AP and the Core Network acting as an interface between the WLAN world and the UTRAN, without any impact on both the two technologies and taking advantage of the same MSC/SGSN services already provided to UTRAN by the Core Network.

This way, we can exploit the essential fact that neither the adjacent RNCs nor the Core Network need any knowledge about the internal resource management of RNSs. Moreover, we think that an Iur logical interface between the standard RNC and the RNC emulator is not required (see figure 4.4): since macro-diversity between cells under Serving-RNC (UMTS cells under standard RNC) and cells under Drift-RNC (WLAN AP under RNC emulator) can not be exploited, Iur user plane would not be used.

Regarding signalling for inter-system handover procedure, we could then use the Serv-

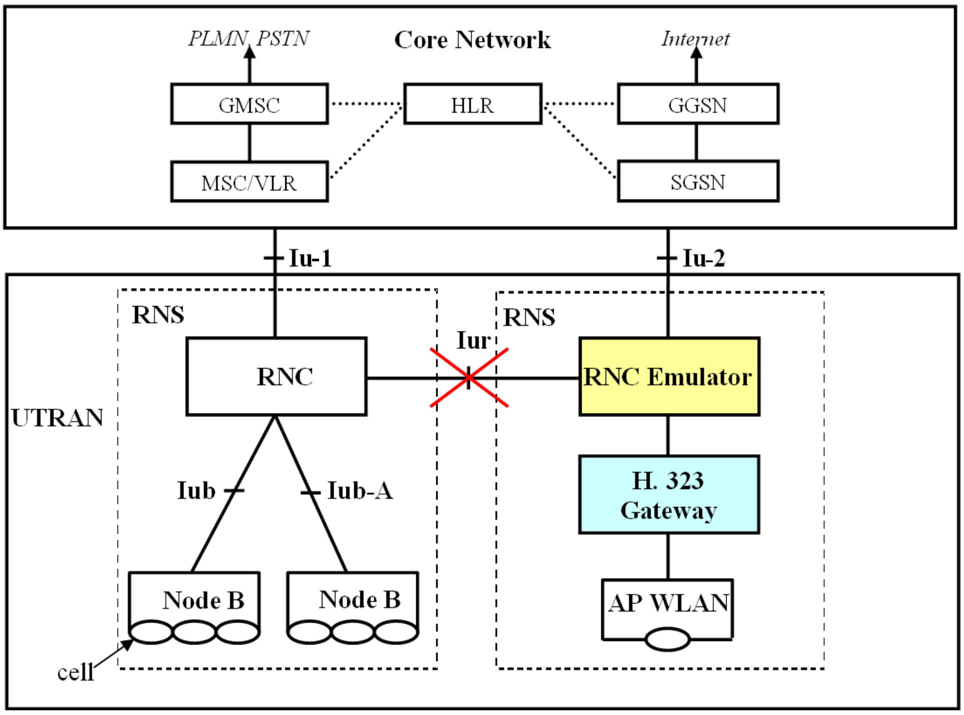


Figure 4.4: UMTS - WLAN Inter-working architecture: proposed logical scheme

ing RNS Relocation function already provided by 3GPP specifications, that manages the mobility of the Iu connection from a RNS to another (e.g., from Iu-1 to Iu-2 in figure 4.4) without loss of Packet Data Protocol -PDP- context or any other session management information (see [28] for a possible message flow allowing the execution of the relocation procedure): in this way, a **seamless handover UMTS-WLAN** without perceiving any service interruption can be performed.

Concerning data flows, two cases have to be considered: interactive/background services and conversational services (voice or video) [62]. In the first case, when the mobile terminal is camped within the WLAN hot-spot the Iu-PS data packets coming from SGSN and carried over Iu-2 (see figure 4.4) should be forwarded by the RNC Emulator to the AP and vice versa; no procedure change is needed when, after a Serving RNS Relocation, the new Iu-PS is Iu-1.

Regarding conversational services, our proposal is to exploit real-time transport characteristics of Iu-CS also when the mobile terminal is connected to the WLAN AP and data are packetized, as in the case of VoIP calls: the main issue is how to adapt RTP/UDP/IP packets coming from the mobile terminal to a circuit switched based context.

To this aim, without any impact on the mobile terminal functionality, an H.323 [45] gateway between the AP and the RNC Emulator could be inserted [46], providing protocol translation and media transcoding between the endpoint of the PS domain (the WLAN AP) and endpoint of the CS domain (RNC Emulator).

As far as the protocol translation is concerned, the easiest solution could be that Non-Access Stratum messages from MSC (call setup messages, for instance) are transparently forwarded by the RNC Emulator towards the H.323 Gateway; vice versa for messages coming from the AP WLAN. Note that Non Access Stratum messages in UMTS only scenarios are already exchanged between mobile terminals and MSC without RNC knowledge (Radio Access Network Application Part -RANAP- messages on Iu and Radio Resource Control -RRC- Direct Transfer messages on Uu, for instance).

An H.323 Gatekeeper should also be introduced to translate E.164 addresses (i.e., phone numbers) into Transport Addresses (e.g., IP address and port address).

4.3.2 Implementation issues

Having defined the conceptual architectural solution for UMTS/WLAN integration, a feasible practical solution has to be investigated in order to conveniently provide a logical and a physical link between the RNC emulator and the Core Network.

Here we imagine that the same network provider will manage both the WLAN and the UMTS networks, hence it is possible that the WLAN AP, as well as the H.323 Gateway and the RNC emulator will be co-located with an UMTS Node-B, to reduce costs and to ease maintenance.

In this case, the physical contiguity between the RNC emulator and the Node-B suggests to adopt already existing UTRAN interfaces for the physical and logical WLAN-UMTS integration. Due to the fact that Layer 1 Iub and Iu comply with the same requirements [47], and that the physical layer provides the same ATM services according to ITU-T I.361, we could multiplex different logical interfaces over the same physical link. In particular, an ATM switch could be inserted in the physical connection already existing between the Node-B and the RNC, in order to switch the flows to/from the Node-B (logical Iub interface) and the flows to/from the RNC Emulator (logical Iu interface) (see figure 4.5).

In this scenario the ATM switch should be located close to the Node-B and consequently also close to the AP, which is also convenient for network management and maintenance.

Thus, the physical link between the ATM switch and the standard RNC carries two different logical interfaces: the Iub from Node-B and the Iu from WLAN RNS. The latter should then transparently cross the standard RNC exploiting the RNC switching capability, thus making the RNC emulator directly connected to the Core Network, according

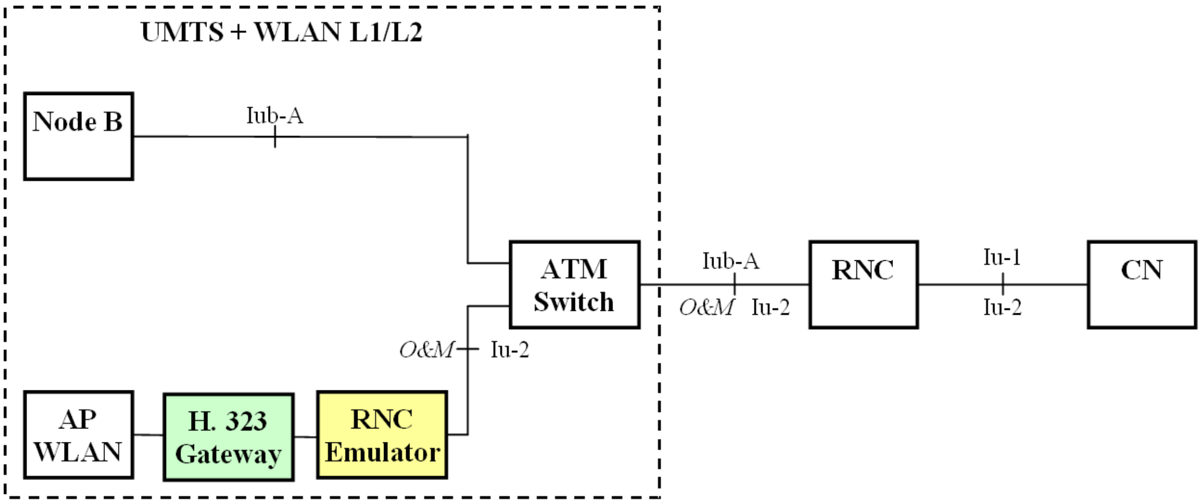


Figure 4.5: Node-B and WLAN integration: implementation

to the conceptual scheme of figure 4.4.

It follows that the link between the standard RNC and the CN conveys Iu for both the standard RNS and the WLAN RNS.

It is worth noting that, apart from the short physical links connecting the AP to the the ATM switch through the H.323 Gateway and the RNC emulator, no significant physical interconnections have been introduced by this solution; obviously, the capacity of the link between the ATM switch and RNC should be adapted in order to convey also the WLAN traffic.

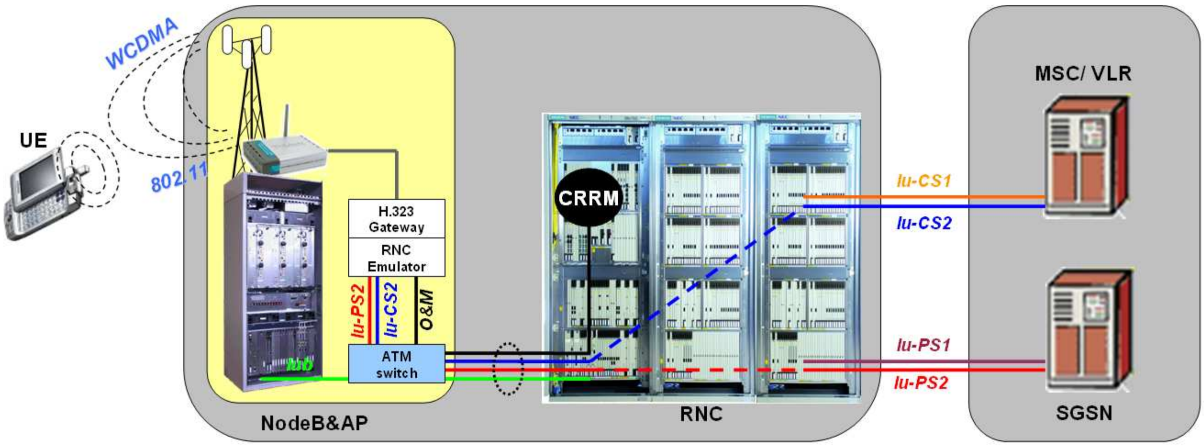


Figure 4.6: Evolved multi-standard (UMTS & WLAN) NodeB communicating to the Common Radio Resource Management (CRRM)

Finally, another logical link could be added between the standard RNC and the RNC emulator, (the path would be the same used for Iu by WLAN RNS - see figure 4.6): this link could be dedicated for Operations and Maintenance (O&M) functions: we expect that a Common Radio Resource Management functionality between different RNSs and particularly between a standard RNS and a WLAN RNS could be carried out through this O&M communication link, leading to an increase of network performance.

In the figure 4.6, we considered the CRRM element located in the RNC, since it could communicate at the same time to several WLAN RNSs concentrated to the same RNC. As will be discussed in detail in the next chapter, the CRRM communicates with both the UMTS Radio Resource Management (RRM) block located in the standard RNC and the WLAN RRM block.

In the next chapter we'll show the benefit introduced with our architectural solution by the proposed inplementation of this new network element, the CRRM entity.

Chapter 5

Common Radio Resource Management UMTS & WLAN

In this chapter, we will investigate the possible advantages introduced by the Common Radio Resource Management (CRRM) for a heterogeneous integrated and interworking UMTS-WLAN network; the performance evaluation will be carried out through simulations, performed adopting our advanced SW platform called SHINE illustrated in Appendix A.

The work has been organized in this way: in section 5.2, the required interactions between the CRRM entity and the local RRM entities will be described; in section 5.3, the main functions of the RRM UMTS and the RRM WLAN will be presented; afterwards, in section 5.4, the CRRM algorithm we projected will be explained in detail. After a proper definition of performance measurement in section 5.6 for such a heterogeneous network, we will present the studied scenario in section 5.7 based on a hotspot of high density traffic covered by both UMTS and WLAN radio access networks. In the numerical results section 5.8, the benefit in the system capacity provided with the implementation of a CRRM QoS based will be shown.

5.1 The CRRM challange

In the last few years several projects were dedicated to the issue of 3G/WLAN interworking networks (e.g. [36], [37]) and many papers appeared, investigating specific aspects. Most of them are related to architectural and/or signalling issues (see [27], [28], [38]) while some papers were focused on TCP performance in the presence of vertical handover (i.e. handover among different technologies) [39]. Handover is indeed a very relevant aspect in such networks and has been investigated also in [29], [30], [40] and [41]. The management of the integrated network, with particular attention to profiles and parameters to be monitored, has also been investigated in [42] and [43].

However, although many operating schemes have been proposed, no deep investigation of the impact of networks integration on the overall performance experienced at application level by the final user has been carried out so far. In this study it is envisioned the possibility to serve by means of the WLAN technology, which can operate at bit rates considerably higher than 3G systems, those voice calls that within the hot spot cannot be served by the UMTS network (and would therefore be blocked) in case of saturation of its radio resources.

A format conversion from circuit-switched (CS) voice flows into packet-switched (PS) Voice over IP (VoIP) flows (and vice versa) is required in order to allow the transcoding between the speech bearers over the Iu-CS interface and RTP/UDP (see chapter 4).

In previous chapter, we already described in detail the required aspects for a tight integration between UMTS and WLAN. As a reminder, here we summarize the main points:

- Availability of a dual-mode mobile UMTS - 802.11, capable of switching between the two technologies.
- All type of services (i.e. speech and data transmission) shall be offered by the two networks.
- Handover "seamless" between UMTS and WLAN.
- Feasible interworking architecture: re-use of already existing logical interface Iu-CS/Iu-PS and creation of a new Network Element providing policies for the Common Radio Resource Management (CRRM).

In this chapter, we focus on the functionalities and benefits provided by the CRRM block. The Common Radio Resource Management 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. We'll show algorithms we implemented in the CRRM in a position to improve the overall system performance. The performance evaluation of a heterogeneous integrated and interworking UMTS-WLAN network will be carried out through simulations, performed adopting the SHINE platform.

Actually, although some tools developed for the simulation of heterogeneous network have been already realized and described (see [37] and [44], for instance), we differentiate from previously proposed solutions since we investigate the performance from the users' point of view, reproducing in details all aspects related to all levels of the protocol stack affecting the perceived Quality of Service (QoS), from the application to the physical layers, including channel characteristics of each different radio technology (that is, for example, unlike the just quoted tools which work on pathloss information only and a rough model of the physical channel, we consider also the fast fading effect in our home-made simulation tool). Moreover, none of the published papers considers the voice flows somehow served by WLANs, thus limiting the capacity enhancement to data packets only.

The feasibility and the benefits of 3G/WLAN interworking are here investigated with reference to the Chinese TD-SCDMA UMTS technology [47] and to the IEEE802.11e [31] WLAN technology (adopting the IEEE802.11a physical layer [13, 10]).

5.2 Interactions between CRRM and RRMs

Before discussing the algorithms to be implemented in the CRRM block, the type of relation between the CRRM block and the already existing Radio Resource Management (RRM) blocks has to be defined.

Let us recall that one RRM function is already located in each Radio Access Technology (RAT) to optimally manage the resources available in the single radio technology. In our vision of different interworking networks, a centralized CRRM is needed to coordinate and define policies to be followed at local level by each RRM. For the sake of simplicity we'll consider a system constituted of a single CRRM entity and many (2 or more) RRM entities, that is we're not interested in this study to the communication among different CRRM entities. One feasible realization of this system is to locate the CRRM in the UMTS RNC (see chapter 4); the local RRMs shall be located in the RNC (UMTS) and in the RNC Emulator (WLAN).

We estimate that the interactions between CRRM and RRM entities (that is the arrows between each RRM block and the CRRM in figure 5.1) involve mainly these functions we'll present in the next subsections:

- Network topology information.

