Third Generation Mobile Technology

and

its evolution towards Fourth Generation

by

Jorge Vicente Gómez

Supervisor: **Lecturer Ari Rantala**

Tampere, Finland

February, 2009

To my parents,
because I owe everything that I am to you

"Any sufficiently advanced technology is indistinguishable from magic."
Arthur C. Clarke.

Abstract

Second Generation telecommunication systems, such as the Global System for Mobile Communications (GSM), enabled voice traffic to go wireless: the number of mobile phones exceeds the number of landline phones and the mobile phone penetration is approaching 100% in several markets. The data-handling capabilities of Second Generation systems are limited, however, and Third Generation systems are needed to provide the high bit-rate services that enable high-quality images and video to be transmitted and received, and to provide access to the Web with higher data rates. These Third Generation mobile communication systems are referred to in this thesis as the Universal Mobile Telecommunication System (UMTS). Wideband Code Division Multiple Access (WCDMA) is the main Third Generation air interface in the world, and deployment is ready in Europe and Asia in the same frequency band, around 2 GHz. WCDMA has also been deployed in the USA in the US frequency bands. WCDMA air interface is available in more than 150 commercial networks and most of those networks have already launched the next phase of WCDMA, High Speed Downlink Packet Access (HSDPA). The growth continues, as there are more commercial networks, more terminals across all categories and more data services being deployed. The large market for WCDMA and its flexible multimedia capabilities will create new business opportunities for manufacturers, operators, and the providers of content and applications.

The thesis is structured as follows. First, in the first chapter we review the previous technologies to the WCDMA, from the most primitive way of telecommunication till evolved Second Generation, next to Third Generation.

Next, the second chapter introduces the Third Generation Technology, the research works from the beginning in 1988 until its standardisation in 1998, the spectrum allocation, the requirements, a brief resume of the improvements and the main differences from Second Generation.

Then the third chapter describes the main parameters in WCDMA air interface, including spreading, correlation receiver, Rake receiver, power control and handovers. It also shows the architecture of the radio access network, interfaces within the radio access network between base stations and radio network controllers.

Thereafter the evolution towards Fourth Generation is described in the fourth charapter with the Release 5 (HSDPA), Release 6 (HSUPA), Release 7 (HSPA+) and Release 8 (LTE). The most important features of each release are explained in the corresponding subsections as well as the improvements in the throughput, the specifications and the modifications regarding previous releases.

Acknowledgements

My greatest debt is to my parents, Pilar and Florentino, and my brother, Óscar, for their unconditional love and support, which got me so far. My parents encouragement and faith in me have truly kept me on track and ensured my progress. I thank my lifetime friends from Madrid, Rober, Acosta, Chiqui, Uge, Juan. They supported me when I made the decision of coming to Finland and cheered me up in my first days here in Tampere. I am specially thankful to María, her love and trust in me make me face any challenge. I do not forget my university classmates from Madrid: Javier, Kike, David, Alex. They helped me so much during my studies. We spent a lot of time together in the library, helping each other. I would like to express my sincere gratitude to my teachers from EUITT Madrid who taught me all I know.

I am really grateful to my home university, Universidad Politécnica de Madrid EUITT, for giving me the opportunity of coming to Finland. It was a dream I had for a long time. I thank my tutor, Josep Lluis Sans for guiding me when I came to Finland. I would have been so lost without his advices. I want also thank my receiving university, TAMK, for the wonderful hospitality you showed me during the entire process. I would like to appreciate the excellent library facilities there are in Tampere, specially in TAMK and TUT, which made finding books and references extremely easy.

I am very grateful to my International Coordinators from Madrid, Mari Carmen and Irina who helped me so much with the Erasmus paperwork and also to my International Coordinator in Tampere, Mirja Onduso, who also helped me with the paperwork and guided me in the beginning of my stay. Thank my supervisor, Ari Rantala for suggesting me this topic for my Final Thesis. It was really interesting and I learnt a lot about it writing this thesis.

I would also like to mention my flatmates, Ossi and Sharif. I could not have imagined a better place for living than this. I have shared so many nice moments here with them during my stay. Although we are far from each other, I will never forget you. And also I thank all my friends here in Tampere: Javi, Marc, Noelia, Patricia, Nacho, Aurélien, Hawon, Dennis, Romain and so many other for greatly enriching my life outside my studies. It was wonderful to know so many different cultures.

Kiitos!!

Contents

Abstract v

Acknowledgements vii

List of Figures xiii

List of Tables xv

Abbreviations xvii

1. Background 1

1.1 0G: Mobile Radio Telephone 1
1.2 1G: First Generation Wireless Telephone Technology 3
1.3 2G: Second Generation wireless telephone technology 4
1.3.1 GSM Technology 5
1.3.2 EDGE Technology 6
1.3.3 Interim Standard 95, cdmaOne 8
References 8

2. Introduction 9

2.1 Beginning of Third Generation Technology 9
2.2 Creation of 3GPP (Third Generation Pattern Technology) 10
2.3 Creation of 3GPP2 11
2.4 IMT-2000 Process in ITU 11
2.5 Requirements for Third-Generation Systems 12
2.6 Spectrum Allocations for Third-Generation Systems 14
References 15

3. Third Generation Technology 17

3.1 Wideband Code Division Multiple Access (WCDMA). 17
3.1.1 Introduction and summary of the main parameters in WCDMA 17
3.1.2 Generic principles of CDMA operation 18
3.1.3 Multipath Radio and Rake Reception 21
3.1.4 Power Control 24
3.1.5 Different kinds of Handovers 26
3.1.5.1 Softer Handover 26
3.1.5.2 Soft Handover 27
3.1.5.3 Hard Handover 27
3.2 Radio Access Network Architecture 28
3.2.1 Introduction to the System Architecture 28
3.2.2 User Equipment (UE) 29
3.2.3 UTRAN Architecture 30
3.2.3.1 The Node B 30
3.2.3.2 The Radio Network Controller (RNC) 31
3.2.3.3 UTRAN Interfaces 31
3.2.4 GSM Core Network architecture 32
3.2.5 External networks 33
3.2.6 General Protocol Model for UTRAN Terrestrial Interfaces 33
3.2.6.1 Horizontal Layers 33
3.2.6.2 Vertical Planes 33

4. The path towards Fourth Generation **37**

4.1 High-Speed Downlink Packet Access (HSDPA) 37
4.1.1 Introduction to HSDPA 37
4.1.2 HSDPA Terminal Capability and Achievable Data Rates 40
4.1.3 Mobility with HSDPA 41
4.1.3.1 Intra-Node B HS-DSCH to HS-DSCH Handover 42
4.1.3.2 Inter-Node-Node B HS-DSCH to HS-DSCH Handover 43
4.1.3.3 HS-DSCH to DCH Handover 44
4.1.4 HSDPA Performance 45
4.1.4.1 Factors Governing Performance 45
4.1.4.2 Spectral Efficiency, Code Efficiency and Dynamic Range 45
4.2 High-Speed Uplink Packet Access (HSUPA) 46
4.2.1 Introduction to HSUPA 46
4.2.2 HSUPA Feasibility 48
4.2.3 HSUPA Physical Layer Structure 49
4.2.4 HSUPA Physical Layer Operation Procedure 50
4.2.5 HSUPA Terminal Capability 51
4.2.6 HSUPA Performance 52
4.2.6.1 Physical Layer Retransmission Combining 52
4.2.6.2 Node B-Based Scheduling 52
4.3 High-Speed Packet Access Evolution (HSPA+) 53
4.3.1 Introduction to HSPA+ 53
4.3.2 Downlink MIMO for HSPA+ 53
4.3.3 Mobile Power Consumption Reduction with Continuous Packet Connectivity . . 57
4.3.4 Voice-over-IP (VOIP) Capacity Enhancements 57
4.3.5 Flat Architecture 58
4.4 UTRAN Long-Term Evolution (LTE) 60
4.4.1 Background and introduction to LTE Technology 60

4.4.2 LTE Multiple Access 64
4.4.2.1 OFDMA Principles 64
4.4.2.2 SC-FDMA Principles 68
4.4.3 Performance 70
4.4.3.1 Peak Bit Rates 70
4.4.3.2 Spectral Efficiency 72
References 74