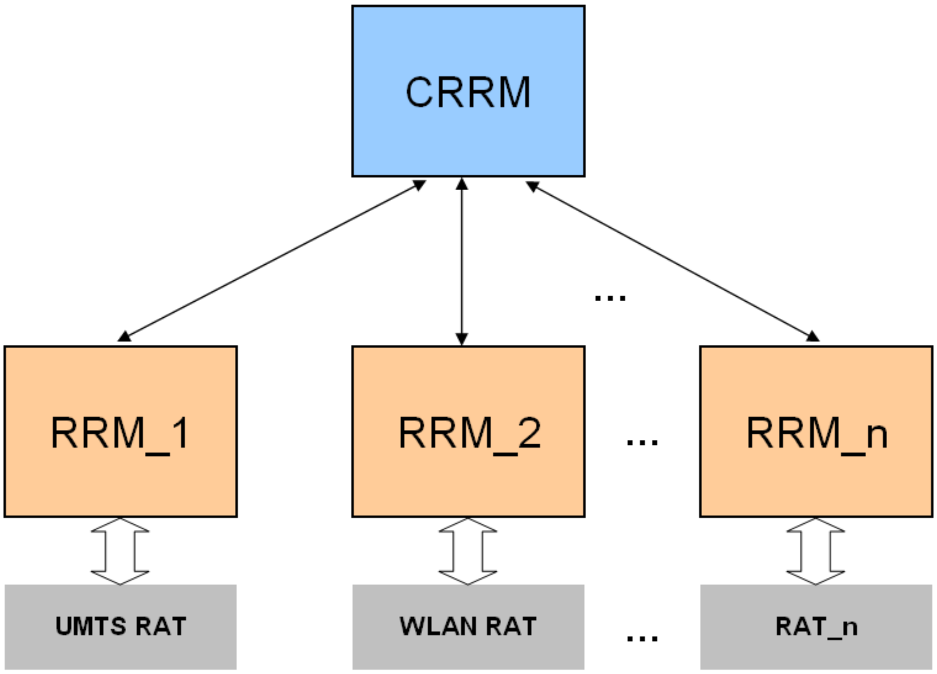


Figure 5.1: RRM & CRRM relations

- Network load report.
- RRM report.
- CRRM decision.

5.2.1 Network topology information

The CRRM block needs to know the geographic position of the UMTS cells and of the WLAN APs. This allows the CRRM to automatically create the external adjacent cell relations for the Inter-system Handover between WLAN and UMTS. As an alternative, CRRM may use the Location-area identification comprising the mobile country code, mobile network code, and location area code (LAC) corresponding to the cell/AP [58] (i.e. by using a different LAC for each AP WLAN).

This type of information may be exchanged between the RRM and the CRRM only once, at the setup of the cell/AP.

In SHINE, our simulation platform for heterogeneous networks, the geographic information is communicated at the start up of the simulation.

5.2.2 Network load report

Each already existing RRM shall be enhanced to report to the CRRM some merit figures directly in relation with the load of the network. The purpose of these reports is that the CRRM algorithm should use this information to increase the overall system capacity, therefore a thoughtful choice of the merit figures shall be done. Each wireless technology has different characteristics in terms of air interface technology, cell-size, coverage, etc., therefore the merit figure reported may differ from a RRM to another: for example the UMTS cell could report the level of uplink interference, the value of transmitted carrier power, the OVSF code tree usage, the latency time of HS-DSCH queues in the UMTS NodeB, whereas the WLAN AP could report the channel occupation rate or the collision rate.

Because of the huge set of merit figures that each local RRM can choose in order to report the load of its cells, the tasks of the CRRM could become extremely complicated. As an example, the CRRM may have difficulties to interpret the value of the transmitted carrier power (TCP) of a UMTS cell in order to take the right action (i.e. to move users from UMTS to WLAN because of high TCP in the UMTS cell): if the UMTS cell reports TCP equal to 10W, the CRRM doesn't understand if it is a high or low value because the UMTS cell could be a micro cell (for example, maximum TCP equal to 12.5W) or a macro cell (for example, maximum TCP equal to 40W). One solution in order to make easier the task of the CRRM is to ask the local RRMs to report percentage values and not absolute values (in our example, compared to the maximum TCP, the micro cell would report 80%, the macro cell would report 25%).

Another point to be discussed is the amount of network load information to be exchanged between the RRM and the CRRM: main concern is on the RRM located in the UMTS RNC which already handles several hundreds of UMTS cells. In order to reduce the complexity of the UMTS RRM and the quantity of signalling between the RRM and the CRRM, we decided the two following points:

1. the local RRMs send triggered measurements to the CRRM, that is reports from the RRM to the CRRM aren't periodic, but only when a threshold is exceeded (above or below) a report is generated;

2. only the cells which are deployed in areas covered by different radio technologies shall send their network load information through RRM (i.e. areas covered by UMTS cells only don't require the UMTS RRM sends any report to the CRRM).

Final remark about the network load information exchange is on the time scale: in order to better the system capacity, the CRRM needs to exploit in real time the complementary characteristics offered by the different radio access technologies. Therefore each local RRM has to provide to the CRRM up-to-date information, averaged over short time intervals (i.e. 100ms, 1s): for example, in order to solve a radio congestion situation in one system, it's required the CRRM quickly receives this information and immediately takes the decision to command an intersystem handover to an alternative radio technology before any call is dropped.

In SHINE, as basic report, we decided the WLAN RRM entity reports the channel occupation rate of the AP, whereas the UMTS RRM entity reports the downlink (NodeB to UE) and uplink (UE to NodeB) OVSF code tree usage rate. However, the platform may be extended to work with other merit figures.

5.2.3 RRM report

The local RRM entities shall send to the CRRM some of the outcome of their RRM algorithms.

First of all, in order to implement a CRRM "Service Based" (see subsection 5.4.1), it's necessary each local RRM informs the CRRM when a call setup is attempted in its own RAT: the RRM shall send the QoS profile received from the Core Network in the *RANAP RAB Assignment* message [59]. With this information, the CRRM "Service Based" communicates to the local RRM whether the Radio Bearer shall actually be setup on its own RAT, or whether a Directed retry procedure shall be started: directed retry is the process of assigning a user to a radio resource that does not belong to the serving RRM by employing relocation procedures. Let us highlight that in the latter case, the relocation procedure between the 2 RRMs uses standard 3GPP signalling [59] between the RNC/RNC Emulator and the Core Network, that is the implementation effort is reduced.

Secondly, some of the local RRM functions (see section 5.3), like Admission Control, may decide to restrict the service availability in one system because of a radio resource shortage condition. If interworking UMTS-WLAN is available, the RRM shall also be enhanced to inform the connected CRRM that such a type of event has happened and therefore ask "help" to the CRRM. In this case, there is a possibility that a blocked call on one system, may be re-directed by the CRRM to an alternative system.

For this issue, the main RRM blocks reporting their decisions to the CRRM should be the Resource Allocation (both admission control and code allocation), the Congestion Control, the Pre-Emption and the Restriction Control.

In SHINE, the RRM decision report we use is the Resource Allocation function. However, the platform may support other RRM reports.

5.2.4 CRRM decision

The CRRM has the fundamental function of coordinating the use of the available radio resources in heterogeneous networks scenarios in order to improve the overall system performance. The CRRM decisions, like moving one user from one congested system to another, are communicated by the CRRM to the local RRM; the RRM shall be enhanced in order to start the standard Intersystem Handover procedure once it receives this order from the CRRM.

In SHINE, the CRRM decisions are the initial RAT selection and the Intersystem Handover command.

In figure 5.2, a diagram resuming the main interactions between the CRRM and each RRM. At the system setup, every RRM sends the NETWORK_TOPOLOGY_INFORMATION; at each call establishment the local RRM ask through RRM_ DECISION whether the call shall be established on the own RAT or an alternative RAT: this decision is communicated by the CRRM through CRRM_DECISION. Finally, each RRM may send the triggered NETWORK_LOAD_REPORT to drive the CRRM_DECISION of an Intersystem Handover based on the network load. It is evident, moreover, that no direct communication between the two RRMs is required; actually this approach allows to use this CRRM block connected to many different RAT (UMTS, WLAN, WIMAX, etc.).

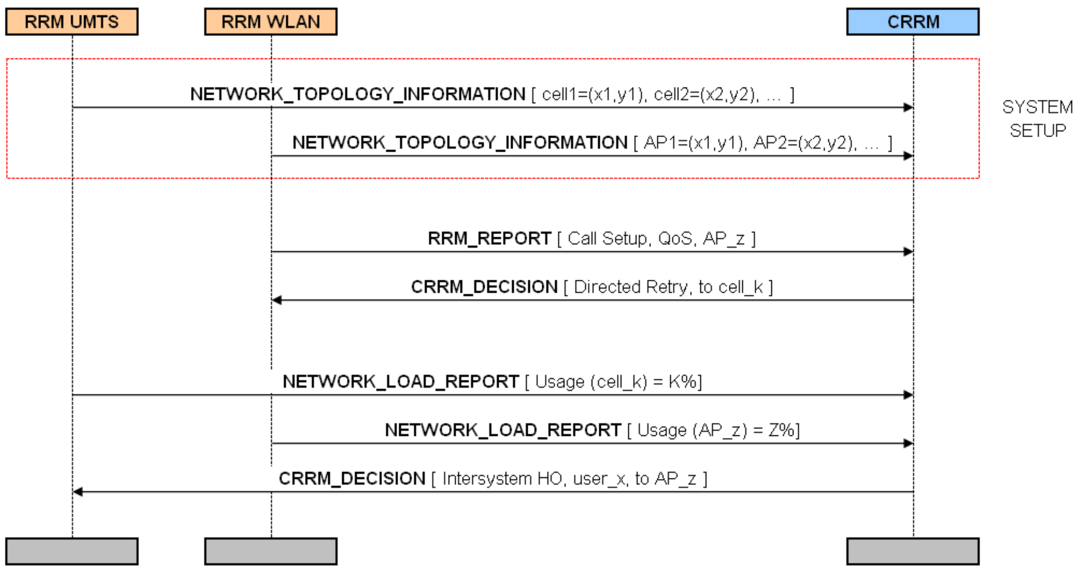


Figure 5.2: Flow diagram of interactions between CRRM and RRMs

In the next two sections we present the functions normally implemented in the local

RRMs and the functions we propose to implement in the CRRM.

5.3 Local RRM functions

In order to understand the impact introduced by the CRRM, it's appropriate to give a brief description of the principal local RRM functions implemented in the UMTS and the WLAN networks.

Main RRM functions of a UMTS network

- *Radio Bearer Translation:* to establish a radio access bearer between UE and Core Network (CN), this function maps the Radio Access Bearer (RAB) QoS parameters from the Core Network to the Radio Bearer (RB) parameters. This mapping has to be performed for each new incoming bearer request.

- *Radio Bearer Control:* This set of functions aims to optimize the usage of radio resources by adapting the amount of resources assigned to an UE depending on its traffic load. This is achieved by managing the bit rate adaptation of the radio bearers to the source bit rate and quality of service (QoS) requirements. The mapping has to take into account the actual system load as well as the actual bit rate and quality of service requirements of the considered radio bearer.

- *Power Control:* Since in UMTS all subscribers share the same frequency band, lowering the interference within the cell while maintaining transmission quality is of paramount importance. For this purpose power control adjusts the transmission power in uplink and downlink.

- *Handover Control*: For UEs moving from a cell to another, this function transfers connected calls to the new cell. The functionality is different for different scenarios: Intrafrequency Handover Control (hard or soft handover within the same frequency), Interfrequency Handover Control (handover between different radio frequency layers), Intersystem Handover Control (3G-2G).

- *Resource Allocation (Admission Control):* The admission control function decides whether a new radio link can be admitted in a particular cell based on the cell load and the available channelization codes.

- *Pre-Emption:* Pre-emption function allows the establishment of new higher-priority calls in the face of high load by degrading lower-priority bearers. It interacts with admission control.

- *Congestion Control:* Congestion control is a vital function, which detects an overload situation in a particular cell based on the load information from the Node Bs. It resolves congestion by invoking functions for reducing cell load, such as channel type switching, bit rate adaptation or bearer dropping.

- *Restriction Control:* This functions permits to restrict the available bit rates within a cell by restricting the allowed minimum spreading factor.

Main RRM functions of a WLAN network

- *Enhanced Distributed Channel Access (EDCA):* IEEE 802.11e allows service differentiation delivery.

- *Call Admission Control (CAC):* in order to ensure the QoS, a CAC algorithm may decide whether an incoming user canor cannot access the network.

5.4 CRRM Algorithm

In our platform, we defined three main algorithms for the Common Radio Resource Management UMTS-WLAN (see the interaction from CRRM to RRM "CRRM_DECISION" in section 5.2):

- CRRM "Service Based".

- CRRM "Coverage Based".

- CRRM "QoS based"

The three CRRM options aren't exclusive, in the section 5.8 we'll show results with different combination of CRRM algorithms. The final solution will be to use them all together.

5.4.1 CRRM "Service Based"

The algorithm of the CRRM, first of all, defines the RAT to be selected at the call setup for each type of service. Given that UMTS can serve data traffic at excellent rates with HSDPA technology, however the choice we made in this project is to preferably serve the data traffic requests coming from the high traffic hotspot area with the WLAN network because it can offer a higher bandwidth than UMTS. On the contrary, our choice for speech users is to preferably serve them with the UMTS network, because of the better speech quality UMTS can provide over its dedicated channels.

Bearer	**Preferred RAT**	**Retry**
Conversational	UMTS	WLAN (VoIP)
Best Effort	WLAN	UMTS

Table 5.1: Initial-RAT selection algorithm in the hotspot.

In SHINE, the CRRM Service Based policy can be easily changed, i.e. we could define that for all calls in the hotspot the preferred RAT is the WLAN.

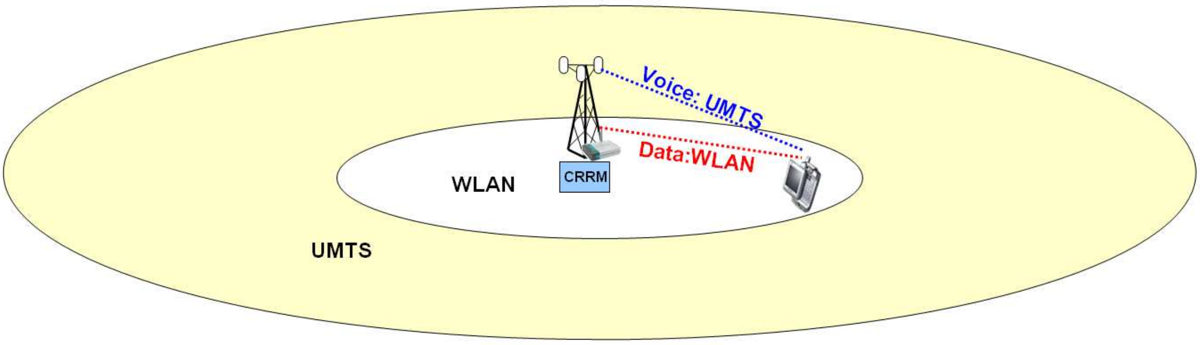


Figure 5.3: CRRM Service Based

Moreover, the CRRM allows to retry the call setup on an alternative RAT, in case of radio resource shortage, i.e. if a voice call is rejected in the UMTS network, the CRRM may order a "Directed Retry" in the WLAN network.

In figure 5.4, we show three possible scenarios for CRRM Service Based.

In case (A), the RRM UMTS informs the CRRM that a user requiring QoS_1 is attempting to establish the call on cell_k; if QoS corresponds to a speech call, the CRRM confirms that the call shall be established in UMTS.

In case (B), the RRM WLAN communicates that a user is attempting to setup a call, for example a speech call, when connected to the AP_z of WLAN; the CRRM Service Based, because of the choices in table 5.1, shall inform the RRM WLAN that a Directed Retry procedure to the UMTS cell_k has to be performed: automatically the RRM WLAN entity shall start a relocation to UMTS.

Finally, in case (C), the RRM UMTS attempts to establish a speech call, the CRRM confirms the call shall be established in UMTS; after that, the RRM UMTS realize there aren't enough radio resource on cell_k, therefore it communicates this Admission Control rejection to the CRRM, which eventually decides to attempts the voice call establishment in WLAN by sending to the RRM UMTS a CRRM_DECISION of Directed Retry to WLAN AP_z.

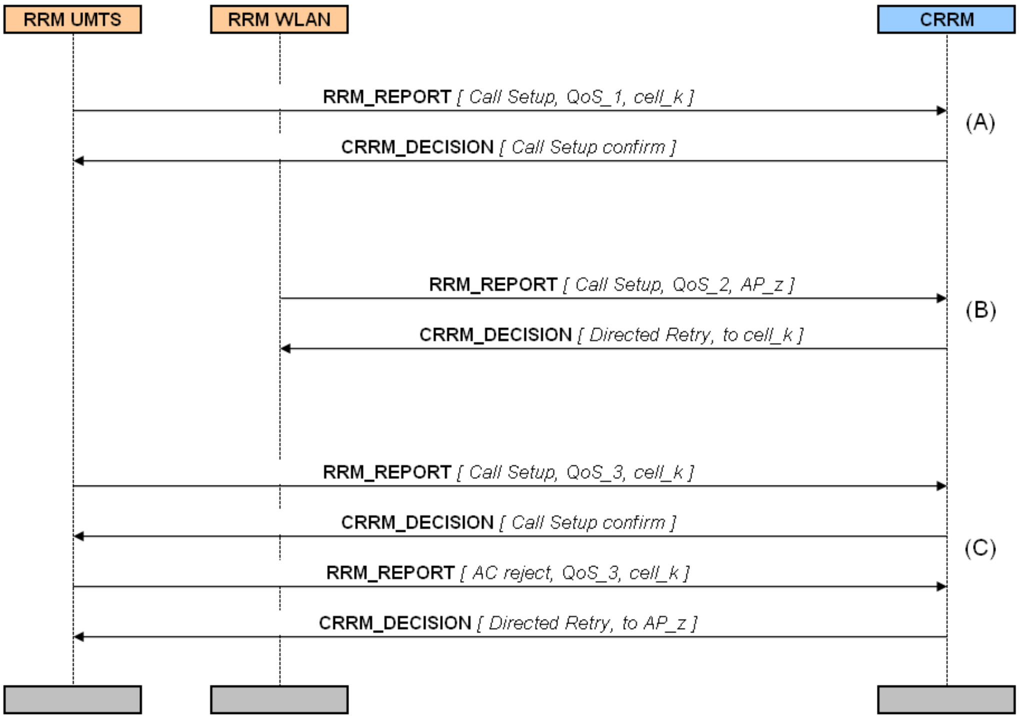


Figure 5.4: Flow diagram of CRRM Service Based

5.4.2 CRRM "Coverage Based"

Secondly, the UMTS-WLAN interworking is "coverage based": when a user connected to the WLAN AP exits from the WLAN coverage, without service interruption, through Intersystem Handover (see chapter 4), it is moved to the UMTS system.

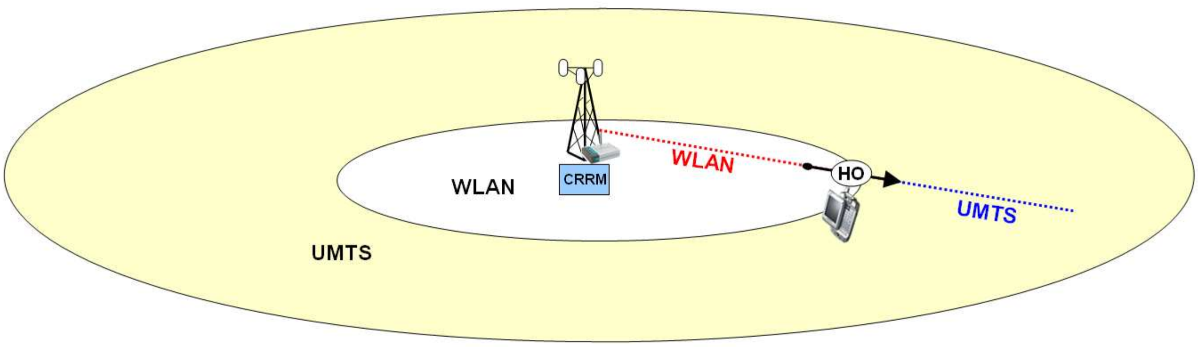


Figure 5.5: CRRM Coverage Based

If UMTS rejects the relocation (for example, because of Admission Control), the call is dropped and the relative counter is incremented: we decided in this case to avoid to maintain the call in the WLAN, because of the well-known negative effect introduced by WLAN stations far away from the AP.

5.4.3 CRRM "QoS Based"

Finally, the UMTS-WLAN interworking is "QoS based": together with the huge effort to develop the SHINE simulation environment, this part is one of the most important contribution of this work. The purpose of the CRRM QoS based is to serve the voice users in the RAT that currently, based on its load, best fits the QoS requirements and to avoid to congest a single RAT when there are many RAT available in the same area.

In section 5.2.2, we described how each local RRM shall send NETWORK_LOAD_REPORT to the CRRM. Here we explain how actually the CRRM uses these radio measurements coming from each RAT. First of all, based on the threshold mechanism, the CRRM locates each UMTS cell and each WLAN AP in one of the following three states:

- Unloaded state.
- Highly loaded state.
- Congested state.

The interaction with the basic CRRM Service Based is the following: the CRRM QoS based decides that each conversational call (that is the most demanding service regarding the packet delay and error rate) is setup on the less loaded RAT (in ascending order, unloaded - highly loaded - congested); in case the UMTS cell and the WLAN AP are in the same state, it is the CRRM Service Based that decides the preferred RAT as defined in table 5.1, that is the speech call is setup in UMTS.

The CRRM QoS based we defined specifies also that the best effort calls in the hotspot shall be always preferably served by the WLAN, that is for the best effort service, the CRRM continues to follow the policy in table 5.1. Nevertheless, the CRRM QoS based doesn't setup any call when the WLAN state is "congested": in this case, the call is established in UMTS.

Besides initial RAT reselection, the CRRM QoS based may command Intersystem Handover procedures: when the CRRM QoS based detects that a UMTS cell or a WLAN AP change the state from "highly loaded" to "congested", the CRRM checks if there is an alternative RAT in not congested state and in case this is true it starts to move users through standard 3GPP RANAP Relocation procedure [59] from the congested RAT to the alternative RAT (one handover every ΔT) till the critical condition is solved.

In figure 5.6 the scenario of interest for the study of the behavior of the CRRM QoS based: in a congested UMTS cell, the CRRM moves with Intersystem Handover some users in the hotspot from UMTS to WLAN. In figure 5.7, it's schematized the interaction between the RRMs and the CRRM: after a NETWORK_LOAD_REPORT with "congestion" measurement in UMTS cell_k, since the related WLAN AP_z is unloaded, the CRRM orders through CRRM_DECISION to handover a user to AP_z every ΔT. Once the UMTS RRM informs in NETWORK_LOAD_REPORT that the cell_k is not anymore congested, the critical situation is solved and the QoS of all the calls in the system (both UMTS and WLAN) has been maintained.

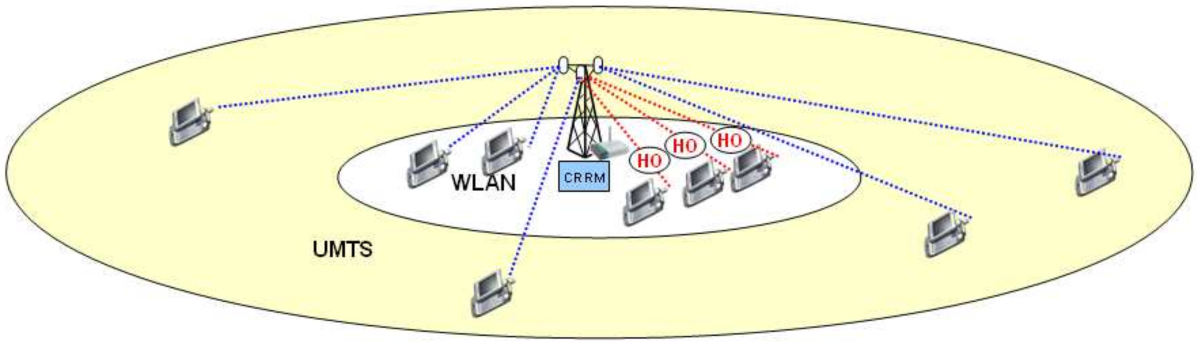


Figure 5.6: CRRM QoS Based

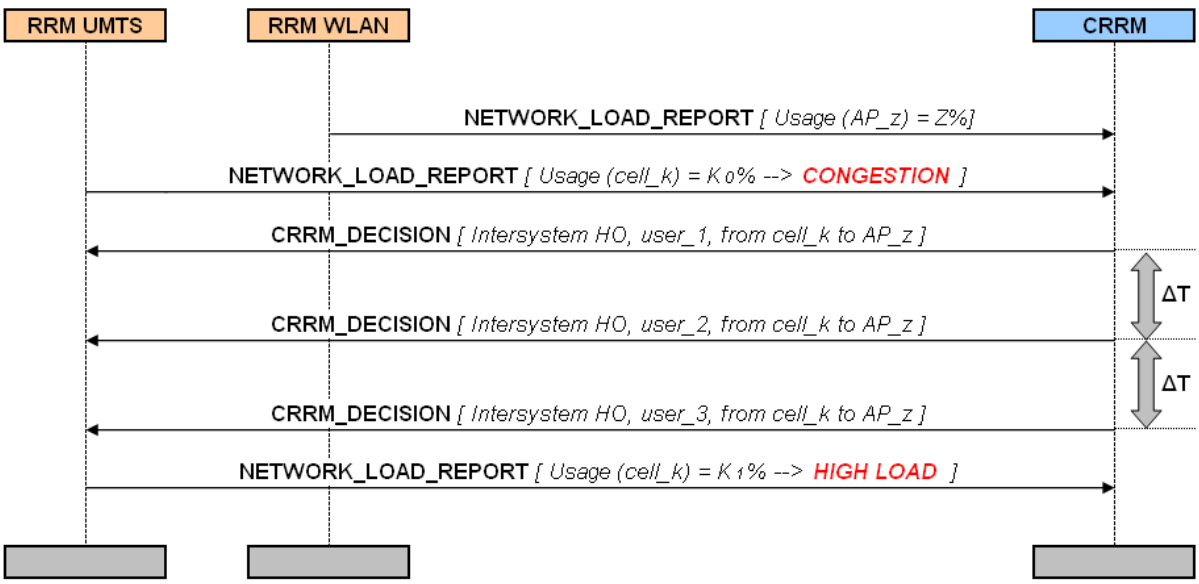


Figure 5.7: Flow diagram of CRRM QoS Based

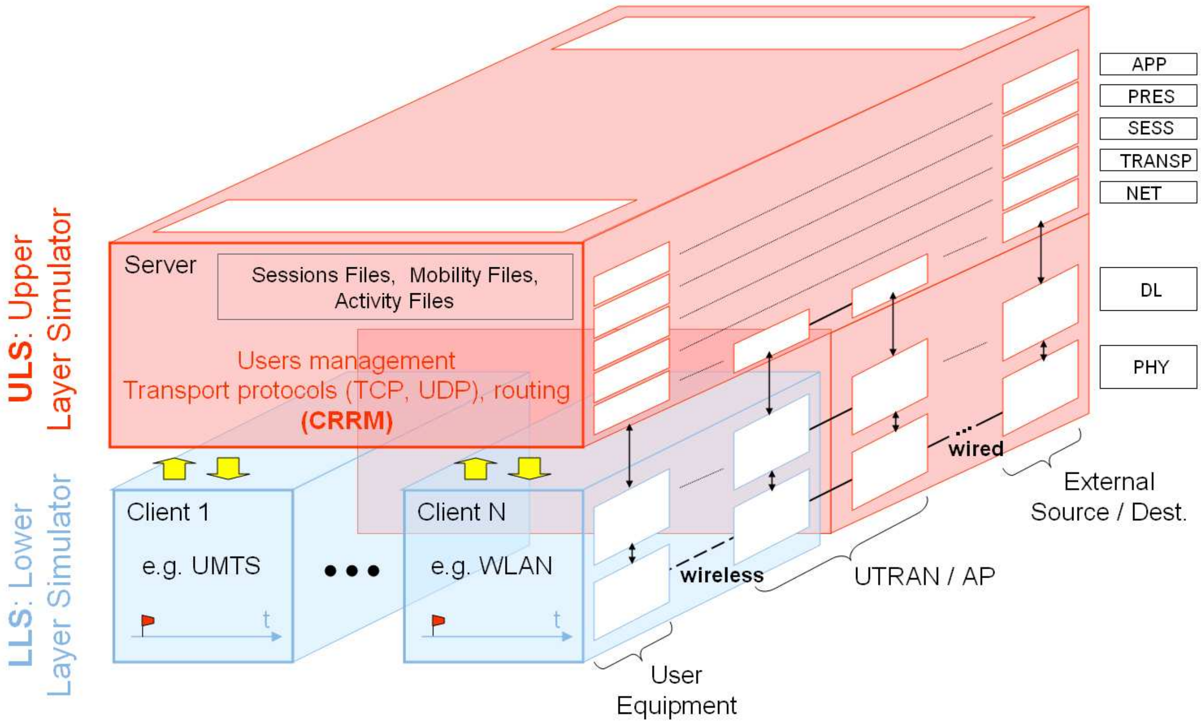


Figure 5.8: Simulation platform architecture.

5.5 Software simulation platform settings

The simulation platform that we developed and adopted for the study of Heterogeneous Networks is explained in Appendix A. In this subsections, a brief recall and some further details on the parameters adopted for the next numerical results are given. As a reminder, in figure 5.8 the block scheme of the platform is reported: the main concept is to divide the overall problem in two areas, simulated by one Upper Layer Simulator (red area) and two or more Lower Layer Simulators (blu areas). The Lower Layer Simulators used in this study are the radio access network simulators for UMTS and WLAN.

5.5.1 Upper Layers Simulator-ULS

The main tasks of ULS are hereafter recalled:

- it sets the starting instant of each new traffic session originated by users, according to the statistics of the traffic class it belongs to, as well as the users' positions within the investigated scenario;
- it generates the bit-flows up(down)loaded by each user in each traffic session according to the statistics of its class of traffic;

- it reproduces the transport protocol behavior: both UDP and New Reno TCP are implemented;
- it performs packet segmentation and reassembly;
- it executes all **CRRM functions**: it selects through which technology should each user be connected on the basis of user-defined rules and the available networks' information; it can also decide to reject a connection or to move it from a LLS to another (that is, from a given technology to another) at any time, thus simulating the network interworking;
- it finally collects all simulation outcomes and generates the outputs (throughput, packet delivery delays,...) from an end-to-end point of view.

In our simulation platform each LLS manages its own time axis; the ULS, for its part, communicates to the LLSs the next instant in which some event concerning the upper protocol levels happens (sessions begin, start of bit transfers, TCP timeouts, etc.). In this way, when the time counter of an LLS reaches that instant, the related LLS simulation stops and a "call" to the ULS is performed asking for the event-related information and providing to the ULS, at the same time, information on the lower protocol levels events. After the ULS replies, the simulation of the calling LLS resumes, the consequent actions (new session start, MAC level frame queuing,etc.) are executed and the instant of the next ULS event (TCP time-outs,etc.) is updated.

An LLS can perform a call to the ULS not only when the instant of the next ULS event has come but also whenever an LLS event which is of interest for the ULS takes place, such as, for instance, the correct transmission of a MAC level frame.

The above described stop-and-wait procedure is the basis for the coordination among LLSs, which is obviously needed when simulating interworking networks: in case an ULS event is of interest for more than one LLS, no ULS reply is granted to the calling LLSs until all the interested LLSs have stopped waiting for it, then the reply is issued on the basis of the LLSs reports provided to the ULS. It follows that although LLS executions can take place at different speeds (LLSs complexity could be different and they could be running on different PCs), the faster LLSs periodically stop and wait for the LLSs they are interworking with, thus re-synchronizing simulations.

5.5.2 UMTS LLS

The UMTS simulation tool (that is, the UMTS LLS) I developed reproduces the main characteristics of the TD-SCDMA radio interface reported in [47], Release 4.

UMTS Physical level Protocol

TD-SDCMA, which stands for Time Division Synchronous Code Division Multiple Access, combines an advanced TDMA/TDD system with a synchronous CDMA. This technology is characterized by a chip rate of 1.28 Mcps and a bandwidth of 1.6 MHz, allowing 3 carriers in a 5 MHz band. Channelization code is OVSF (Orthogonal Variable Spreading Factor) and a QPSK modulation scheme is adopted.

TD-SCDMA frame is constituted by 2 subframes of duration of 5ms: each subframe is subdivided in 7 time slots, which can be flexibly assigned either to several users or to a single user who may require multiple time slots.

Main features of our TD-SDCMA LLS are:

- power attenuation due to propagation was taken into account according to the Walfish-Ikegami model, hence the pathloss (PL) dependence on the propagation distance d is given by $PL(d) = K_1 + K_2 log10(d)$. Here we assumed $K_1 = 15.3$ and $K_2 = 37.6$;
- the shadowing is modelled by means of lognormal random variables with zero mean and an exponential time correlation function; the spread here assumed for the log-normal shadowing process is $\sigma = 5dB$;
- fast fading: different ITU channels are simulated (pedestrian, vehicular, indoor, both A and B) [54]; here the pedestrian channel model has been assumed with an user speed of 3.5Km/h;
- both fast closed loop power control (rate 200 Hz) and slow outer loop power control are implemented; each bearer service is characterized by a proper value of initial target signal-to interference ratio and different Transmitted Power Command step sizes (Δ) can be selected;
- the maximum transmitted power for each cell is 10W;
- both multi-slot and multi-code combinations are supported.

UMTS Layer 2 Protocols

Basic procedures are defined for the MAC simulator, providing a logical channel service, and for the RLC simulator, providing a Radio Bearer service.

MAC main task is the priority-based management of the shared resource among users; RLC provides, on the other hand, segmentation functionality and different reliability modes: for CS speech connections RLC transparent mode is used and the quality of service is estimated by averaging bit error rate measurements over long periods (here we assumed 1.0 s); in case of PS sessions the simulator evaluates for each transport block belonging to a data packet whether it is correctly received or not through the value of the related measured block error rate (see chapter 3).

For PS calls a pair of RLC instances in acknowledge mode is used, providing a reliable radio bearer service which includes error correction by means of automatic retransmission of the transport block in case of reception failure.