5. Conclusions **75**

Bibliography **77**

List of Figures

1.1 Half-Duplex . . . 1

2.1 Standardisation and commercial operation schedule for WCDMA and its evolution . . 11
2.2 Spectrum allocation . . . 15

3.1 Spreading and despreading in DS-CDMA . . . 18
3.2 Principle of the CDMA correlation receiver . . . 19
3.3 Multipath propagation leads to a multipath delay profile . . . 20
3.4 Fast Rayleigh fading as caused by multipath propagation . . . 21
3.5 Block diagram of the CDMA Rake receiver. . . . 22
3.6 Closed loop power control . . . 23
3.7 Closed-loop power control compensates a fading channel . . . 24
3.8 Softer handover . . . 25
3.9 Soft handover . . . 26
3.10 UMTS System Architecture . . . 27
3.11 User Equipment . . . 28
3.12 UTRAN architecture . . . 29
3.13 GSM Core Network . . . 31
3.14 General protocol model for UTRAN terrestrial interfaces . . . 33

4.1 Example of intra-Node B HS-DSCH to HS-DSCH handover . . . 40
4.2 Example of inter-Node B HS-DSCH to HS-DSCH handover . . . 41
4.3 Example of HS-DSCH to DCH handover . . . 42
4.4 HSUPA Physical Layer Structure . . . 48
4.5 MIMO for HSPA+ . . . 52
4.6 Different types of constellations . . . 53
4.7 Voice capacity evolution with Release 7 VoIP . . . 56
4.8 Evolution towards flat architecture . . . 57
4.9 Evolution of the mobile technologies . . . 58
4.10 Uplink and downlink data rates compared for HSPA and LTE . . . 60
4.11 Frequency domain scheduling . . . 61
4.12 Comparison of the UMTS and LTE architectures . . . 62
4.13 Adjacent subcarrier with OFDMA . . . 63
4.14 OFDMA transmitter and receiver diagrams . . . 64
4.15 IFFT/FFT principle . . . 65
4.16 Use of cyclic prefix for removing inter-symbol interference . . . 66
4.17 Simplified SC-FDMA transmitter and receiver chains . . . 66
4.18 SC-FDMA with frequency-domain generation . . . 67
4.19 FDMA multiplexing of users with FFT/IFFT implemented in the transmitter 68
4.20 Relative spectral efficiency of LTE compared with HSPA R6 . . . 72

List of Tables

2.1 Main differences between WCDMA and GSM networks 13

4.1 Comparison of fundamental properties of DCH and HS-DSCH 37
4.2 HSDPA terminal capability categories 38
4.3 Theoretical bit rates with 15 multi-codes for different TFRCs 39
4.4 Comparison of fundamental properties of DCH and E-DCH 45
4.5 Comparison of fundamental properties of HSDPA and HSUPA 46
4.6 Selected HSDPA terminal categories 54
4.7 Selected HSUPA terminal categories 54
4.8 Downlink peak bit rates 69
4.9 Uplink peak bit rates 70

Abbreviations

3GPP	3rd Generation Partnership Project
AAL	ATM Adaptation Layer
ALCAP	Access Link Control Application Part
AMC	Adaptive modulation and coding
AMPS	Advanced Mobile Phone Service
AMR	Adaptive multirate
ARIB	Association of Radio Industries and Business
ARP	Autoradiopuhelin
ARQ	Automatic Repeat reQuest
ATDMA	TDMA-based Advanced TDMA Mobile Access
ATM	Asynchronous Transfer Mode
AWS	Advanced Wireless Services
BB	Baseband
BER	Bit error rate
BLER	Block error rate
BoD	Bandwidth on Demand
BPSK	Binary phase shift keying
BSC	Base Station Controller
BTS	Base Transceiver Station
CBR	Constant Bit Rate
CBU	Control Basic Unit
CDMA	Code Division Multiple Access
CIR	Carrier to Interference Ratio
CODIT	CDMA-based Code Division Testbed
CPCH	Common Packet Channel
CQI	Channel Quality Indicator
CRC	Cyclic redundancy check
CS	Circuit Switching
DAC	Digital to Analog Converter
DC	Direct Current
DCS	Digital Cellular System (GSM 1800)
DCH	Dedicated Channel
DL	Downlink
DPCCH	Dedicated Physical Control Channel
DPDCH	Dedicated Physical Data Channel
DRX	Discontinuous Reception
DS	Direct Sequence
DSCH	Downlink-shared Channel
E-DCH	Enhanced DCH
E-DPCCH	Enhanced DPCCH

E-DPDCH	Enhanced DPDCH
EDGE	Enhanced Data rates for GSM Evolution
ETSI	European Telecommunications Standards Institute
EUTRAN	Evolved UTRAN
FACH	Forward Access Channel
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FFT	Fast Fourier Transform
GMSK	Gaussian minimum-shift keying
GSM	Global System for Mobile communications
HARQ	Hybrid ARQ
HLR	Home Location Register
HPF	High Pass Filter
HS-DPCCH	High-Speed Dedicated Physical Control Channel
HS-DSCH	High Speed Downlink-shared Channel
HS-PDSCH	High Speed Physical downlink shared channel
HS-SCCH	High-Speed Shared Control Channel
HSDPA	High Speed Downlink Packet Access
HSUPA	High Speed Uplink Packet Access
IC	Integrated Circuit
IEEE	Institute of Electrical and Electronics Engineers
IMTS	Improved Mobile Telephone Service
IMT-2000	International Mobile Telecommunication
ISDN	Integrated Services Digital Network
ITU	International Telecommunication Union
LAN	Local Area Network
LPF	Low Pass Filter
LTE	Long Term Evolution
MAC	Medium access control
MC	Multi-carrier
MCU	Multipoint Control Unit
MIMO	Multiple Input Multiple Output
MMS	Multimedia Message
MSC	Mobile Services Switching Centre
MT	Mobile Termination
NBAP	Node B Application Part.
NMT	Nordic Mobile Telephone System
OFDMA	Orthogonal Frequency Division Multiple Access
OLT	Offentlig Landmobil Telefoni
PAR	Peak to average
PCPICH	Primary Common Pilot Channel
PCS	Personal Communication Services
PDU	Protocol data units
PLL	Phase Locked Loop
PLMN	Public Land Mobile Network
PS	Packet Switching
PSTN	Public Switched Telephone Network
PTT	Push to Talk
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RACE	Research of Advanced Communication Technologies in Europe

RACH	Random Access Channel
RAN	Radio Access Network
RBS	Radio Base Station
RLC	Radio Link Control
RNC	Radio Network Controller
RNS	Radio Network Subsystem
RRC	Radio Resource Control
S-CCPCH	Secondary common control physical channel
SC-FDMA	Single Carrier FDMA
SF	Spreading Factor
SINR	Signal to interference and noise ratio
SIR	Signal-to-Interference
SMS	Short Message Service
SRB	Signaling Radio Bearer
TCP	Transport Control Protocol
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TE	Terminal Equipment
TFCI	Transport Format Combination Indicator
TFRC	Transport Format and Resource Combination
TPC	Transmission Power Control
TTI	Transmission Time Interval
UE	User Equipment
UL	Uplink
UMTS	Universal Mobile Telecommunications System
USIM	UMTS subscriber Identity Module
UTRAN	UMTS Terrestrial Radio Access Network
VCO	Voltage Controller Oscillator
VLR	Visitor Location Register
VoIP	Voice-over-IP
WARC	World Administrative Radio Conference
WCDMA	Wideband-Code Division Multiple Access
WLAN	Wireless Local Area Network

Chapter 1

Background

1.1 0G: Mobile Radio Telephone

These systems were the predecessors of the first generation of cellular telephones; this is why they are known as pre cellular or zero generation systems. Maybe, the first prototype of this generation was the Handie Talkie H12-16, used on the Second World War by the US Army. Anyway, this kind of mobile phones cannot being considered like the current mobile phones due to was not possible a channel frequency change automatically while one person was moving; i.e. if that person went out of the coverage offered by the antenna, the call would be missed if the operator did not change the frequency by himself.

Between these systems we can find Push to Talk (PTT), Mobile Telephone System (MTS), Improved Mobile Telephone Service (IMTS), Advanced Mobile Telephone System (AMTS), *Offentlig Landmobil Telefoni* (OLT), Mobile telephony system D (MTD), Autoradiopuhelin (ARP). Next, these systems will be briefly explained.