At top of Layer 2, Radio Resource Control (RRC) block implements Call Admission Control: a new radio link is successfully setup provided that the necessary OVSF codes are available, the estimated interference is less than a given threshold and the initial power required in Node-B is available.

The simulator decides to drop a call using a "leaky bucket" algorithm, which compares the average measured bit error rate, BER, with the threshold $BER_{drop} = 2 \cdot 10^{-2}$: if $BER > BER_{drop} \rightarrow$ increase *counter* (+1), otherwise $\rightarrow$ decrease *counter* (-2) ; if *counter* ≥ 4, the call is dropped

Different classes of traffic are supported:

- CS Adaptive Multi Rate -AMR- speech at 12.2 Kbps;
- PS Best Effort: three Unconstrained Delay Data bearer services at 64/64 Kbps, 64/144 Kbps and 64/384 Kbps are considered, where x/y Kbps stands for x Kbps for the uplink and y Kbps for the downlink.

In the following, when dealing with UMTS data services, we always considered the 64/384 Kbps bearer.

5.5.3 WLAN LLS

As hereafter detailed, the WLAN network simulator (that is, the WLAN LLS) carefully reproduces the IEEE802.11a [13, 10] PHY level behavior as well as the MAC protocol of the IEEE 802.11e [31] technology. Moreover, a Call Admission Control (CAC) strategy has been implemented on top of IEEE802.11e MAC layer in order to prevent WLAN saturation.

WLAN Physical level Protocol

As for the PHY issues, let us recall that IEEE802.11a is based, at the Physical level, on eight operating modes adopting in the 5 GHz band the Orthogonal Frequency Division Multiplex (OFDM) technique to counteract the effects of frequency selective fading [10]. Each mode is characterized by a different combination of modulation scheme and code rate (R) of the punctured convolutional code adopted to correct transmission errors. Mode 1 (the slowest one), for instance, adopts a BPSK-based OFDM scheme with $R = \frac{1}{2}$, while Mode 8 (the fastest one) adopts a 64QAM-based OFDM scheme with $R = \frac{3}{4}$.

All IEEE 802.11a PHY level aspects (propagation, modulation, channel coding, ...) have been carefully taken into account by the WLAN LLS, in particular:

- according to ETSI recommendation, the multipath channel is generated in a time and frequency correlated way following [11]. In this work we considered the channel model A (see [11]) that corresponds to an indoor environment with 18 Rayleigh distributed paths. The time-correlated channel variations are taken into account considering a relative user speed of 3 km/h for the Clarke [12] Doppler spectrum of each path;

- the Auto Rate Fallback (ARF) [32] link adaptation algorithm is assumed to select the proper transmission mode (i.e., the combination of modulation scheme and coding rate) following channel variations; MAC level frames are discarded after 7 consecutive failed transmissions; hard decision convolutional decoding is assumed;

- according to the technical specifications of most of commercially available WLAN devices, here we assumed 19dBm for EIRP, 2dB for the antennas' gains and 12dB for the receiver noise figure. The signal power has been assumed decaying with the fourth power of the distance.

WLAN Medium Access Control Protocols

As for the MAC protocol of IEEE 802.11e, here we considered the contention-based access mode (Distributed Coordinated Function DCF for version 'a', Enhanced DCF EDCF for version 'e', see [13, 31] for details) which is based on the Carrier Sensing Medium Access with Collision Avoidance (CSMA-CA) strategy; the hidden terminal problem is supposed to be negligible, hence the two-ways handshake procedure [13] is assumed in the following.

Let us recall that IEEE802.11e differs from IEEE802.11a, as well as from IEEE802.11b/g, since it introduces the concept of Quality of Service in the frame delivery procedure at the MAC level.

The idea at the basis of the IEEE802.11e MAC is to manage traffic flows with different throughput/delay requirements by means of different queues, which are given different probabilities to gain the access to the channel: the more the requirements of the traffic assigned to a queue are stringent, the higher is the probability its queue is given to gain the access to the channel. IEEE802.11a MAC layer, on the contrary, does not differentiate among different traffic flows.

Let us recall, now, that IEEE802.11e specifications [31] define only the MAC level strategies, which can be combined with anyone of the IEEE802.11a, IEEE802.11b or IEEE802.11g PHY layers; here, when dealing with IEEE 802.11e we mean the combination of IEEE802.11e MAC level and IEEE802.11a PHY level.

When considering IEEE802.11e, the different traffic classes are assigned different priorities: in table 5.2 we reported the values of IEEE802.11e MAC protocol parameters adopted in our simulation to differentiate the access probability of each queue (see [31] for details on their meaning). According to the values of table 5.2 the VoIP traffic is given the highest priority (that is, the highest probability to gain the access to the channel) while FTP traffic is given the lowest priority.

	Voice			**Web browsing**			**Ftp**		
Parameter	AIFS	CWmin	m	AIFS	CWmin	m	AIFS	CWmin	m
value	1	8	2	4	16	6	4	32	6

Table 5.2: Set of parameters adopted for IEEE 802.11e

Call Admission Control Protocol

In order to prevent the saturation of WLAN resources, that would lead to a poor QoS level perceived by final users, here we implemented on the IEEE802.11a/e simulator a simple but efficient CAC solution, based on a centralized evaluation of the current network congestion level. Let recall that this is done by monitoring the channel occupation through the assessment of the *Channel Occupation Rate* parameter, C_O [33], defined as the ratio between the amount of time the medium is busy, T_B, and the related observation time ΔT:

$$C_O = \frac{T_B}{\Delta T}.$$

The evaluation of the channel occupation rate is particularly simple to be implemented in existing Access Points, owing to their carrier sensing capability. The AP simply adds the busy medium sensed time T_{SB} to its transmission time T_{AP}, as follows:

$$T_B = T_{AP} + T_{SB}.$$

In order to improve the accuracy of the estimation of T_B, also the mandatory idle times DIFS, SIFS [13] and AIFS [31] are considered in the assessment of T_{AP} and T_{SB}: in particular, a DIFS period is considered for each uncorrect transmission (i.e. each packet not acknowledged) and both DIFS (or the appropriate AIFS, for IEEE802.11e) and SIFS periods are considered for each correct transmission; please note that as long as the transmission is successfully performed, the AP automatically knows the correspondent class of traffic and thus its AIFS interval.

For high values of C_O the frame collision probability increases and consequently the end to end delay and the throughput experienced by users get worse. The CAC algorithm has to keep the Channel Occupation under the saturation point, hereafter denoted as Congestion Threshold (C_T). When C_T is exceeded any admission is denied, till C_O goes below a Decongestion Threshold (D_T).

5.6 Performance measurements

An integrated UMTS-WLAN system needs also an integrated definition of Quality of Service requirements and merit figures. The main idea is that the end-to-end user experience should be independent from the radio technology - UMTS or WLAN - chosen by the CRRM, i.e. the integrated system wouldn't be attractive if in UMTS the speech quality were satisfactory whereas in the hotspot the VoIP over WLAN were highly affected by jitter and packet loss.

The merit figures shall provide the performance of the overall UMTS&WLAN system, for example a call established in UMTS, moved to WLAN because of congestion, and moved again to UMTS because of coverage shall be considered as a single call, and not as three different calls. The only entity that can evaluate the global merit figures is the CRRM, because it's the only network element which can keep track of each call in the heterogeneous network.

Quality of Service requirements

The quality requirements of the services provided in the scenario we will simulate are the followings:

- *Speech traffic.* When served by UMTS, the speech traffic is a circuit switched bidirectional flow. A speech user is assumed to be satisfied if its call is neither blocked nor dropped, and the overall outage time is lower than 5% of the call duration in each direction (an outage event for CS voice calls occurs when the average value of the bit error rate exceeds the threshold $BER_{out} = 10^{-2}$).

  When served by WLAN AP, the speech traffic is a VoIP flow; the user in this case is satisfied if 97% of packets are received in less than 0.15 sec in each direction [46]; following [46], since the voice/VoIP conversion adopting a G.729 codec in the H.323 gateway takes around 120 ms (including coding, decoding, bufferization delay, etc.) and assuming that the circuit switched Core Network introduces a negligible delay, the maximum tolerable delay introduced by the WLAN is thus 30 msec.

It is clear the parameters used to measure the quality of speech calls are different in UMTS and WLAN; in order to measure the outage of calls that were served by UMTS and WLAN in different periods, we decided to count the number of time intervals (for example, slot duration of 1 second) in which the user experienced outage in UMTS or in WLAN, then to divide this number by the total duration of the call and finally to compare this ratio with a 5% threshold.

- *Web browsing traffic.* Characterized by bursts (packet calls) of small application-level packets followed by inactivity periods (reading time). A web browsing user is satisfied if 90% of packets are received in less than 5 sec;

- *FTP traffic.* Bursty traffic characterized by requests of huge application-level packets followed by inactivity periods (reading time). An FTP user is satisfied if the average experienced throughput is higher than 800 kbit/s.

Merit figures

The merit figures we will consider in the numerical results section 5.8 cover the different phases of Quality of Service [26] aspects during service use from the customer's point of view. They are here defined:

- *Call Setup success rate (CSSR)*: it regards the **Service Access**, that is if the customer wants to use a service, the network operator of an integrated UMTS-WLAN network should provide him as much as possible access to the service. *CSSR* describes the ratio between the calls successfully connected to the system (either UMTS or WLAN) and the overall number of call attempts:

$$CSSR = \frac{N_{attempt} - N_{block}}{N_{attempt}} \tag{5.1}$$

  where $N_{attempt}$ is the number of call attempts and N_{block} is the number of blocked calls (for example, for Admission Control). Note that if a call attempt in the hotspot fails in both UMTS and WLAN, this is counted only once by the CRRM: the perspective is to provide a performance indication from the user point of view; each single system may internally count this failure for its own statistics, but at the end the CRRM will count only one blocked call as only one user "pushed the button" to start a call.

- *Drop Call rate (DCR)*: it regards the **Service Retainability**, that is it describes the termination of services in accordance with or against the will of the user. *DCR* is the ratio between the abnormal releases (i.e. dropped calls for poor radio conditions or for congestion) and the overall number of call releases (that is both the normal and the abnormal releases):

$$DCR = \frac{N_{drop}}{N_{release}} \tag{5.2}$$

where N_{drop} is the number of dropped calls and $N_{release}$ in the number of call releases.

- *Outage rate (OutR)*: only for speech calls, it regards the **Service Integrity**, that is the Quality of Service during service use. *OutR* describes the ratio between the number of calls which, although normally released, perceived an unacceptable outage time lasting more than 5% of the duration of the call, and the overall number of call releases (that is both the normal and the abnormal releases):

$$OutR = \frac{N_{outage}}{N_{release}} \tag{5.3}$$

where N_{outage} is the number of calls with unacceptable outage time.

- *Satisfaction rate (SatR)*: only for speech calls, it summarize the degree of voice user satisfaction, by considering altogether Service Access, Service Retainability and Service Integrity. A call is considered satisfied if it isn't blocked, nor dropped, nor it didn't feel outage:

$$\begin{aligned} SatR &= CSSR \cdot (1 - (DCR + OutR)) = & (5.4) \\ &= \frac{N_{attempt} - N_{block}}{N_{attempt}} \cdot \frac{N_{release} - (N_{drop} + N_{outage})}{N_{release}} = & (5.5) \\ &= \frac{N_{attempt} - N_{block}}{N_{attempt}} \cdot \frac{N_{sat}}{N_{release}} & (5.6) \end{aligned}$$

where N_{sat} is the number of satisfied users, that is the calls which were successfully connected to the network and were normally released without perceiving significant outage during the call.

In case of a stationary system, in which $N_{attempt} = N_{block} + N_{release}$, from equation 5.6 it is easy to obtain an easier formula for the satisfaction rate SatR:

$$SatR = \frac{N_{sat}}{N_{attempt}} \tag{5.7}$$

Figure 5.9: Investigated scenario: WLAN APs in hotspot of high density traffic

- *Average perceived throughput*: only for best effort calls, it is the average value of the application level throughput perceived by users; the application level throughput is defined as the ratio between the amount of bits of each application-level 'packet' and the time elapsed from the instant in which the TCP packet containing the first fragment of the application-level 'packet' reached the head of the TCP transmission queue and the instant in which the TCP packet containing the last fragment of the same application-level 'packet' was completely successfully received;

- *Average delivery delay*: it is the average time interval occurring from the generation of an application-level 'packet' in the transmitter to its successful reception, thus including in the delay calculation the time spent in the TCP and MAC levels queues.

5.7 Traffic scenario

The architecture for an integrated UMTS-WLAN system we've proposed in chapter 4 and the CRRM project we've explained in the previous sections 5.2-5.4, cover the general scenario of a system in which both UMTS and WLAN radio access network are deployed and there is an interest to exploit the complementary characteristics provided by the two

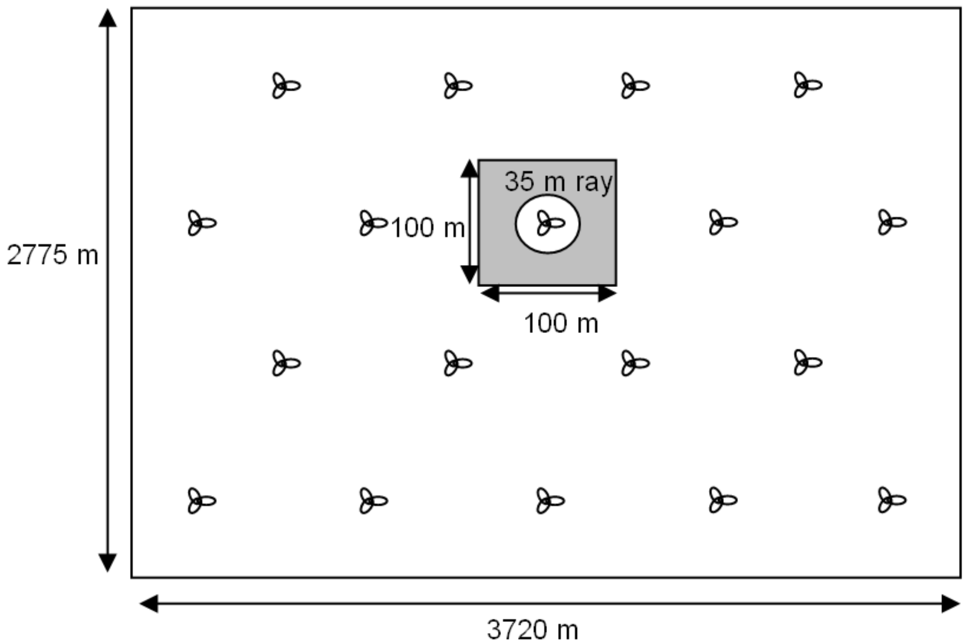


Figure 5.10: Simulated scenario. The grey square indicates the area considered for numerical results

technologies.

We haven't yet mentioned the location or the number of NodeBs and the number of WLAN APs: in figure 5.9 an example, in which 2 WLAN APs cover high traffic areas (red circles), inside a global UMTS coverage. In this section we define the network topology and the traffic scenario we used for the simulation with our platform SHINE. Actually, in order to carry out significant investigations on the benefits of networks integration and to derive meaningful conclusions, particular attention has to be paid not only to the accuracy of the simulation platform but also to the realistic modelling of all aspects characterizing the operating conditions of real networks, such as the scenario, the variety of services requested by users and the statistics and geographical distribution of traffic flows.

5.7.1 Network topology

The scenario investigated is depicted in figure 5.10 and consists of 18 UMTS Nodes-B with tri-sectorial antennas over an area of 3720x2775 m^2 which also includes a WLAN hot spot. The WLAN AP is assumed co-located with a Node-B (one of the central ones in figure 5.10) and is equipped with a dipole antenna offering an omnidirectional coverage in the horizontal plane; to avoid border effects all merit figures investigated in the numerical results section refer to an area of 100x100 m^2 centered in the AP (see figure 5.10).

5.7.2 Traffic distribution

In order to reproduce the variety of services provided in a real network, here we assumed that users could perform voice calls as well as web browsing and file transfer sessions, thus generating three different kinds of traffic whose statistical description have been carefully reproduced.

Regarding the geographical distribution of traffic, here we considered the superimposition of a *background traffic* generated by users uniformly distributed in the whole 3720x2775 m^2 region, and of a *hot spot traffic* generated by users uniformly distributed in a circular area (hereafter denoted as the *hot spot region*) centered in the AP, with a radius of $35m$, which could represent a highly crowded area where the WLAN hot spot is deployed (see figure 5.10):

- *Background traffic*: it is generated by "background users" that are assumed to perform voice calls or web-browsing sessions in the whole area covered by UMTS.
- *Hot spot traffic*: it is the further traffic contribution which is added to the "background traffic" only in the hot spot region and it is constituted by a huge additional amount of voice, web-browsing and file transfer (hereafter, FTP) traffics.

The straightforward consequence of the above reported assumptions is that the hot spot region is characterized by a higher user density than the surrounding region: this is typical of crowded areas such as, for instance, an airport gate or a shopping mall where, actually, the realization of a hot spot is envisioned.

The above reported characteristics are summarized in the first two columns of table 5.3, in the third one the mobility type is reported (we assumed only the voice users enters/exits from the Hot spot traffic), whereas in the fourth column the average call arrival rate is reported.

Traffic	**Area**	**Mobility**	**Arrival rate**
Voice$^{backgr.}$	whole scenario	Pedestrian/Vehicular [54]	2.7 calls/s
Web br.$^{backgr.}$	whole scenario	Static	1.08 sessions/s
Voice^{HS}	hot spot	Static	varying: $k \cdot f_a(t)$
Web br.^{HS}	hot spot	Static	0.02 sessions/s
FTP^{HS}	hot spot	Static	0.012 sessions/s

Table 5.3: Traffic distribution and arrival rates

We assumed the additional average voice call arrival rate in the hot spot ($Voice^{HS}$) is a variable quantity depending on the time, $f_a(t)$, in order to simulate the behavior of a hot spot of traffic in which the density of users change dynamically: in figure 5.11, it is shown that in 4800 seconds of simulation there are two peaks of traffic of the same

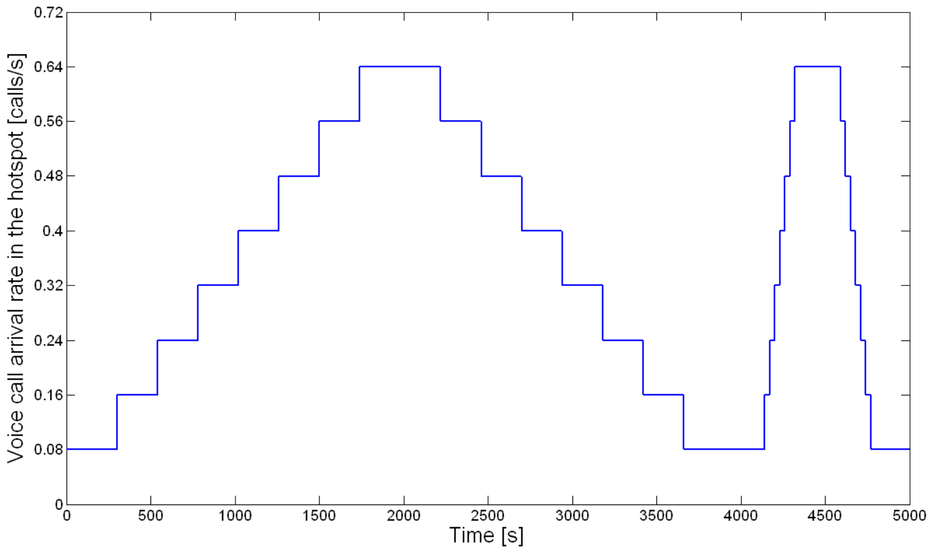


Figure 5.11: $f_a(t)$: voice call arrival rate Voice^{HS} in the hotspot

maximum value (0.64 calls/s, corresponding to 76.8 Erlang, based on traffic parameters of table 5.4), but with different variation step. In the numerical results section, the merit figures are evaluated as a function of k, that is the multiplier factor of $f_a(t)$ in table 5.3: for each run of simulation, the parameter k will assume the values 0.125, 0.25, 0.5, 0.75 and 1.

For the sake of clarity we finally summarize in table 5.4 the traffic categories we considered (first column), their characteristics (second column), the references where the models of each traffic class were taken (third column) and the satisfaction thresholds (fourth column).

Class	**Characteristics**	**Refer to**	**Satisfaction thresholds**
CS Voice	Poissonian duration, 120 sec in average		Natural conclusion with outage lower than 5% in each direction
VoIP	CBR traffic, 20 bytes packets with rate 8 kbit/s per direction	[46] (codec G.729)	97% of packets received in less than 0.03 sec (wireless link) in each direction
Web B.	1 to 8 packet calls of 70 kbytes in average, divided into 1 to 30 packets each; 60 sec average reading time	[55] with $\alpha = 0.6$, average reading time halved	90% of packets received in less than 5 sec
FTP	1 to 6 packet calls of 500 kbytes in average; 180 sec average reading time	Each packet call follows [34], with $\mu = 13.06$ and $\sigma = 0.12$	Average throughput of 800 kbit/s

Table 5.4: Adopted traffic classes: parameters and requirements for satisfaction

5.8 Numerical results

Given that we defined the CRRM strategy in sections 5.2-5.4 and a realistic traffic scenario in the previous section, here the system performance of such an integrated UMTS-WLAN network is finally investigated through our simulation platform SHINE; the main simulation setting were described in section 5.5. Actually, the final scope of this work is to evaluate the possible benefit of a CRRM approach when in the hotspot covered by WLAN AP and UMTS there is a high distribution of voice call users that cannot be completely served by UMTS only.

In the figures 5.12-5.16, the impact of different levels of integration between UMTS and WLAN on Quality of Service is investigated in the 100x100 squared meters area around the location of the WLAN AP. The results are provided for four different CRRM functions, of progressive complexity:

- **No CRRM**: the voice calls in the hotspot are served only by UMTS: during the peaks of traffic, the calls that are not admitted to the system won't be redirected to WLAN. This is the basic situation when no interworking between UMTS and WLAN is realized.

- **CRRM Service Based** (see subsection 5.4.1): the voice calls in the hotspot are preferably served by UMTS, but in case of resource shortage, call setup is redirected to WLAN. Moreover CRRM Coverage Based (see subsection 5.4.2) is executed: users exiting from WLAN hotspot do Intersystem handover to UMTS.

- **CRRM QoS Based - 50%** (see subsection 5.4.3): same algorithm used for previous point *(CRRM Service Based)*; moreover, the CRRM QoS based algorithm is enabled too. Once the UMTS system reports a load of 50% (see subsection 5.2.2 for *Network load report*) the congestion of UMTS is declared. This is the most advanced proposed CRRM option, allowing users to be served by the best system in terms of congestion of the radio interface.

- **CRRM QoS Based - 85%**: same as previous point, but with higher threshold - 85% - for UMTS congestion detection.

In figure 5.12, it's shown the **Satisfaction Rate** ($SatR$) of voice users in the grey square of figure 5.10, including the area covered by the WLAN AP; it's evaluated the performance when varying the multiplier factor k of the traffic function $f_a(t)$ (see Voice^{HS} traffic definition in table 5.3). The worst result for $SatR$ is obtained when no UMTS-WLAN integration is established *(No CRRM)*. The first substantial improvement is

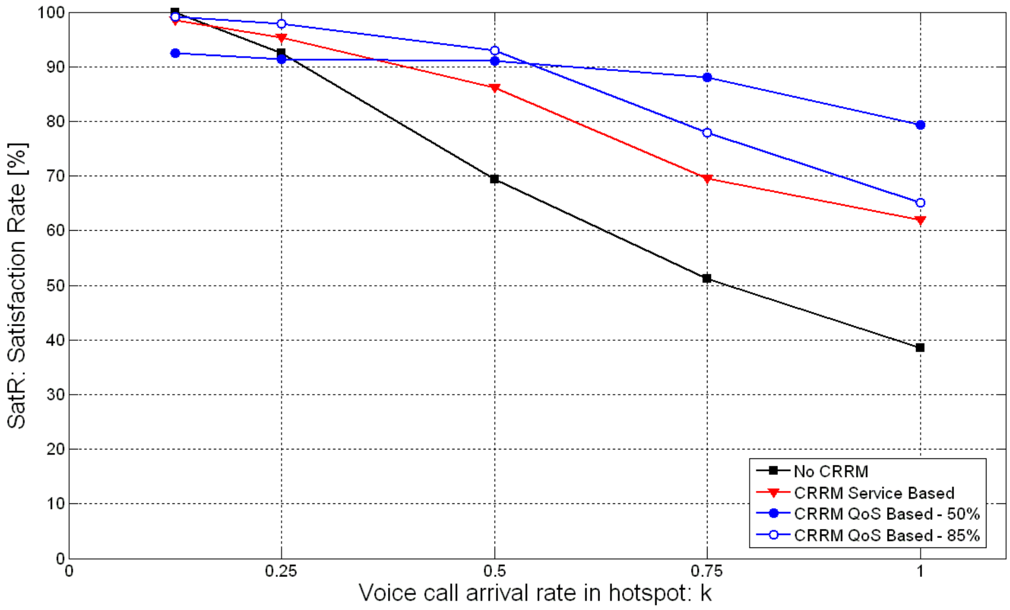


Figure 5.12: Voice QoS in the investigated 100x100 m^2 area: users' Satisfaction Rate *SatR*

achieved with the *CRRM Service Based*: during the peaks of traffic, the voice calls rejected by UMTS maybe redirected to WLAN, ensuring an increase of the number of voice calls that can be managed by the system, as could be expected since without UMTS-WLAN integration the available bandwidth is clearly lower. For $k = 0.5$ (corresponding to traffic peaks of 38.4 Erlang), *SatR* increases from 70% to around 85%, whereas for the higher value $k = 1$ (corresponding to traffic peaks of 76.8 Erlang) the Satisfaction Rate increases from 40% to 60%.

The next step in the performance improvement is obtained thanks to the full CRRM algorithm implementation, both Service and QoS based (for convenience, here we denote this option only with the label *CRRM QoS Based*): many parameterizations have been tried, in this analysis two of the most interesting ones from the results point of view have been reported. By using two different thresholds at 50% and 85% for Network Load report of congestion in UMTS, the performance of the *CRRM QoS Based* in terms of *SatR* improves the results of the *CRRM Service Based*.

For different ranges of traffic in the hotspot (k), a different parameterization of *CRRM Service Based* improves the Satisfaction Rate: with $k < 0.5$, the *CRRM QoS Based - 85%* option provides the best performance, whereas for $k > 0.5$, the *CRRM QoS Based - 50%* option outperforms both the *CRRM QoS Based* and the *CRRM QoS Based - 85%*. In

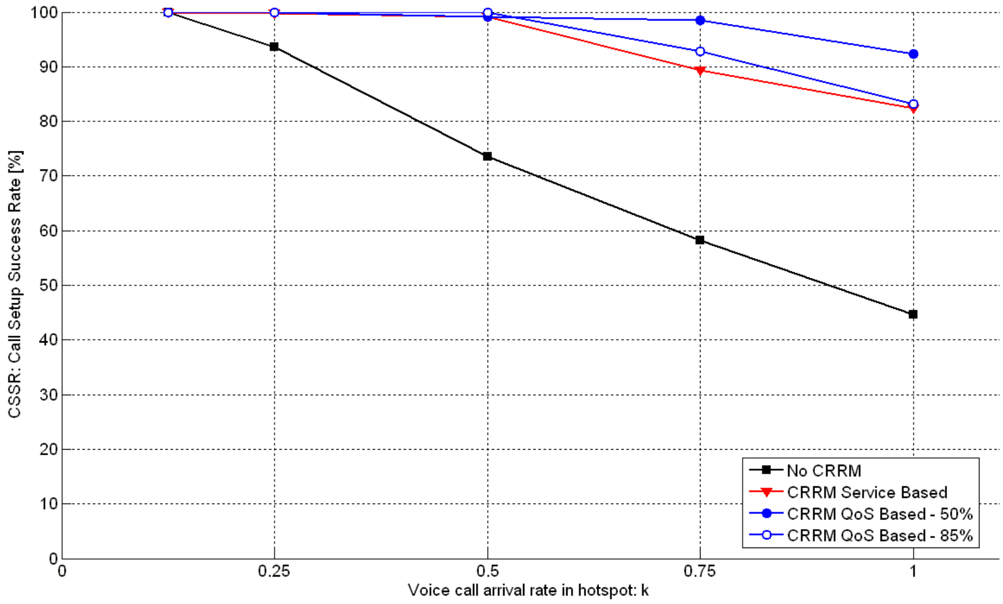


Figure 5.13: Speech Service Access QoS in the investigated 100x100 m^2 area: users' Call Setup Success Rate $CSSR$

particular, when $k = 1$, the Satisfaction Rate is about 80% with the *CRRM QoS Based - 50%*: this result is the double compared to the performance of the *No CRRM* option, as could be expected in general terms, but it's also remarkable this *CRRM QoS Based* option satisfies about more 20% users than the simpler *CRRM Service Based* option.

A single *CRRM QoS Based* algorithm changing dynamically its congestion thresholds based on the current traffic (k) will be evaluated in future studies, in order to determine if it is possible to achieve or even improve with a single algorithm parameterization the results provided by the envelope of the curves obtained adopting the two proposed parameterizations (CRRM QoS based 50% and 85%).

Let us recall that the merit figure $SatR$ is a Performance Measurement providing the result of the combination of several phase of the service: access, retainability and integrity. Now we separately study each of these aspects to investigate the root causes of the trends of the Satisfaction Rate.

In figure 5.13, the results of **Call Setup Success Rate** ($CSSR$) of voice users in the hotspot are presented. It can be observed that the system without interworking UMTS-WLAN *(No CRRM)*, starts to block voice users for $k = 0.25$: from this point on, the service access can be improved only with the CRRM. At $k = 0.5$, both the *CRRM Service Based* and the *CRRM QoS Based* provide $CSSR > 99\%$, that is this integrated network would satisfy an ordinary Service Level Agreement (i.e. $CSSR > 96\%$); this is due to the fact that the calls blocked in UMTS can be served by WLAN.

In the highest simulated traffic scenario, $k = 1$, apart from giving service access to almost 50% additional users compared to the *No CRRM* scenario (93% vs 45%), it is notable that the *CRRM QoS Based - 50%* provides service to more 10% users compared to the *CRRM Service Based* algorithm. At first, this improvement of the *CRRM QoS Based* compared to the *CRRM Service Based* might surprise since both these two CRRM options have at their disposal the same amount of radio resources of UMTS and WLAN, therefore it could be generally expected that the *CRRM QoS Based* should impact only on service retainability and integrity as will be explained later, and not on service access. From these simulation results, on the contrary, we have realized that the strategy of *CRRM QoS Based* to establish and move calls on the best system in terms of radio interface load leads to a better radio resource management and at last to provide access in the system to a larger number of users: instead of exhausting the UMTS resources and then forwarding the next call setups to WLAN (*CRRM Service Based* algorithm), it's preferable to serve calls in the less congested system.

Moreover, in the *CRRM QoS Based* configuration, once the UMTS cell in the hotspot reports a High Load state, the CRRM starts to move calls in the hotspot from UMTS to WLAN: this procedure allows the UMTS RRM to free resources for next incoming calls that cannot be served by WLAN, for example, calls established just outside the area covered by the AP WLAN.

In figure 5.14, the performance in terms of **Call Drop Rate** (DCR) of voice users in the hotspot is shown. The basic configuration with *No CRRM* presents the worst results: due to the high traffic load concentrated in the hotspot area, the service retainability is reduced; calls can be accepted by the UMTS RRM, but because of peaks of uplink interference, calls maybe abnormally terminated. When $k = 0.5$, the performance of *No CRRM* is similar to the result of *CRRM Service Based* (DCR about 4.5%), even if the latter can use the radio resources of both UMTS and WLAN; nevertheless, even if the level of service retainability is similar, figure 5.13 at $k = 0.5$ shows that with *No CRRM* 25% less users accessed the system, therefore the system load is reduced, thus affecting the comparison in terms of service retainability.

At $k = 0.5$, the *CRRM QoS Based* (both 50% and 85% options) halves the Drop Call Rate to 2.4%, a value that can be acceptable for a Service Level Agreement; in this case the comparison with the DCR of the *CRRM Service Based* is relevant since both of them offer the same voice traffic quantity ($CSSR > 99\%$). This clear improvement is provided by the combined management of the load levels of the two radio interfaces UMTS and

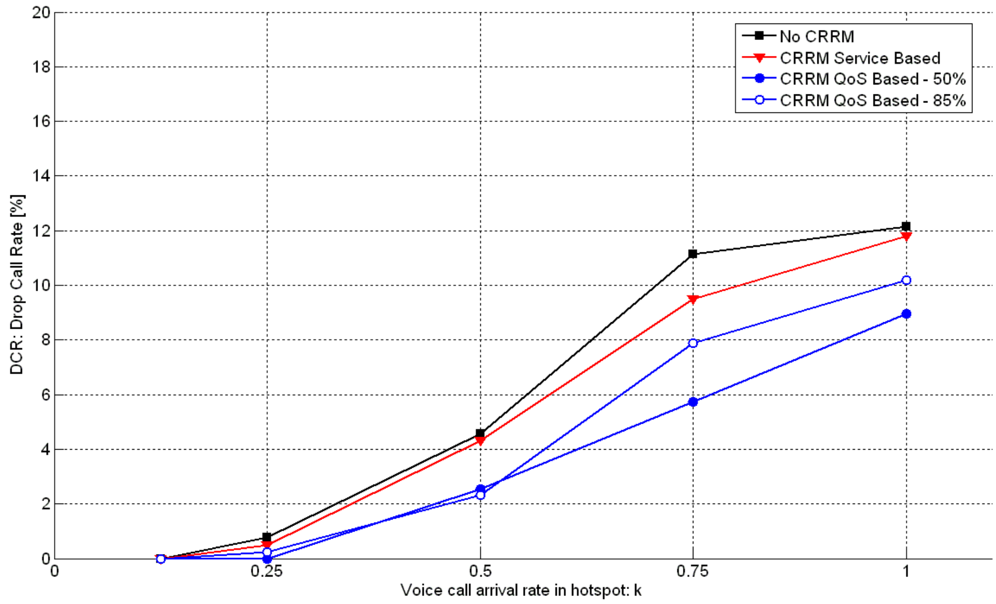


Figure 5.14: Speech Service Retainability QoS in the investigated 100x100 m^2 area: users' Drop Call Rate DCR

WLAN: when one system achieves the congestion level, before any call is abnormally terminated, the *CRRM QoS Based* attempts to move voice calls to the alternative radio access technology, if that one isn't congested.