PTT was a system based on half-duplex communication; i.e. while one person speaks the other listen and is not able to simultaneously talk and hear the other party, emulating walkie-talkie communications.

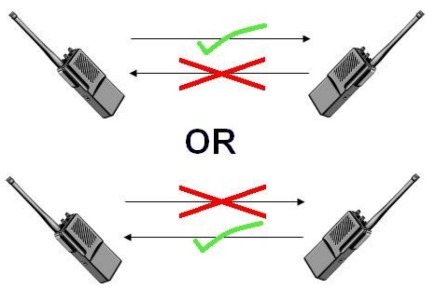

Figure 1.1. Half-Duplex

MTS was operator assisted both directions, meaning that if you were called from a land line the call would be routed to a mobile operator, who would route it to your phone. Similarly, to make an outbound call you had to go through the mobile operator, who would ask you for your mobile number and the number you were calling, and then would place the call.

This system was first used in Saint Louis, Missouri on 17 June, 1946 and was developed by Bell. At first, this service was used for a good number of North American people. Then, it was replaced for IMTS system.

IMTS was a big improvement with regard to MTS system because it made possible a direct-dial instead connections through an operator. Besides this system was full duplex, then users could talk and hear the other person at the same time.

IMTS base station sites generally covered an area 40-60 miles in diameter and it had 7 or 8 channels in larger cities while rural stations had as few as one or two channels. Each telephone connection required the exclusive use of a channel. Then, the most important problem of this system was that it limited the number of subscribers.

AMTS or Advanced Mobile Telephone System was a 0G method of radio communication, mainly used in Japanese portable radio systems. It operated on the 900 MHz band.

OLT (Norwegian for Offentlig Landmobil Telefoni, Public Land Mobile Telephony), was the first land mobile telephone network in Norway. It was established December 1, 1966, and continued until 1990. In 1981, there were 30,000 mobile subscribers, which at the time made this network the largest in the world.

The network operated in the 160 MHz VHF band, using frequency modulation (FM) on 160-162 MHz for the mobile unit, and 168-170 MHz for the base station. Most mobile sets were semi-duplex, but some of the more expensive units were full duplex. Each subscriber was assigned a five digit phone number. In 1976, the OLT system was extended to include UHF bands, incorporating MTD, and allowing international roaming within Scandinavian countries.

MTD (Swedish abbreviation for Mobiltelefonisystem D, or Mobile telephony system D) was a manual mobile phone system for the 450 MHz frequency band. It was introduced in 1971 in Sweden, and lasted until 1987. The MTD network had 20,000 users at its peak, with 700 people employed as phone operators.

ARP (Autoradiopuhelin, "car radio phone") was the first commercially operated public mobile phone network in Finland. The network was proposed in 1968 and building began in 1969. It was launched in 1971, and reached 100% geographic coverage in 1978 with 140 base stations. The ARP network was closed at the end of 2000.

ARP was a success and reached great popularity (10,800 users in the year 1977, with a peak of 35,560 in 1986), but the service eventually became too congested and was gradually replaced by the more modern NMT technology. However, ARP was the only mobile phone network with 100% percent coverage for some time thereafter, and it remained popular in many special user groups.

ARP operated on 150 MHz frequency (80 channels on 147.9 - 154.875 MHz band). Transmission power ranged from 1 watt to 5 watts. It first used only half-duplex transmission. Later, full-duplex car phones were introduced. Being analog, it had no encryption and calls could be listened to with scanners. It started as a manually switched service, but was fully automated in 1990; however, by that time the number of subscribers had dwindled down to 980 users. ARP did not support handover, so calls would disconnect when moving to a new cell area. The cell size was approximately 30 km.

The first ARP mobile terminals were extremely large for the time and could only be fitted in cars' trunks, with a handset near the driver's seat. ARP was also expensive. In the 1990s, handhelds were introduced in ARP but they never became popular as more modern equipment was already available in other systems like NMT.

1.2 1G: First Generation Wireless Telephone Technology

1G is short for first-generation wireless telephone technology, cellphones. These are the analog cellphone standards that were introduced in the 1980s and continued until being replaced by 2G digital cellphones. The main difference between two succeeding mobile telephone systems, 1G and 2G, is that the radio signals that 1G networks use are analog, while 2G networks are digital.

Although both systems use digital signaling to connect the radio towers to the rest of the telephone system, the voice itself during a call is encoded to digital signals in 2G whereas 1G is only modulated to higher frequency, typically 150MHz and up.

One such standard is NMT (Nordic Mobile Telephone), used in Nordic countries, Switzerland, Netherlands, Eastern Europe and Russia. Others include AMPS (Advanced Mobile Phone System) used in the United States and Australia [1], TACS (Total Access Communications System) in the United Kingdom, C-450 in West Germany, Portugal and South Africa , Radiocom 2000 in France, and RTMI in Italy. In Japan there were multiple systems. Three standards, TZ-801, TZ-802, and TZ-803 were developed by NTT, while a competing system operated by DDI used the JTACS (Japan Total Access Communications System) standard.

NMT (Nordic Mobile Telephony) is the first fully-automatic cellular phone system. It was specified by Nordic telecommunications administrations (PTTs) starting in 1970, and opened for service in 1981 as a response to the increasing congestion and heavy requirements of the manual mobile phone networks: ARP (150 MHz) in Finland and MTD (450 MHz) in Sweden, Norway and Denmark.

NMT is based on analog technology (first generation or 1G) and two variants exist: NMT-450 and NMT-900. The numbers indicate the frequency bands uses. NMT-900 was introduced in 1986 because it carries more channels than the previous NMT-450 network.

The cell sizes in an NMT network range from 2 km to 30 km. With smaller ranges the network can service more simultaneous callers; for example in a city the range can be kept short for better service. NMT used full duplex transmission, allowing for simultaneous receiving and transmission of voice. Car phone versions of NMT used transmission power of up to 15 watt (NMT-

450) and 6 watt (NMT-900), handsets up to 1 watt. NMT had automatic switching (dialing) and handover of the call built into the standard from the beginning, which was not the case with most preceding car phone services, such as the Finnish ARP. Additionally, the NMT standard specified billing as well as national and international roaming.

AMPS (Advanced Mobile Phone System) is the analog mobile phone system standard developed by Bell Labs, and officially introduced in the Americas in 1983 and Australia in 1987. AMPS is a first-generation cellular technology that uses separate frequencies, or "channels", for each conversation; that is, Frequency Division Multiple Access (FDMA). It therefore requires considerable bandwidth for a large number of users.

What really separates AMPS from older systems is the "back end" call setup functionality. In AMPS, the cell centers can flexibly assign channels to handsets based on signal strength, allowing the same frequency to be re-used in various locations without interference. This allowed a larger number of phones to be supported over a geographical area. AMPS pioneers fathered the term "cellular" because of its use of small hexagonal "cells" within a system.

The anatomy of each channel is composed of 2 frequencies. 416 of these are in the 824–849 MHz range for transmissions from mobile stations to the base stations, paired with 416 frequencies in the 869–894 MHz range for transmissions from base stations to the mobile stations. Each cell site will use a subset of these channels, and must use a different set than neighboring cells to avoid interference. This significantly reduces the number of channels available at each site in real-world systems. Each AMPS channel is 30 kHz wide.

1.3 2G: Second Generation wireless telephone technology

Second generation 2G cellular telecoms networks were commercially launched on the GSM standard in Finland by Radiolinja (now part of Elisa) in 1991. Three primary benefits of 2G networks over their predecessors were that phone conversations were digitally encrypted, 2G systems were significantly more efficient on the spectrum allowing for far greater mobile phone penetration levels; and 2G introduced data services for mobile, starting with SMS text messages.

2G technologies can be divided into TDMA-based and CDMA-based standards depending on the type of multiplexing used. The main 2G standards are:

- GSM (TDMA-based), originally from Europe but used in almost all countries on all six inhabited continents (Time Division Multiple Access). Today accounts for over 80% of all subscribers around the world.

- IS-95 or cdmaOne, (CDMA-based, commonly referred as simply CDMA in the US), used in America and parts of Asia. Today accounts for about 17% of all subscribers globally. Over a dozen CDMA operators have migrated to GSM including operators in Mexico, India, Australia and South Korea.