Moreover, in the *CRRM QoS Based* configuration, since its general strategy is to avoid to use all UMTS resources until calls start to be blocked, the voice call users exiting from the area covered by the WLAN AP will likely perform a successful Intersystem Handover to UMTS; on the contrary, with the *CRRM Service Based* only, the possibility to exit from the WLAN area and then to not find enough radio resources in UMTS and therefore being dropped is higher.

In figure 5.15, the results of **Outage Rate** ($OutR$) of voice users in the hotspot are presented. The service integrity provided with *No CRRM* shows the best results, but as explained above, this basic option offers a lower voice traffic, therefore a straight comparison is tricky. Only in the most unloaded scenario ($k = 0.125$) $CSSR = 100\%$ for all four curves (see figure 5.13), therefore a reasonable analysis can be performed. In this specific case we can observe that $OutR$ obtained with *CRRM QoS Based - 50%* is much higher than with *No CRRM*: the reason is that, since UMTS radio channels are optimized to transport the voice flows, for lower traffic values the best choice is always to serve the voice calls in UMTS (*No CRRM*). On the contrary, when the *CRRM QoS Based - 50%*

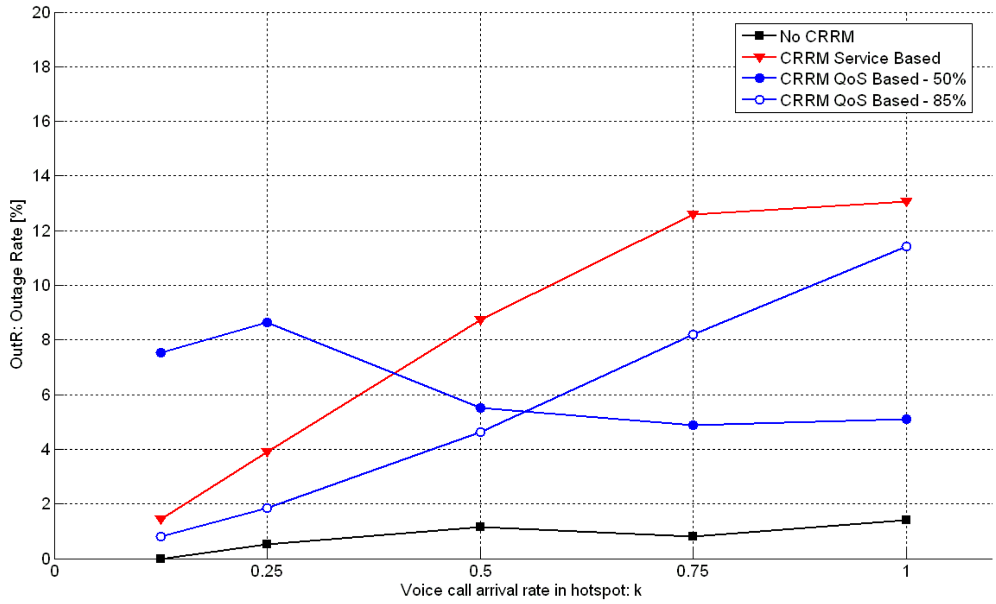


Figure 5.15: Speech Service Integrity QoS in the investigated 100x100 m^2 area: users' Outage Rate $OutR$

is used, some short spikes in the UMTS cell load may trigger the CRRM to move voice calls to WLAN, in which voice users will likely perceive a worst VoIP experience. This is confirmed by the trends in figure 5.16 - *Number of Intersystem Handover procedures per voice call* - in which it is reported that with the option *CRRM QoS Based - 50%*, the users execute about 0.5 Intersystem Handovers per call when $k = 0.125$, much more than with the other configurations.

The same phenomenon is less relevant with *CRRM QoS Based - 85%*: in this case, the higher threshold for UMTS congestion detection is less sensitive to spikes of cell load, therefore for low values of k ($k < 0.5$) less voice calls are moved to WLAN, that is only during real high load situations the CRRM will react and order Intersystem Handover. The drawback of the option of *CRRM QoS Based* at 85% is that by reacting more slowly to radio congestion events or by recognizing a lower number of critical network load situations, when the offered traffic increases ($k > 0.5$), this CRRM option may solve too late some congestion situation, therefore in this case $OutR$ is much higher compared to the *CRRM QoS Based* at 50%.

This distinction finally explains the trends detected in the first figure 5.12 of numerical results, in which the highest Satisfaction Rate $SatR$ is obtained with the *CRRM QoS Based* algorithm, by choosing the two different thresholds for congestion at 50% and at

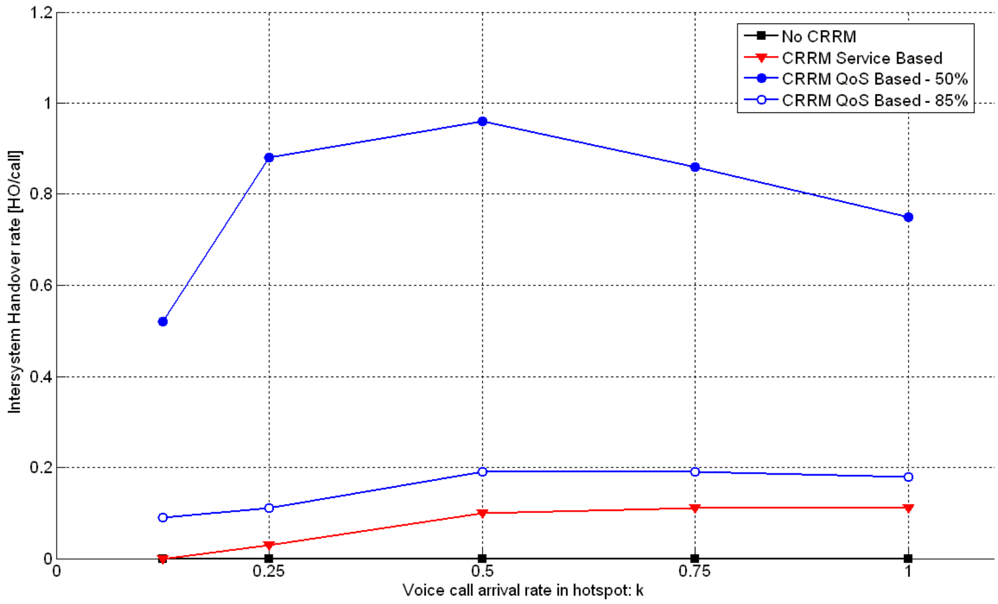


Figure 5.16: Number of Intersystem Handover procedures per voice call

85% in different ranges of traffic (k).

Apart from the number of Intersystem Handover procedures between the two radio access networks, another interesting merit figure of the interworking between UMTS and WLAN is the distribution of voice calls between the different networks. In the figures 5.17 and 5.18, the average subdivision of voice call users between UMTS and WLAN is reported, respectively for $k = 0.5$ and for $k = 1$.

From figure 5.17, it's immediate to verify how the global performance in terms of Satisfaction Rate (see figure 5.12) has been improved from the condition of *No CRRM* to the implementation of the *CRRM QoS Based - 85%*: the latter serves 30% of the users in WLAN, most of them wouldn't otherwise obtain service access. On the contrary the option of *CRRM QoS Based* with threshold at 50% triggers a higher number of Intersystem Handover from UMTS to WLAN: even 55% of the voice users are served by WLAN, therefore, because of the lower performance of VoIP in WLAN, the service integrity of the system is reduced compared to the case of *CRRM QoS Based - 85%*, .

In the case of high traffic load, $k = 1$, in figure 5.18, almost 65% of the voice calls established in the hotspot are served by WLAN with the *CRRM QoS Based - 50%*: this is firstly due to the high number of call attempts that cannot be handled by UMTS because of Admission Control; secondly, the high UMTS network load reports trigger several intersystem handover of voice call connections from UMTS to WLAN. The *CRRM QoS*

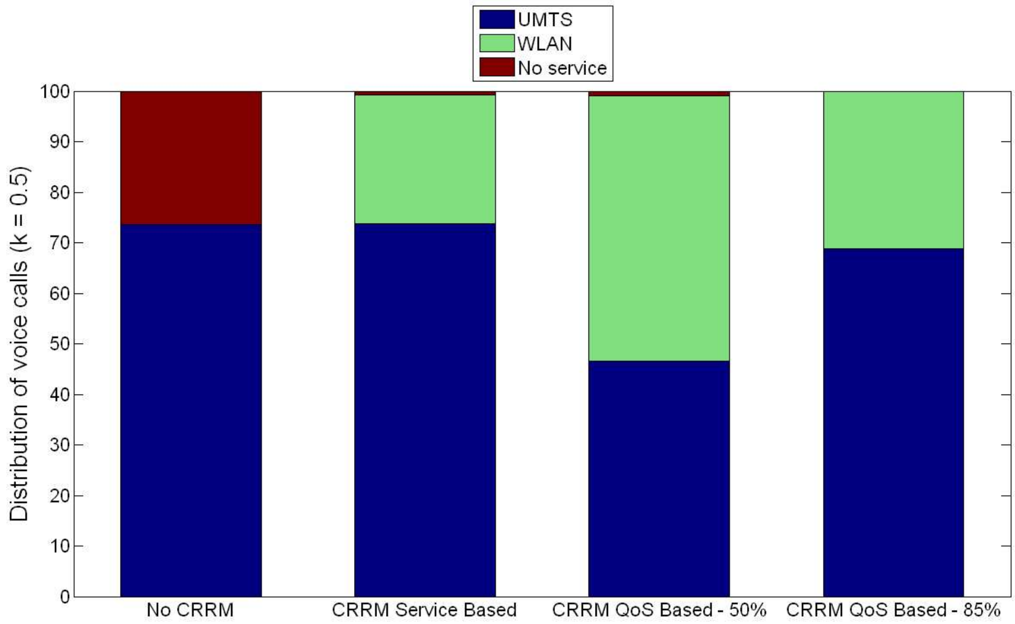


Figure 5.17: Distribution of voice call users in the two networks covering the hotspot ($k = 0.5$)

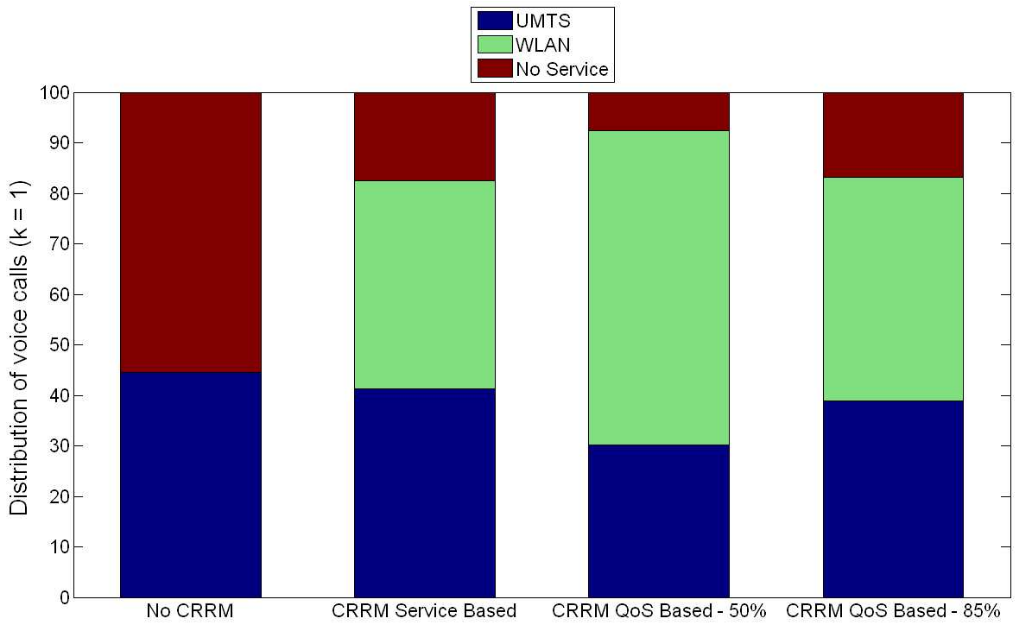


Figure 5.18: Distribution of voice call users in the two networks covering the hotspot ($k = 1$)

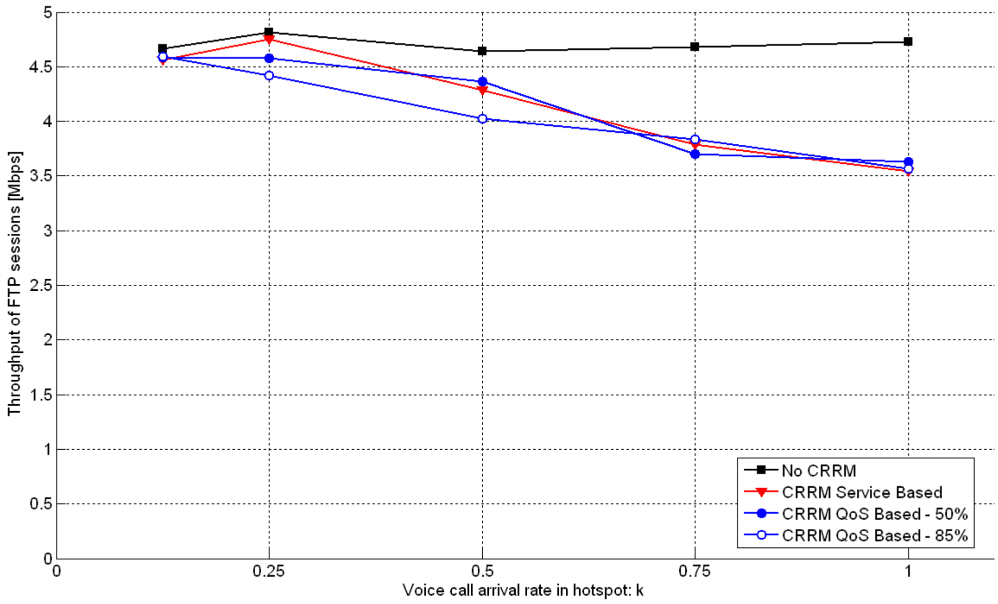


Figure 5.19: Ftp sessions in the hot spot: average perceived throughput

Based option at 85%, on the contrary, triggers less handovers, thus reducing the balancing of radio resources between UMTS and WLAN and finally leading to a higher call block and a worst service integrity and retainability.

In the last figure of this work - figure 5.19 - it is reported the main drawback introduced by adopting a common radio resource management approach to improve the quality of service for speech users in the hotspot, which was the main purpose of this study. As could be expected, the average perceived throughput during FTP sessions in the hotspot is reduced from the case in which the WLAN is dedicated to serve only data traffic (*No CRRM*) to the case of any CRRM configuration in which the WLAN channel is much more busy to "help" the UMTS to serve a larger amount of voice users.

However, we consider acceptable in the worst case scenario ($k = 1$) a reduction of less than 25% for the average FTP throughput, that is a best effort service. As a matter of fact, for $k = 1$, the benefit obtained thanks to the *CRRM QoS Based* is that the Satisfaction Rate of voice users (see figure 5.12) is doubled.

Just for information, when *No CRRM* is used, the performance of WLAN is independent from the voice traffic k in the hotspot; the small oscillations of average FTP throughput in this configuration are due to the fact that each simulation uses seeds generated by different random sequences.

In this chapter we've presented the interactions between the Common Radio Resource Management (CRRM) entity and the local RRM entities in UMTS and WLAN; different CRRM algorithms have been proposed, and simulated with our advanced platform (SHINE) in a realistic scenario of a hotspot of high density traffic covered by both UMTS and WLAN radio access networks. In this final section of the numerical results we have shown that a *CRRM QoS Based* algorithm can largely improve the performance of such an integrated UMTS-WLAN network in terms of served voice users.

Actually, the aim of this study wasn't to find the exact optimum, but it was to prove that it is suitable to realize an integrated heterogeneous wireless network and that with an appropriate CRRM algorithm combining radio interface measurements of both UMTS and WLAN, the quality of service provided by the system can be clearly increased.

Conclusions

In this thesis, the problem of performance evaluation of third-generation and heterogeneous wireless networks is faced, focusing the attention on all main aspects of the radio interface protocol layers, that is, from the physical to the application layer.

To this aim, a TD-SCDMA (Time-Division Synchronous Code Division Multiple Access) system simulator has been realized (TD-SCDMA is one if the three UMTS standard modes): the main functional blocks of this tool are the *Link Level simulator*, the *Network Level simulator*, an *Interface module* between the link and the network levels and finally the *Upper Layer simulator*, which receives the input from a *Mobility simulator* and an *User Activity simulator*.

Thanks to an "instant value" interface approach, the impact of the physical layer parameters for the power control algorithm on the overall end-to-end user performance has been analyzed: we showed that these parameters should be carefully chosen in order to balance the quality of voice and data users in a third-generation mobile network.

Afterwards, the study of system performance has been extended to a wider and really topical scenario, the wireless heterogeneous networks, in which many wireless radio access technologies with complementary characteristics cover the same high traffic areas: the TD-SCDMA and the WLAN IEEE802.11x simulators were used to reproduce this situation.

Let me highlight that at the beginning of my activity, three years ago, only few works on the topic of UMTS-WLAN integration were available in the technical literature and most of them were on signalling and architectural issues only. Now, a growing number of research groups is focusing on this topic and international projects are active on it.

The main contribution of my work in this area has been the design of a *Common Radio Resource Management* algorithm, which can be Coverage Based, Service Based and Quality of Service (QoS) Based. Thanks to an innovative simulation platform called SHINE, the full dynamic simulation of a UMTS-WLAN wireless heterogeneous network has been carried out: the proposed CRRM QoS Based algorithm appears to largely improve the performance of such an integrated UMTS-WLAN network in terms of served users.

This important result shall not be considered only as the final arrival point of this activity, but also as the demonstration that an intelligent tight interworking in heterogeneous networks allows to widely improve the system capacity and that further studies focused on CRRM algorithms based on radio interface measurements can be a prolific field of research.

Appendix A

SHINE: Simulation platform for Heterogeneous Interworking Networks

The timeliness of the issue of access-networks interworking is witnessed by the large number of papers appeared in the open Literature in the recent past on this topic (see for instance [27, 28, 29, 30]), especially with reference to the integration of UMTS and wireless local area networks (WLANs).

Indeed, this issue is very challenging for researchers since investigations on heterogeneous interworking networks can hardly be carried out analytically, owing to the high complexity of technologies involved as well as the large number of aspects to be simultaneously considered, besides those specifically related to the access technologies (traffic characteristics, TCP protocol behavior, realistic channel modelling, ...).

It follows that the only feasible way to accurately investigate the performance of interworking access-networks, in order for instance to design effective common radio resource management strategies, is to realize a software simulator able to capture all aspects of access-networks behaviors, including, as is obvious, the networks interworking strategies.

The realization of such a simulator is still, however, a hard task and should be carried out having two important issues in mind: flexibility and time efficiency.

The former is mainly related to the reusability of most of the software independently on the particular interworking technologies to be investigated, in such a way to make possible to move, for instance, from an UMTS-WLAN simulator to a GPRS-WiMAX one without rewriting almost the entire code. In few words, flexibility is related to a change of perspective in the design of the simulation tool which should be conceived as a simulation platform rather than a dedicated simulator.

As for the time efficiency, it is straightforward to understand that simulating the behavior of two or more interworking access-networks, which obviously operate simultaneously, could be really prohibitive in terms of simulation time, unless parallel execution is performed. The point is: can parallel execution be performed with commonly available computers and without any experience on parallel programming?

In order to achieve both the above mentioned objectives, we developed the simulation platform SHINE (Simulation platform for Heterogeneous Interworking NEtworks), hereafter described.

A.1 Simulation platform Structure

SHINE was developed with the objective to reproduce the behavior of interworking access-networks, taking care of all aspects related to every single protocol level of each access technology affecting the achieved performance.

Its realization was planned in order to overcome the limitations of off-the-shelf network simulation tools, such as, for instance, Opnet [1] or ns-2 [2], in terms of capability to simulate the behavior of heterogeneous networks operating simultaneously and exchanging users.

Hereafter the main criteria we adopted in order to design SHINE in a flexible and time efficient manner are reported.

A.1.1 Flexibility.

To achieve the objective of flexibility we adopted a client-server structure for the simulation platform, which is constituted, in particular, by one server-core simulator (hereafter called Upper Layers Simulator, ULS) and a client simulator for each access technology considered (Lower Layers Simulators, LLSs) (see figure A.1 where, for the sake of clarity, only one LLS is depicted).

The ULS simulator is, in its turn, constituted by an access network(s) side and a Core Network side: at the access network(s) side the ULS takes care of all information related

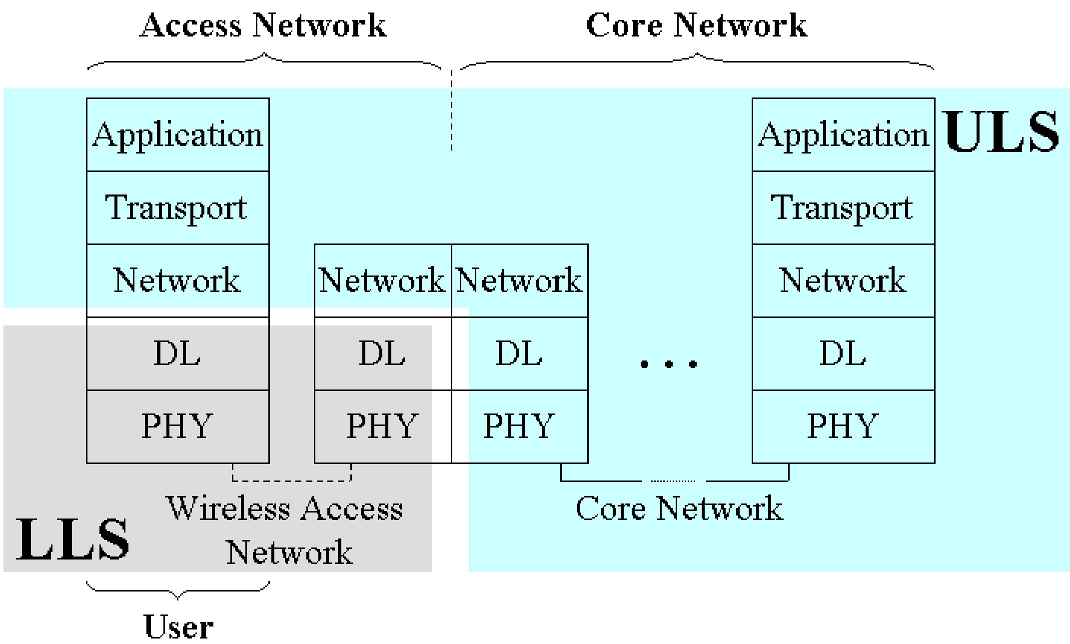


Figure A.1: Simulator platform global architecture.

to those users operating within the region covered by the simulated access-networks, such as their mobility, class of service, etc. and of the end-to-end aspects of each connection, such as the generation of the application-level traffic and the users' TCP or UDP dynamics; at the Core-Network side, instead, the ULS takes care of all aspects concerning communications.

Focusing the attention on the access network(s) side, it is worth noting that the ULS structure, being related to the end-to-end aspects of communications, is independent on the particular access technology (UMTS, WLAN, ...) adopted to establish the user connection.

All aspects related to the access technologies adopted, hence related to the data-link and physical layers and the local Radio Resource Management, are managed by the LLSs, which are the client simulators and are specific for each access technology, so that our simulation platform provides the presence of a LLS for each radio technology adopted in the investigated scenario (see figure A.2).

What is really remarkable is that ULS and LLSs are distinct executables; nonetheless the ULS communicates run time with LLSs through the TCP sockets of the computer operating-system, thus simulating both vertical communications among the protocol layers of the single access technology and, since ULS is on top of all LLSs and manages all end-to-end communications aspects, the access-networks interworking.

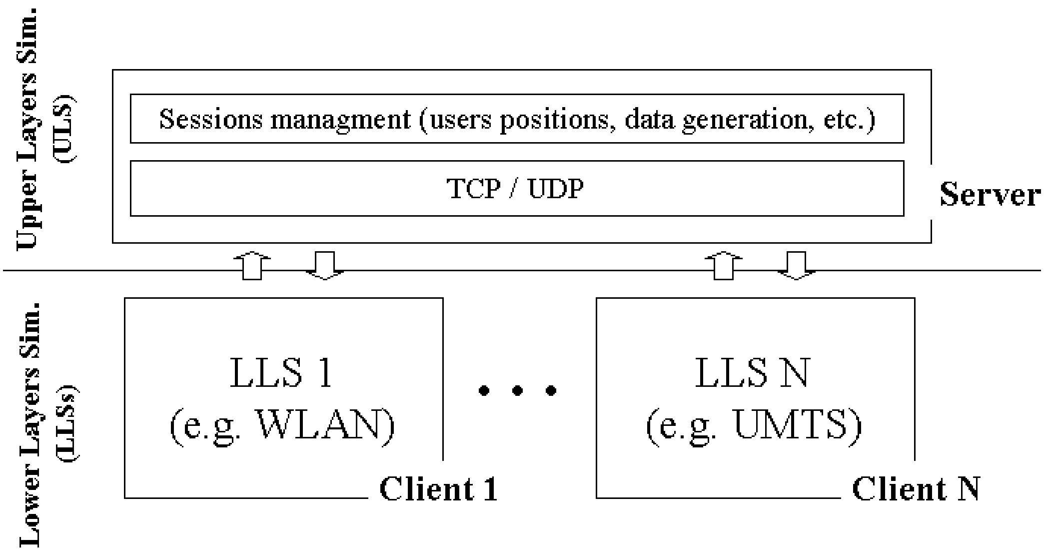


Figure A.2: Simulation platform architecture. Access networks side.

Given this structure, any change in the access technologies investigated requires only to write the related LLS simulators (that is, the data-link and physical level simulators of each new technology including the corresponding RRM entity), which obviously have to be provided with the standardized communication interface to interact with the ULS.

A.1.2 Time efficiency.

As previously stated, the ULS and LLSs are distinct executables which communicates one with the others through the TCP sockets of the PC operating-system; this means that ULS and LLSs can run independently on different personal computers, provided that they are connected to the same computer network. In particular the simultaneous activities of the interworking access-networks is reproduced by the simultaneous executions of the related LLSs.

It is straightforward to understand that with this solution the objective of parallel computing, and therefore of simulation time efficiency, has been easily achieved without resorting to multiprocessor computers and to the paradigms of parallel programming.

A.2 ULS and LLSs main tasks

As previously stated, the ULS manages the end-to-end aspects of each connection (no matter the access technology supporting them at the physical and data-link levels), hence its tasks are mainly concerned with communications management (connections setup and closure, management of application level traffic flows, ...), the simulation of transport level protocols (TCP, UDP, ...) and the processing of simulation outcomes to provide application level performance. In particular, the main tasks of ULS are:

- to set the starting instant of each new traffic session originated by users according to the arrival statistics of the traffic class it belongs to (http, e-mail, voice calls, ...), as well as users positions within the investigated scenario;
- to manage connection setup and closure procedures;
- to generate the bit-flows up(down)loaded by users in each session according to the statistics of their class of traffic;
- to reproduce the transport protocol behavior;
- to perform packet segmentation and reassembly;
- to select through which technology (that is, through which LLS) should each user connect to the network on the basis of user-defined rules and available information on the current status of each network; it can also decide to reject a connection or to move it from an LLS to another (that is, from a given technology to another) at any time, thus simulating vertical (i.e. inter-system) handovers;
- to collect, finally, all simulation outcomes and to generate the outputs (user satisfaction rate, throughput, packet delivery delays,...) from an end-to-end point of view.

As for the LLSs, since they are specific for the particular access technologies investigated, their tasks are mainly concerned with data-link and physical level aspects of communications and, in particular, are:

- to perform, if required, the call admission control specific of the technology it simulates and all technology specific radio resource management;
- to manage, if required, the transmission scheduling at the data-link layer level;
- to perform MAC or RLC fragmentation and reassembly of TCP-IP level packets;
- to simulate MAC/RLC behavior of the given technology;

- to reproduce all physical layer procedures related to each transmission and reception: power control, handover, radio frequency measurements, channel coding, modulation, information detection, decoding, etc.
- to simulate all main Radio Resource Management (RRM) functions (see section 5.3 for more details).
- to collect, finally, all simulation outcomes and to generate the outputs (user satisfaction rate, throughput, packet delivery delays,...) from the wireless links point of view (that is, at data-link and physical level).

It is important to underline that the above reported structure and tasks division allows an easy management of vertical (inter-system) handovers, which have obviously to be simulated when investigating interworking access-networks. Vertical handovers, in fact, essentially determine a change of the access technology adopted by users; from the simulation point of view they mainly implies to switch the management of users physical and data-link levels from one LLS to another, and this can be easily managed by the ULS which is on top of all LLSs.

A.3 Time axis management

In SHINE each LLS manages its own time axis; the ULS, for its part, communicates to the LLSs the next instant in which some event concerning the upper protocol levels happens (connection request, start of bit transfers, TCP timeouts, etc.) properly setting a parameter called *ULProxEvent*, which represents a "time-flag". This way, when the time counter of an LLS reaches that instant, the related LLS simulation stops and a "call" to the ULS is performed asking for the event-related information. After the ULS reply, the simulation of the calling LLS is resumed, the consequent LLS actions (MAC level frame queuing, ...) are taken, the time-flag *ULProxEvent* is updated and the LLS time counter is restarted.

An LLS can perform a call to the ULS not only when the instant of the next ULS event has come but also whenever an LLS event which is of interest for the ULS takes place, such as, for instance, the correct transmission of a MAC frame.

This mechanism is illustrated in figure A.3 where both lower layers (A.3-b) and upper layers events (A.3-c) are depicted.

The above described stop-and-wait procedure is the basis for the coordination among LLSs, which is obviously needed when simulating interworking networks: in case an ULS event is of interest for more than one LLS (for instance a new traffic session searching for

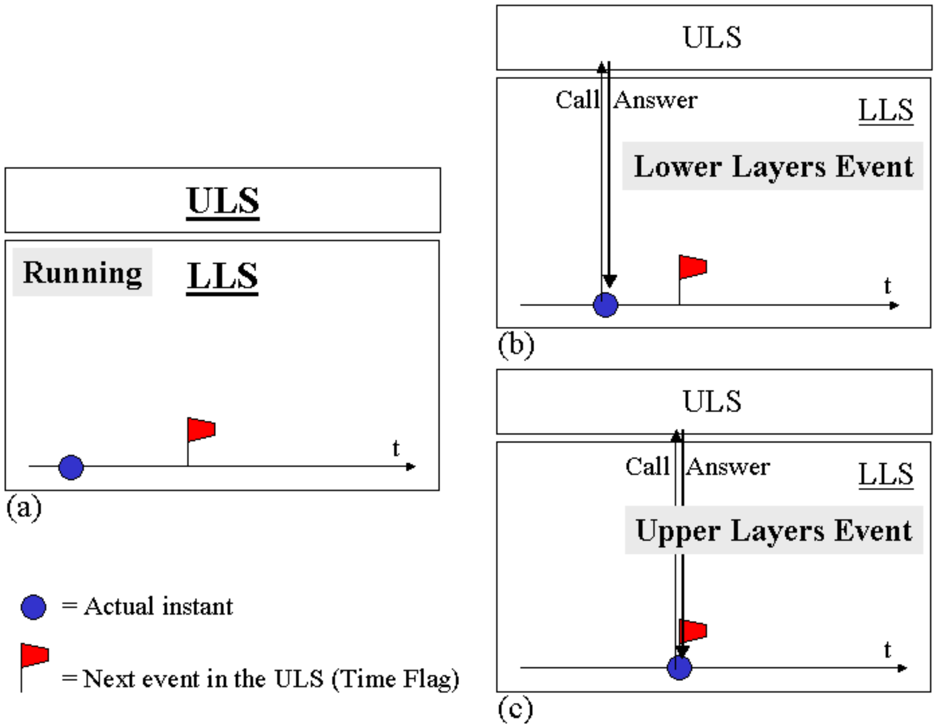


Figure A.3: ULS-LLS communications.

the best connection), no ULS reply is granted to the calling LLSs until all the interested LLSs have stopped waiting for it.

It follows that although LLS executions take place at different speeds (LLSs complexity could be different and they could be running on different PCs), the faster LLSs periodically stop and wait for the LLSs they are interworking with, thus re-synchronizing simulations.

Please note that the depicted architecture leaves the LLS' implementer free to decide whereas an event driven or a time slotted simulator is the best choice for the specific task. For instance, while a WLAN simulator is generally implemented as an event driven simulator, a cellular system, characterized by a fixed radio frame structure, may be more easily reproduced adopting a slotted time axis.