1.3.1 GSM Technology

GSM (Global System for Mobile communications) is the most popular standard for mobile phones in the world. Its promoter, the GSM Association, estimates that 82% of the global mobile market uses the standard [2]. GSM is used by over 3 billion people across more than 212 countries and territories [3] [4]. GSM differs from its predecessors in that both signaling and speech channels are digital, and thus is considered a second generation (2G) mobile phone system. This has also meant that data communication was easy to build into the system.

GSM is a cellular network, which means that mobile phones connect to it by searching for cells in the immediate vicinity. GSM networks operate in four different frequency ranges. Most GSM networks operate in the 900 MHz or 1800 MHz bands. The rarer 400 and 450 MHz frequency bands are assigned in some countries, notably Scandinavia, where these frequencies were previously used for first-generation systems.

GSM-900 uses 890–915 MHz to send information from the mobile station to the base station (uplink) and 935–960 MHz for the other direction (downlink), providing 124 RF channels (channel numbers 1 to 124) spaced at 200 kHz. Duplex spacing of 45 MHz is used. Time division multiplexing is used to allow eight full-rate or sixteen half-rate speech channels per radio frequency channel. There are eight radio timeslots (giving eight burst periods) grouped into what is called a TDMA frame. Half rate channels use alternate frames in the same timeslot. The channel data rate for all 8 channels is 270.833 kbit/s, and the frame duration is 4.615 ms.

The transmission power in the handset is limited to a maximum of 2 watts in GSM850/900 and 1 watt in GSM1800/1900.

GSM has used a variety of voice codecs to squeeze 3.1 kHz audio into between 5.6 and 13 kbit/s. Originally, two codecs, named after the types of data channel they were allocated, were used, called Half Rate (5.6 kbit/s) and Full Rate (13 kbit/s). These used a system based upon linear predictive coding (LPC). In addition to being efficient with bitrates, these codecs also made it easier to identify more important parts of the audio, allowing the air interface layer to prioritize and better protect these parts of the signal.

There are five different cell sizes in a GSM network: macro, micro, pico, femto and umbrella cells. The coverage area of each cell varies according to the implementation environment. Macro cells can be regarded as cells where the base station antenna is installed on a mast or a building above average roof top level. Micro cells are cells whose antenna height is under average roof top level; they are typically used in urban areas. Picocells are small cells whose coverage diameter is a few dozen meters; they are mainly used indoors. Femtocells are cells designed for use in residential or small business environments and connect to the service provider’s network via a broadband internet connection. Umbrella cells are used to cover shadowed regions of smaller cells and fill in gaps in coverage between those cells.

Cell horizontal radius varies depending on antenna height, antenna gain and propagation conditions from a couple of hundred meters to several tens of kilometers. The longest distance the GSM specification supports in practical use is 35 kilometers.

The modulation used in GSM is Gaussian minimum-shift keying (GMSK), a kind of continuous-phase frequency shift keying. In GMSK, the signal to be modulated onto the carrier is first smoothed with a Gaussian low-pass filter prior to being fed to a frequency modulator, which greatly reduces the interference to neighboring channels (adjacent channel interference).

The network behind the GSM system seen by the customer is large and complicated in order to provide all of the services which are required. It is divided into a number of sections and these are each covered in separate articles.

- the Base Station Subsystem (the base stations and their controllers).
- the Network and Switching Subsystem (the part of the network most similar to a fixed network). This is sometimes also just called the core network.
- the GPRS Core Network (the optional part which allows packet based Internet connections).

GSM was designed with a moderate level of security. The system was designed to authenticate the subscriber using a pre-shared key and challenge-response. Communications between the subscriber and the base station can be encrypted. The development of UMTS introduces an optional USIM, that uses a longer authentication key to give greater security, as well as mutually authenticating the network and the user, whereas GSM only authenticated the user to the network, and not vice versa. The security model therefore offers confidentiality and authentication, but limited authorization capabilities, and no non-repudiation. GSM uses several cryptographic algorithms for security. The A5/1 and A5/2 stream ciphers are used for ensuring over-the-air voice privacy. A5/1 was developed first and is a stronger algorithm used within Europe and the United States; A5/2 is weaker and used in other countries. Serious weaknesses have been found in both algorithms: it is possible to break A5/2 in real-time with a ciphertext-only attack, and in February 2008, Pico Computing, Inc revealed its ability and plans to commercialize FPGAs that allow A5/1 to be broken with a rainbow table attack [5]. The system supports multiple algorithms so operators may replace that cipher with a stronger one.

1.3.2 EDGE Technology

Enhanced Data rates for GSM Evolution (EDGE), Enhanced GPRS (EGPRS), or IMT Single Carrier (IMT-SC) is a backward-compatible digital mobile phone technology that allows improved data transmission rates, as an extension on top of standard GSM. EDGE can be considered a 3G radio technology and is part of ITU's 3G definition, but is most frequently referred to as 2.75G. EDGE was deployed on GSM networks beginning in 2003— initially by Cingular (now AT&T) in the United States.

EDGE is standardized by 3GPP as part of the GSM family, EDGE is an upgrade that provides a potential three-fold increase in capacity of GSM/GPRS networks. The specification achieves higher data-rates by switching to more sophisticated methods of coding, within existing GSM timeslots. Introducing 8PSK encoding, EDGE is capable of delivering higher bit-rates per radio channel in good conditions.

EDGE can be used for any packet switched application, such as an Internet connection. High-speed data applications such as video services and other multimedia benefit from EGPRS' increased data capacity. EDGE Circuit Switched is a possible future development.

Evolved EDGE was added in Release 7 of the 3GPP standard. This is a further extension on top of EDGE, providing reduced latency and potential speeds of 1Mbit/s by using even more complex coding functions than the 8PSK originally introduced with EDGE.

EDGE/EGPRS is implemented as a bolt-on enhancement for 2G and 2.5G GSM and GPRS networks, making it easier for existing GSM carriers to upgrade to it. EDGE/EGPRS is a superset to GPRS and can function on any network with GPRS deployed on it, provided the carrier implements the necessary upgrade.

Although EDGE requires no hardware or software changes to be made in GSM core networks, base stations must be modified. EDGE compatible transceiver units must be installed and the base station subsystem needs to be upgraded to support EDGE. If the operator already has this in place, which is often the case today, the network can be upgraded to EDGE by activating an optional software feature. Today EDGE is supported by all major chip vendors for GSM. New mobile terminal hardware and software is also required to decode/encode the new modulation and coding schemes and carry the higher user data rates to implement new services.

In addition to Gaussian minimum-shift keying (GMSK), EDGE uses higher-order PSK/8 phase shift keying (8PSK) for the upper five of its nine modulation and coding schemes. EDGE produces a 3-bit word for every change in carrier phase. This effectively triples the gross data rate offered by GSM. EDGE, like GPRS, uses a rate adaptation algorithm that adapts the modulation and coding scheme (MCS) according to the quality of the radio channel, and thus the bit rate and robustness of data transmission. It introduces a new technology not found in GPRS, Incremental Redundancy, which, instead of retransmitting disturbed packets, sends more redundancy information to be combined in the receiver. This increases the probability of correct decoding.

EDGE can carry data speeds up to 236.8 kbit/s (with end-to-end latency of less than 150 ms) for 4 timeslots (theoretical maximum is 473.6 kbit/s for 8 timeslots) in packet mode. This means it can handle four times as much traffic as standard GPRS. EDGE will therefore meets the International Telecommunications Union's requirement for a 3G network, and has been accepted by the ITU as part of the IMT-2000 family of 3G standards.

EDGE Evolution improves on EDGE in a number of ways. Latencies are reduced by lowering the Transmission Time Interval by half (from 20 ms to 10 ms). Bit rates are increased up to 1 MBit/s peak speed and latencies down to 100 ms using dual carriers, higher symbol rate and higher-order modulation (32QAM and 16QAM instead of 8-PSK), and turbo codes to improve error correction. And finally signal quality is improved using dual antennas improving average bit-rates and spectrum efficiency. EDGE Evolution can be gradually introduced as software upgrades, taking advantage of the installed base. With EDGE Evolution, end-users will be able to experience mobile internet connections corresponding to a 500 kbit/s ADSL service.

1.3.3 Interim Standard 95, cdmaOne

Interim Standard 95 (IS-95), is the first CDMA-based digital cellular standard pioneered by Qualcomm. The brand name for IS-95 is cdmaOne. IS-95 is also known as TIA-EIA-95.