A.4 ULS-LLSs communications

As previously stated, ULS is the server simulator which always waits for LLSs calls (that is, the clients calls), either because something happened at the lower layers (case *b* in figure A.3), or because the LLSs current times reached *ULProxEvent* (the flag was reached in case *c* of figure A.3).

All communications from LLS to ULSs and vice-versa can be classified adopting a

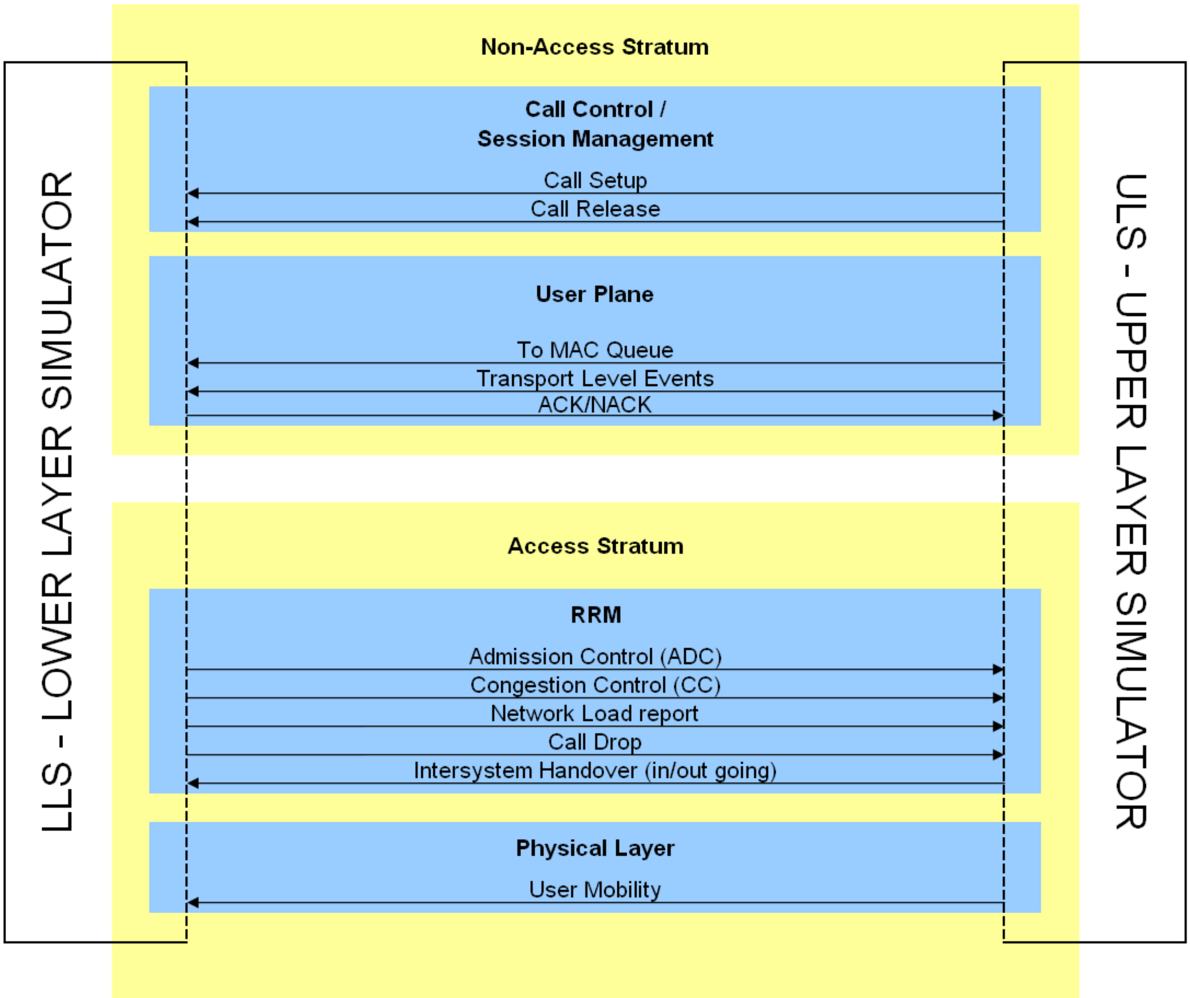


Figure A.4: ULS-LLSs communications scheme (cellular network notation) implemented in SHINE cellular network formalism, as shown in figure A.4. On the left side of the figure the generic LLS is represented, whereas the ULS is depicted on the right side. From top to bottom of the central part of the figure, we find the logical planes of the overall system: from the call control/session management down to the physical layer, that is, from all the Non Access Stratum functions to the Access Stratum ones. Thus, the arrows and their directions clearly show all possible communications occurring between the ULS and LLSs. Obviously, the arrows from left to right are related to lower layers events to be communicated to the ULS, while the others are due to the flag reaching, hence due to ULS events to be communicated to the LLSs.

Bibliography

[1] OPNET Technologies, Inc. http://www.opnet.com

[2] NS-2 Network Simulator. http://www.mash.cs.berkley.edu/ns/

[3] Glomosim Simulator. http://pcl.cs.ucla.edu/projects/glomosim/

[4] M.Wellens, M.Petrova, J.Riihijrvi, P.Mhnen, "*Building a Better Wireless Mousetrap: Need for More Realism in Simulations*", Proc. of WONS 2005, St. Moritz, Switzerland, January 2005.

[5] R.J.Punnoose, P.V.Nikitin, D.D.Stancil, "*Efficient Simulation of Ricean Fading within a Packet Simulator*", Proc. IEEE 2000 Fall Vehic. Tech. Conf., pp. 764-767.

[6] J.Heidemann et al., "*Effects of Detail in Wireless Network Simulation*" SCS Multiconference on Distributed Simulation, Phoenix, Arizona, USA January, 2001.

[7] D.Cavin, Y.Sasson, A.Schiper, "*On the accuracy of MANET Simulators*", POMC'02, October 30-31, 2002, Toulouse, France.

[8] M.Takai, J.Martin, R.Bagrodia, "*Effects of Wireless Physical Layer Modeling in Mobile Ad Hoc Networks*", Proc. MobiHoc 2001.

[9] G.Yeung et al, "*Detailed OFDM modeling in network simulation of mobile ad hoc networks* ", 18th Workshop on Parallel and Distributed Simulation, 2004. PADS 2004. May 2004.

[10] IEEE Std. 802.11a, "*High-Speed Physical Layer in the 5 GHz Band*", 2000.

[11] J. Medbo, H. Andersson, P. Schramm, H. Asplund, J-E. Berg, "*Channel Models for HIPERLAN/2 in different indoor scenarios*,1998-04".

[12] W.C. Jakes, "*Microwave Mobile Communications*". Piscataway, NJ. IEEE Press, 1994.

[13] IEEE Std. 802.11, "*Wireless LAN Medium Access Control (MAC) And Physical Layer (PHY) Specifications*", 1999.

[14] K. Kammerlander, "Benefits and implementation of TD-SCDMA," International Conference on Communication Technology, WCC-ICCT 2000., pp. 1013-1016, 21-25 Aug. 2000.

[15] J. Zhang, W. Wang, "Dynamic system level simulation for radio resource management of TD-SCDMA high speed downlink packet access," International Conference on Communications, Circuits and Systems and West Sino Expositions, pp. 187-191, 29 June-1 July 2002.

[16] Y. Zhang, Y. Chang, D Yang, "Link performance of TD-SCDMA systems," Spread Spectrum Techniques and Applications, 2002 IEEE Seventh International Symposium on , pp.14-18, 2002.

[17] S. Hmlinen, P. Slanina, M. Hartman, A. Lappetelinen and H. Holma, "A Novel Interface Between Link and System Level Simulations", ACTS Mobile Telecommunications Summit'97, Aalborg, Denmark, in proceedings pp. 599-604, October 7-10, 1997.

[18] ETSI TS 123 107 v 5.10.0, "Universal Mobile Telecommunication System (UMTS); Quality of Service (QoS) concept and architecture", septembre 2003

[19] P. Slanina, F. Raji, C. Mecklenbräuker, C. Hölzl, G. Ostermayer, T. Stadler, T. Gruhn, "Performance of hybrid ARQ types for TDD mode", Proc. ACTS 4^{th} Mobile Communications Summit, 1999

[20] V.V. Veeravilli, A. Sendonaris, "The coverage-capacity tradeoff in cellular CDMA systems", IEEE Transactions on Vehicular Technology, Vol. 48, No.5, September 1999, pp. 1443-1450

[21] H. Holma, A. Toskala, "WCDMA for UMTS Radio Access for Third Generation Mobile Communications", WILEY, September 2000.

[22] A. Mehrotra, "GSM System Engineering", Artech House, 1997.

[23] E. Malkamaki, F. de Ryck, C. Mourot, "A Method for Combining Radio Link Simulations and System Simulations for a Slow Frequency Hopped Cellular System", IEEE Veh. Technol. Conf. 1994 (VTC'94), pp. 1145-9, Phoenix, AZ, May 1994.

[24] H. Olofsson, M. Almgren, C. Johansson, M. Hook, F. Kronestedt, "Improved Interface Between Link Level and System Level Simulations Applied to GSM", IEEE Int. Conf. On Univ. Pers. Comm. (ICUPC'97), Page(s): 79 -83, vol.1, S. Diego, CA, Oct. 1997.

[25] Z. Kostic, X. Qiu, J. Chuang, L.F. Chang, K. Chawla, N. Sollenberger, "Dynamic Frequency Hopping in Cellular Systems with Network Assisted Resource Allocation", IEEE VTC'2000spring, pp. 2459-2463, Tokyo, May, 2000.

[26] ITU-T Recommendation E.800, "*Terms and definitions related to quality of service and network performance including dependability*". Augusto 1994.

[27] Kaveh Pahlavan et al., "*Handover in Hybrid Mobile Data Networks*", IEEE Personal Communications, April 2000.

[28] Shiao-Li Tsao, Chia-Ching Lin, "*Design and Evaluation of UMTS-WLAN Interworking Strategies*", 2002 Computer & Communications Research Laboratories, Industrial Technology Research Institute, Hsinchu, Taiwan, R.O.C.

[29] M. Lott et al., "*Interworking of WLAN and 3G systems*", IEE Proceedings-Communications Volume 151, Issue 5, 24 Oct. 2004 Page(s):507 - 513.

[30] Hongyang Bing,Chen He,Lingge Jiang, "*Performance analysis of vertical handover in a UMTS-WLAN integrated network*", IEEE PIMRC 2003.

[31] IEEE 802.11 WG, IEEE 802.11e/D5.0, "*Draft Supplement to Standard for Telecommunications and Information Exchange Between Systems LAN/MAN Specific Requirements Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS)*, Aug. 2003".

[32] Ad Kamerman and Leo Monteban, "*WaveLAN®- II: A High-Performance Wireless LAN for the Unlicensed Band*", Bell Labs Technical Journal, Summer 1997.

[33] D. Gu, J. Zhang, "*A new measurement-based admission control method for IEEE802.11 wireless local area networks*" Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on , Volume: 3 , 7-10 Sept. 2003 Pages:2009 - 2013 vol.3

[34] IEEE 802.20 Working group on Mobile Broadand Wireless Access, "*Traffic Model for IEEE 802.20 MBWA System Simulations (Baseline Draft)*", N.K.Shankaranarayanan, July 2003.

[35] http://www.umatechnology.org

[36] Http://www.ist-mobydick.org

[37] Http://www.Monasidre.com

[38] N.Vulic, I.Niemegeers, S.H.de Groot, "*Architectural options for the WLAN integration at the UMTS radio access level*", IEEE VTC 2004-Spring.

[39] Kim Sung-Eun, J.A.Copeland, "*Interworking between WLANs and 3G networks : TCP challenges*" IEEE WCNC 2004.

[40] M.Bernaschi, F.Cacace, A.Pescape, S.Za, "*Analysis and experimentation over heterogeneous wireless networks*", Tridentcom 2005.

[41] J.Indulska, S.Balasubramaniam, “*Context-aware vertical handovers between WLAN and 3G networks*”, IEEE VTC 2004-Spring.

[42] http://www.ist-scout.org

[43] http://www.freeband.nl/kennisimpuls/projecten/4gplus

[44] G.F.Riley et al., “*Distributed network simulations using the dynamic simulation backplane*” 21st International Conference on Distributed Computing Systems, 16-19 April 2001.

[45] “*ITU-Reccomendation H.323: Packet-based multimedia communications systems*”, International Telecommunication Union, 1997.

[46] Bur Goode, “*Voice Over Internet Protocol (VoIP)*”, Proceedings of the IEEE, vol.90, No.9, sept. 2002.

[47] 3GPP, TS 25 series.

[48] 3GPP TS 25.221 V5.2.0, ”Physical channels and mapping of transport channels onto physical channels (TDD)”

[49] 3GPP TS 25.222 V5.7.0, ”Multiplexing and channel coding (TDD)”

[50] 3GPP TS 25.223 V5.3.0, ”Spreading and modulation (TDD)”

[51] 3GPP TS 25.331 V5.8.0, ”Radio Resource Control (RRC); Protocol specification”

[52] 3GPP TS 25.102 V5.5.0, appendix A, ”UE Radio Transmission and Reception (TDD)”, june 2003

[53] 3GPP TS 25.105 V5.4.0, appendix A, ”Base Station (BS) radio transmission and reception (TDD)”, april 2003

[54] ITU-R Recommendation M.1225, “*Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000*”. 1997.

[55] 3GPP TR 101 112 V3.2.0, “*Technical Report Universal Mobile Telecommunications System (UMTS); Selection procedures for the choice of radio transmission technologies of the UMTS*”. 1998-04.

[56] CWTS 3G digital cellular telecommunications system TSM , ”Multiplexing and multiple access on the radio path”, 05-02, Dec. 2001., Dec. 2001.

[57] ”cdma2000 Evaluation Methodology”, 3GPP2 C.R1002-0 v1.0, December 10 2004.

[58] 3GPP TS 24.008 V5.16.0, ”Mobile radio interface Layer 3 specification”, June 2006

[59] 3GPP TS 25.413 V5.12.0, ”UTRAN Iu interface RANAP signalling”, June 2005

[60] 3GPP TS 43.318 V7.0.0, "Generic access to the A/Gb interface"

[61] 3GPP TR 22.934 V6.0.0, "*Feasibility study on 3GPP system to Wireless Local Area Network (WLAN) interworking (Release 6)*, 2002"

[62] 3GPP TS 23.107. Technical Specification Group Services and System Aspects; QoS Concept and Architecture

Acknowledgments

It is impossible to acknowledge every one who had contributed in some way to the results of my doctoral course. Nonetheless, to the best of my ability I shall attempt to do so.

I would thus thank Professor Oreste Andrisano, who always supported and trusted my job, also in the periods out of Bologna. Besides these three years of studies, I've undertaken a long journey with him, which started on the summer of 2002 when I joined for the first time the IEIIT-CNR institute for my degree thesis. Working in his research team has been a fantastic opportunity.

Special thanks are dedicated to Gianni Pasolini and Alessandro Bazzi. I'm grateful to have worked with them, they have been really my guidance in this period, and I'm only sorry not being able to spend much more time with them in Bologna; thanks for your help and encouragements to finalize this thesis!

I'm happy to convey my appreciation to the other researchers and professors of the University of Bologna, Alberto Zanella (thanks to let me go to Las Vegas!), Barbara Masini, Davide Dardari, Andrea Conti and Roberto Verdone.

Great thanks to Lorenzo Faggioni from Siemens Mobile. We have enjoyed together many epic battles in the UMTS test lines; he has given me a lot of freedom to deepen my research interests, he has transmitted to me the passion for teamwork, but most of all he's a great friend.

I wish to thanks also my unforgettable colleagues (friends) in Siemens Milano, Giovanni De Pascalis, Marco Guidone, Manuel Piras, Simone Buzzi, Alberto Fontana, Andrea Sala and Arnó Valerio: how much fun we had in these years! And how many discussions about football, potatoes and UTRAN faults.

A special thank to Andrea Brambilla from Siemens. We left Milan together and now we share an amazing apartment in Dusseldorf and an exceptional period of our life; I'm very happy of that!

I would like also to thanks Augusto Tomas and Doctor René Jahnkow from Siemens Dusseldorf. We experienced how to setup a full UMTS network alone in the desert of Saudi Arabia, unrepeatable adventure!

I wish to thanks Gianluigi Liva for his constant support, also when things weren't going in the best way; he's another Italian Ph.D. who immigrated in Germany and I share with him the idea to go back to Romagna, asymptotically.

Thanks to Mauro Gibertoni, for his friendship for a life.

I want finally to express unique thanks to my family. My parents and my two sisters, Laura and Gabriella, supported and always encouraged me during the long period of my studies. By dedicating them this thesis, I know that I'm just giving back a little bit of what they did (and they are still doing) for me. Nevertheless, they know how much I love them.

13th March 2007