It is a 2G Mobile Telecommunications Standard that uses CDMA, a multiple access scheme for digital radio, to send voice, data and signaling data (such as a dialed telephone number) between mobile telephones and cell sites.CDMA or "code division multiple access" is a digital radio system that transmits streams of bits (PN Sequences[1]). CDMA permits several users to share the same frequencies. Unlike TDMA "time division multiple access", a competing system used in 2G GSM, all radios can be active all the time, because network capacity does not directly limit the number of active radios. Since larger numbers of phones can be served by smaller numbers of cell-sites, CDMA-based standards have a significant economic advantage over TDMA-based standards, or the oldest cellular standards that used frequency-division multiplexing.

In North America, the technology competed with Digital AMPS (IS-136, a TDMA technology). It is now being supplanted by IS-2000 (CDMA2000), a later CDMA-based standard.

It is used in the USA, South Korea, Canada, Mexico, India, Israel, New Zealand, Sri Lanka, Venezuela, Brazil, China PRC and Vietnam. In Q4 2007, around 431 Million subscribers worldwide used CDMA CDMA Development Group, which is about 13 % of the 3,300 Million subscribers in the world. For comparison, of the global subscribers, those using GSM and 3GSM total 2,881 Million GSM World or 86% of the worldwide subscriber base. All other technologies, TDMA (D-AMPS), PDC, iDEN and various older analog systems account for less than 1% of the total worldwide subscriber base. Of the CDMA users, a large number of users come from the USA, South Korea, Vietnam, India, Brazil and China PRC).

References

[1] http://www.amta.org.au/

[2] "GSM World statistics". GSM Association (2007). Retrieved on 2008-06-07.

[3] "About GSM Association". GSM Association. Retrieved on 2007-01-08.

[4] "Two Billion GSM Customers Worldwide". 3G Americas (June 13, 2006). Retrieved on 2007-01-08.

[5] http://blog.washingtonpost.com/securityfix/

1. A Pseudo-random number sequence is a sequence of numbers that has been computed by some defined arithmetic process but is effectively a random number sequence for the purpose for which it is required.

Chapter 2

Introduction

2.1 Beginning of Third Generation Technology

Third Generation Technology was developed in order to face up to the new requirements of services what were coming, as high-quality images and video or to provide access to the Web with higher data rates. The data-handling capabilities of second-generation systems are limited and was necessary other mobile technology.

In Europe, a long period of research preceded the selection of third-generation technology. The RACE I (Research of Advanced Communication Technologies in Europe) programme started the basic third-generation research work in 1988. This programme was followed by RACE II, with the development of the CDMA-based Code Division Testbed (CODIT) and TDMA-based Advanced TDMA Mobile Access (ATDMA) air interfaces during 1992-95. In addition, wideband air interface proposals were studied in a number of industrial projects in Europe.

The European research programme Advanced Communication Technologies and Services (ACTS) was launched at the end of 1995 in order to support mobile communication research and development. Within ACTS the Future Radio Wideband Multiple Access System (FRAME) project [1] was set up with the objective of defining a proposal for a UMTS radio access system. The main industrial partners in FRAMES were Nokia, Siemens, Ericsson, France Télécom and CSEM/Pro Telecom, with participation also from several European universities. Based on an initial proposal evaluation phase in FRAMES, a harmonized multiple access platform was defined, consisting of two modes: FMA1, a wideband TDMA [2], and FMA2, a wideband CDMA [3]. The FRAMES wideband CDMA and wideband TDMA proposals were submitted to ETSI (European Telecommunications Standards Institute) candidates for UMTS air interface and ITU IMT-2000 submission.

The proposals for the UMTS Terrestrial Radio Access (UTRA) air interface received by the milestone were grouped into five concept groups in ETSI in June 1997, after their submission and presentation during 1996 and earl 1997.

The following groups were formed:

- Wideband CDMA (WCDMA)
- Wideband TDMA (WTDMA)
- TDMA/CDMA
- OFDMA
- ODMA

All the proposed technologies were basically able to fulfill the UMTS requirements, although it was difficult to reach a consensus on issues such as system capacity, since the results of simulations can vary greatly depending on the assumptions. However, it soon became evident in the selection process that WCDMA and TDMA/CDMA were the main candidates. Also, issues such as the global potential of a technology naturally had an impact in cases where obvious technical conclusions were very limited; in this respect, the outcome of the ARIB (Association of Radio Industries and Businesses of Japan) technology selection in Japan gave support to WCDMA.

ETSI decided between the technologies in January 1998, selecting WCDMA as the standard for the UTRA air interface on the paired frequency bands, i.e. for Frequency Division Duplexing (FDD) operation, and WTDMA/CDMA for operation with unpaired spectrum allocation, i.e. for TDD operation.

It took 10 years from the initiation of European research programs to reach decision of the UTRA technology.

2.2 Creation of 3GPP (Third Generation Pattern Technology)

As similar technologies were being standardized in several regions around the world, it became evident that achieving identical specifications to ensure equipment compatibility globally would be very difficult with work going on parallel. Also, having to discuss similar issues in several places was naturally a waste of resources for the participating companies. Therefore, initiatives were made to create a single forum for WCDMA standardization for a common WCDMA specification.

The standardization organizations involved in the creation of the 3GPP [4] were ARIB (Japan), ETSI (Europe), TTA (Korea), TTC (Japan) and T1P1 (USA). This project was established in December 1998. Later, during 1999, the China Wireless Telecommunication Standard Group (CWTS) also joined 3GPP. It later became CCSA (China Communication Standards Association). The partners agreed on join efforts for the standardization of UTRA, now standing for Universal Terrestrial Radio Access, as distinct from UTRA (UMTS Terrestrial Radio Access) from ETSI, also submitted to 3GPP. Companies such as manufacturers and operators are members of 3GPP through the respective standardization organization to which they belong.

The first full set of 3GPP specifications was called Release 99 due to was completed at the end of 1999. The first commercial network was opened in Japan during 2001 for commercial use in key areas, and in Europe at the beginning of 2002 for the pre-commercial testing phase and for commercial use during 2003.

3GPP specified important evolution steps on top of WCDMA: HSPA for downlink in Release 5 and for uplink in Release 6. The downlink solution, High Speed Downlink Packet Access (HSDPA) was commercially deployed in 2005 and the uplink counterpart, High Speed Uplink Packet Access (HSUPA), during 2007. Further HSPA evolution is specified in 3GPP Release 7, and its commercial deployment is expected by 2009. HSPA evolution is also known as HSPA+. And the newest one is a new radio system called Long-Term Evolution what belongs to Release 8 and will be the replacement for WCDMA in the near future.

The schedule for 3GPP standardisation and for commercial deployment is illustrated in Figure 2.1.

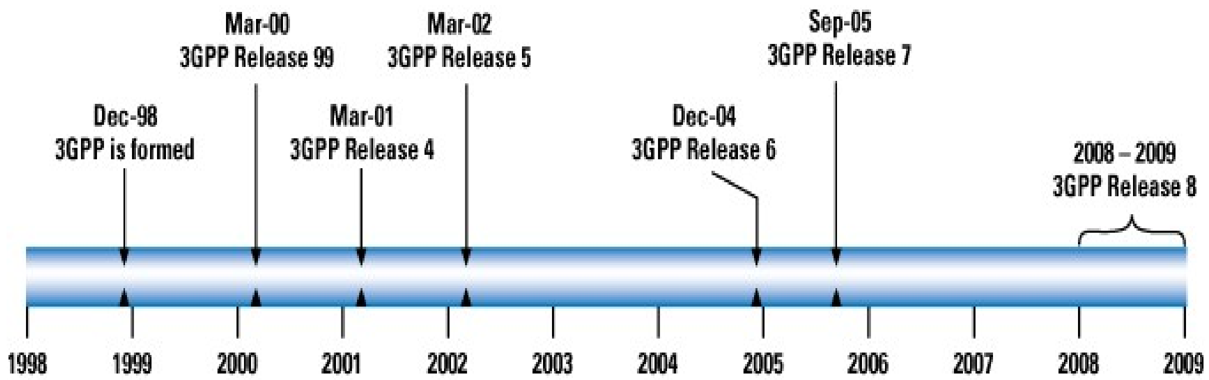

Figure 2.1. Standardisation and commercial operation schedule for WCDMA and its evolution

2.3 Creation of 3GPP2

Work done in TR45.5 and TTA was merge to form 3GPP2, focused on the development of cdma2000 Direct-Sequence (DS) and Multi-Carrier (MC) mode for the cdma2000 third-generation component. This activity has been running in parallel with the 3GPP project, with participation from ARIB, TTC and CWTS as a member organizations. The focus shifted to MC mode after global harmonization efforts, but then later work started to focus further more on the narrowband IS-95 evolution, reflected in the IS-2000 standards series.

2.4 IMT-2000 Process in ITU

In the ITU (International Telecommunications Union), recommendations have been developed for third-generation mobile communications systems, the ITU terminology being called IMT-2000

[5], formerly FPLMTS. In the ITU-R, ITU-R TG8/1 has worked on the radio-dependent aspects, while the radio-independent aspects have been covered in ITU-T SG11.

In the radio aspects, ITU-R TG8/1 received a number of different proposals during the IMT-2000 candidate submission process. In the second phase of the process, evaluation results were received from the proponent organizations as well as from the other evaluation groups that studied the technologies. During the first half of 1999 the recommendation IMT.RKEY was created, which describes the IMT-2000 multimode concept.

The ITU-R IMT-2000 process was finalized at the end of 1999, when the detailed specification (IMT-RSCP) was created and the radio interface specifications were approved by ITU-R. The detailed implementation of IMT-2000 will continue in the regional standard bodies. The ITU-R process has been a important external motivation and timing source for IMT-2000 activities in regional standards bodies. The requirements set by ITU for an IMT-2000 technology have been reflected in the requirements in the regional standards bodies, e.g. in ETSI UMTS 21.01 [6], in order for the ETSI submission to fulfill the IMT-2000 requirements.

The UTRA FDD (WCDMA) and cdma2000 are part of the CDMA interface, as CDMA Direct Spread and CDMA Multi-Carrier respectively. UWC-136 and DECT are part of the TDMA-based interface in the concept, as TDMA Single Carrier and TDMA Multi-Carrier respectively. The TDD part in CDMA consists of UTRA TDD from 3GPP and TD-SCDMA from CWTS. Harmonization has been completed for the FDD part in the CDMA interface, and the harmonization process for the CDMA TDD modes within 3GPP resulted in the 1.28 Mcps being included in the 3GPP Release 4 specifications, completed 03/2001.

2.5 Requirements for Third-Generation Systems

The second-generation systems were built mainly to provide speech services in macro cells. To understand the background to the differences between second-and third-generation systems, we need to look at the new requirements of the third-generation systems, which are listed below:

- Bit rates up to 2 Mbps;
- Variable bit rate to offer bandwidth on demand;
- Multiplexing of services with different quality requirements on a single connection, e.g. speech, video and packet data;
- Delay requirements from delay-sensitive real-time traffic to flexible best-effort packet;
- Quality requirements from 10% frame error rate to 10^{-6} bit error rate;
- Coexistence of second-generation and third-generation systems and inter-system handovers

for coverage enhancements and load balancing;

- Support of asymmetric uplink and downlink traffic, e.g. web browsing causes more loading to downlink than to uplink;

- High spectrum efficiency.

The next table (Table 1.1) lists the main differences between WCDMA/High Speed Packet Access (HSPA) and GSM/Enhanced Data Rates for GSM Evolution (EDGE) networks. The differences reflect the new requirements of the third-generation systems. For example, the larger bandwidth of 5 MHz is needed to support higher bit rates. HSPA Release 7 has also added a Multiple Input Multiple Output (MIMO) multi-antenna solution and higher order modulation 64QAM to support even higher data rates. HSPA pushes more functionalities to the base station and allows flat architecture, which improves the efficiency and the Quality of Service (QoS) capabilities for packet services.

	WCDMA/HSPA	**GSM/EDGE**
Carrier spacing	5 MHz	200 kHz
Frequency reuse factor	1	1-18
Frequency diversity	5 MHz bandwidth gives multipath diversity with Rake receiver	Frequency hopping with frequency diversity
Power control frequency	Up to 1500 Hz	Up to 2 Hz
Circuit and packet switched protocols	Same protocols in radio network	Different protocols
Packet scheduling and retransmission control	In base station (HSPA)	In base station controller
Network architecture	Flat architecture with two network elements in user plane (HSPA Release 7)	Four network elements in user plane
MIMO	2x2 MIMO in downlink	-
Downlink modulation	QPSK, 16QAM, 64QAM (HSPA Release 7)	GMSK, 8PSK

Table 2.1. Main differences between WCDMA and GSM networks

2.6 Spectrum Allocations for Third-Generation Systems

Work to develop third-generation mobile systems started when the World Administrative Radio Conference (WARC) of the international Telecommunications Union (ITU), at its 1992 meeting, identified the frequencies around 2 GHz that were available for use by future International Mobile Telephony 2000 (IMT-2000) mobile systems, both terrestrial and satellite. Within the IMT-2000 framework, five air interfaces are defined for third-generation systems, based on either CDMA or TDMA technology. The original target of the third-generation process was a single global IMT-2000 air interface. In practice, the third-generation systems are closer to this target than were second-generation systems, since WCDMA has clearly turned out to be the most dominant IMT-2000 standard in commercial deployments. The same WCDMA air interface is deployed in Europe, Asia, Australia, in North and South America and in Africa.

Most of the WCDMA deployments use the identified IMT-2000 spectrum around 2GHz: 1920-1980 MHz for uplink and 2110-2170 MHz for downlink. This spectrum is in IMT-2000 use in Europe, Asia (including Japan and Korea) and in Brazil. The first licenses for that spectrum were granted in Finland in March 1999, followed by Spain in March 2000. No auction was conducted in Finland or in Spain. Also, Sweden granted the licenses without auction in December 2000. However, in other countries, such as the UK, Germany and Italy, an auction similar to the US Personal Communication Services (PCS) spectrum was conducted.

WCDMA will also be deployed in the existing second-generation frequency bands that were also identified for IMT-2000 in WRC-2000 and are currently used by GSM or cdma. That approach is called refarming. The WCDMA deployment in the USA started by refarming WCDMA to the existing cellular bands at 850 MHz and to the PCS band at 1900 MHz, since there were no new frequencies available for WCDMA deployment. A new frequency band was auctioned in the USA in 2006. this band, the so-called Advanced Wireless Services (AWS) band, is located at 1700 MHz for uplink and at 2100 MHz for downlink. WCDMA deployment at that band has already started. The AWS band uplink happens to be within GSM1800 uplink band and the downlink is within the UMTS downlink band.

WCDMA refarming to GSM bands has also started in Europe and in Asia. The GSM900 band is attractive, since a lower frequency can provide better coverage than the IMT-2000 band at 2 GHz. The coexistence of GSM and WCDMA in the same frequency band needs to be taken into account in the network planning.

The new IMT-2000 band around 2.6 GHz with a total 190 MHz spectrum will soon be available for the deployment of IMT-2000 and other mobile systems. In Europe, the spectrum includes 2 x 70 MHz for FDD systems and 50 MHz in the middle gap that could be used for TDD deployments due to the operator spectrum being is already more in line with the TDD arrangement.

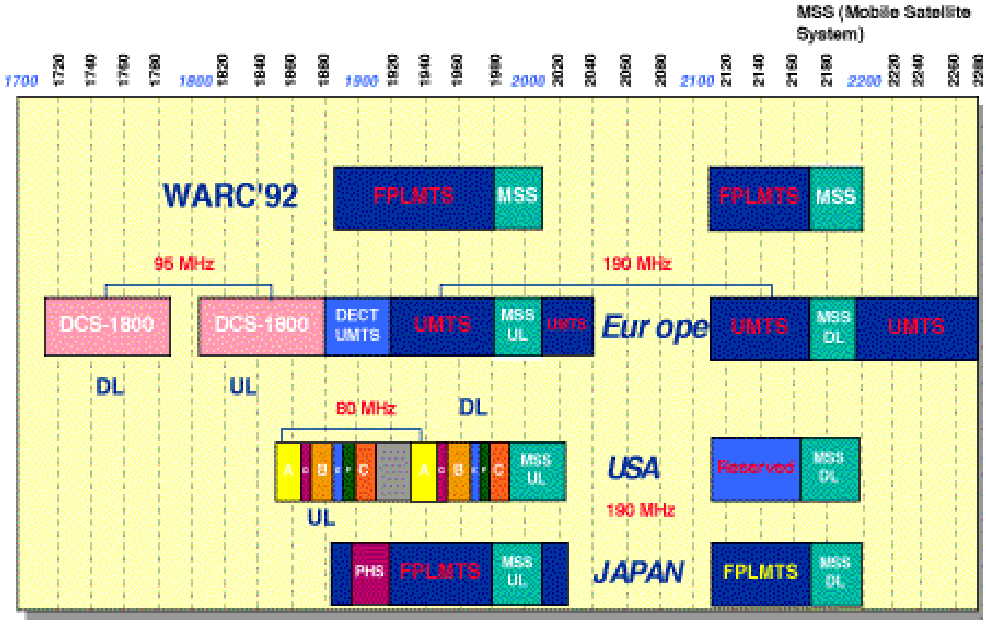

Figure 2.2. Spectrum allocation

References

[1] Nikula, E., Toskala, A., Dahlman, E., Girard, L. and Klein, A., “FRAMES Multiple Access for UMTS and IMT-2000”, IEEE Personal Communications Magazine, April 1998, pp. 16-24

[2] Klein, A., Pirhonen, R., Sköld, J. and Suoranta, R. “FRAMES Multiple Access Mode 1 – Wideband TDMA with and without spreading”, Proc. IEEE Int. Conf. on Personal Indoor and Mobile Radio Communications, PIMRC´97, Helsinki, Finland, 1-4 September 1997, pp. 37-41.

[3] Ovesjö, F., Dahlman, E., Ojanperä, T., Toskala, A. and Klein, A., “FRAMES Multiple Access Mode 2 – Wideband CDMA”, Proc. IEEE Int. Conf. on Personal Indoor and Mobile Radio Communications, PIMRC´97, Helsinki, Finland, 1-4 September 1997, pp. 42-46.

[4] http://www.3GPP.org.

[5] http://www.itu.int/imt/.

[6] Universal Mobile Telecommunications Systems (UMTS), Requirements for the UMTS Terrestrial Radio Access System (UTRA), ETSI Technical Report, UMTS 21.01 version 3.0.1, November 1997

Chapter 3

Third Generation Technology

3.1 Wideband Code Division Multiple Access (WCDMA)

3.1.1 Introduction and summary of the main parameters in WCDMA

This section introduces the principles of the WCDMA air interface. Special attention is drawn to those features by which WCDMA differs from GSM and IS-95.

We present the main system design parameters of WCDMA in this section and give brief explanations for most of them.

- WCDMA is a wideband Direct-Sequence Code Division Multiple Access (DS-CDMA) system, i.e. user information bits are spread over a wide bandwidth by multiplying the user data with quasi-random bits (called chips) derived from CDMA spreading codes. In order to support very high bit rates (up to 2 Mbps), the use of a variable spreading factor and multicode connections is supported.

- The chip rate of 3.84 Mcps leads to a carrier bandwidth of approximately 5 MHz. DS-CDMA systems with a bandwidth of about 1 MHz, such as IS-95, are commonly referred to as narrowband CDMA systems. The inherently wide carrier bandwidth of WCDMA supports high user data rates and also has certain performance benefits, such as increased multipath diversity. Subject to his operating license, the network operator can deploy multiple 5 MHz carriers to increase capacity, possibly in the form of hierarchical cell layers. The actual carrier spacing can be selected on a 200 kHz grid between approximately 4.4 and 5 MHz, depending on interference between the carriers.

- WCDMA supports highly variable user data rates, in other words the concept of obtaining Bandwidth on Demand (BoD) is well supported. The user data rate is kept constant during each 10 ms frame. However, the data capacity among the users can change from the frame to frame. This fast radio capacity allocation will typically be controlled by the network to achieve optimum throughput for packet data services.

- WCDMA supports two basics modes of operation: Frequency Division Duplex (FDD) and Time Division Duplex (TDD). In the FDD mode, separate 5 MHz carrier frequencies are used for the uplink and downlink respectively, whereas in TDD only one 5 MHz is time-shared between the uplink and downlink. Uplink is the connection from the mobile to the base station, and downlink is that from the base station to the mobile.

- WCDMA supports the operation of asynchronous base stations, so that, unlike in the synchronous IS-95 system, there is no need for a global time reference such as a GPS. Deployment of indoor and micro base stations is easier when no GPS signal needs to be received.

- WCDMA employs coherent detection on uplink and downlink based on the use of pilot symbols or common pilot. While already used on the downlink in IS-95, the use of coherent detection on the uplink is new for public CDMA systems and will result in an overall increase of coverage and capacity on the uplink.

- The WCDMA air interface has been crafted in such a way that advanced CDMA receiver concepts, such as multiuser detection and smart adaptive antennas can be deployed by the network operator as a system option to increase capacity and/or coverage. In most second generation systems no provision has been made for such receiver concepts and as a result they are either not applicable or can be applied only under severe constraints with limited increases in performance.

- WCDMA is designed to be deployed in conjunction with GSM. Therefore, handovers between GSM and WCDMA are supported in order to be able to leverage the GSM coverage for the introduction of WCDMA

3.1.2 Generic principles of CDMA operation

Figure 3.1 depicts the basic operations of spreading and despreading for a DS-CDMA system.

User data is here assumed to be a BPSK-modulated bit sequence of rate R, the user data bits assuming the values of +/- 1. The spreading operation, in this example, is the multiplication of each user data bit with a sequence of 8 code bits, called chips. We assume also for the BPSK spreading modulation. We see that the resulting spread data is at a rate of 8 x R and has the same random (pseudo-noise-like) appearance as the spreading code. In this case we would say that a spreading factor of 8 was used. This wideband signal would then be transmitted across a wireless channel to the receiving end.

During despreading we multiply the spread user data/chip sequence, bit duration by bit duration, with the very same 8 code chips as we used during the spreading of these bits. As shown, the original user bit sequence has been recovered perfectly, provided we have (as we shown in figure 3.1) also perfect synchronization between the spread user signal and the (de)spreading code.

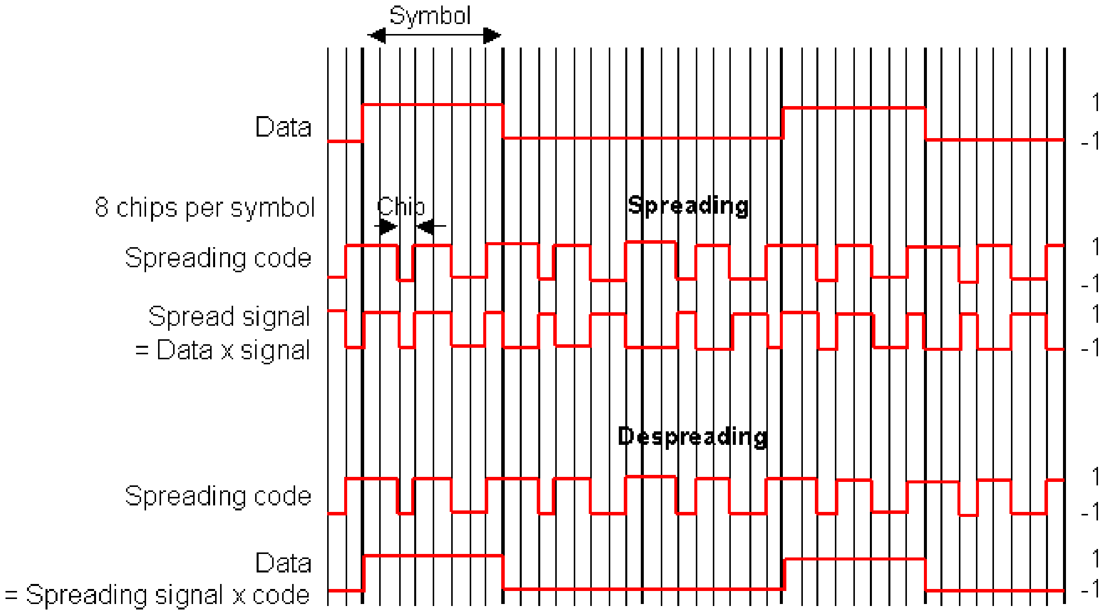

Figure 3.1. Spreading and despreading in DS-CDMA

During despreading we multiply the spread user data/chip sequence, bit duration by bit duration, with the very same 8 code chips as we used during the spreading of these bits. As shown, the original user bit sequence has been recovered perfectly, provided we have (as we shown in figure 3.1) also perfect synchronization between the spread user signal and the (de)spreading code.

The increase of the signaling rate by a factor of 8 corresponds to a widening (by a factor of 8) of the occupied spectrum of the spread user data signal. Due to this virtue, CDMA systems are more generally called spread spectrum systems. Despreading restores a bandwidth proportional to R for the signal.

The basic operation of the correlation receiver for CDMA is shown in Figure 3.2. The upper half of the figure shows the reception of the desired own signal. As in Figure 3.1, we see the despreading operation with a perfectly synchronized code. Then, the correlation receiver integrates (i.e. sums) the resulting products (data x code) for each user bit.

The lower half of Figure 3.2 shows the effect of the despreading operation when applied to the CDMA signal of another user whose signal is assumed to have been spread with a different spreading code. The result of multiplying the interfering signal with the own code and integrating the resulting products leads to interfering signal values lingering around 0.

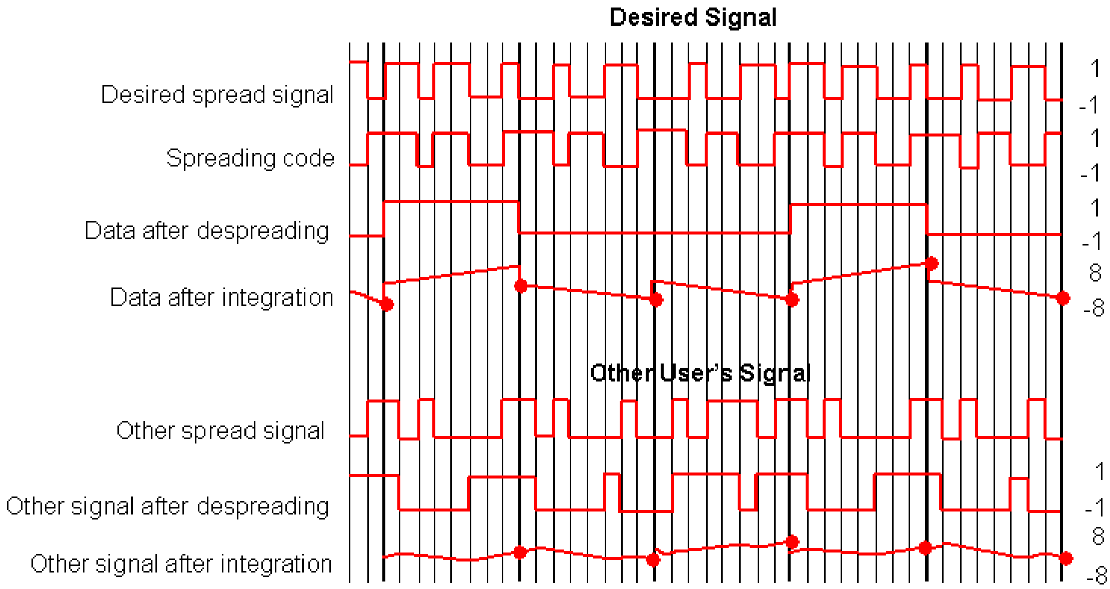

Figure 3.2. Principle of the CDMA correlation receiver

As can be seen, the amplitude of the own signal increases on average by a factor of 8 relative to that of the user of the other interfering system, i.e. the correlation detection has raised the desired user signal by the spreading factor, here 8, from the interference present in the CDMA system. This effect is termed "processing gain" and is a fundamental aspect of all CDMA systems, and in general of all spread spectrum systems. Processing gain is what gives CDMA systems the robustness against self-interference that is necessary in order to reuse the available 5 MHz carrier frequencies over geographically close distances. Let´s take an example with real WCDMA parameters. Speech service with a bit rate of 12.2 kbps has a processing gain of 25 dB = 10 x $\log_{10}$ (3.84e6/12.2e3). After despreading, the signal power needs to be typically few decibels above the interference and noise power. The required power density over the interference power density after despreading is designated as E_b/N_0 in this thesis, where E_b is the energy, or power density, per user bit and N_0 is the interference and noise power density. For speech service E_b/N_0 is typically in the order of 5.0 dB, and the required wideband signal-to-interference ratio is therefore 5.0 dB minus the processing gain = -20.0 dB. In other words, the signal power can be 20 dB under the interference or thermal noise power, and the WCDMA receiver can still detect the signal. The wideband signal-to-interference ratio is also called the carrier-to-interference ratio C/I. Due to spreading and despreading, C/I can be lower in WCDMA than, for example, in GSM. A good quality speech connection in GSM requires C/I = 9-12 dB.

Since the wideband signal can be below the thermal noise level, its detection is difficult without knowledge of the spreading sequence. For this reason, spread spectrum systems originated in military applications where the wideband nature of the signal allowed it to be hidden below the omnipresent thermal noise.

Note that within any given channel bandwidth (chip rate) we will have a higher processing gain for lower user data bit rates than for high bit rates. In particular, for user data bit rates of

2Mbps, the processing gain is less than 2 (=3.84 Mcps/2 Mbps = 1.92 which corresponds to 2.8 dB) and some of the robustness of the WCDMA waveform against interference is clearly compromised.

Both base stations as well as mobiles for WCDMA use essentially this type of correlation receiver. However, due to multipath propagation (and possibly multiple receive antennas), it is necessary to use multiple correlation receivers in order to recover the energy from all paths and/or antennas. Such a collection of correlation receivers, termed “fingers”, is what comprises the CDMA Rake receiver.

3.1.3 Multipath Radio and Rake Reception

Radio propagation in the land mobile channel is characterised by multiple reflections, diffractions and attenuation of the signal energy. These are caused by natural obstacles such as buildings, hills, and so on, resulting in so-called multipath propagation. There are two effects resulting from multipath propagation:

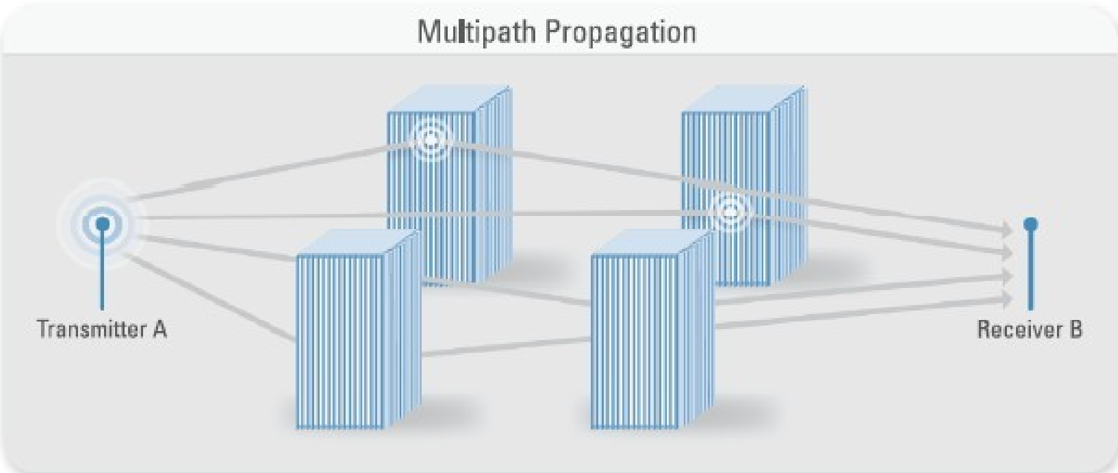

Figure 3.3. Multipath propagation leads to a multipath delay profile

1. The signal energy may arrive at the receiver across clearly distinguishable time instants. The arriving energy is “smeared” into a certain multipath delay profile (Figure 3.3). The delay profile extends typically from 1 to 2 μs in urban and suburban areas, although in some cases delays as long as 20 μs or more with significant signal energy have been observed in hilly areas. The chip duration at 3.84 Mcps is 0.26 μs. If the time difference of the multipath components is at least 0.26 μs, the WCDMA receiver can separate those multipath components and combine them coherently to obtain multipath diversity. The 0.26 μs delay can be obtained if the difference in path lengths is at least 78 m. With a chip rate of about 1 Mcps, the difference in the path lengths of the multipath components must be about 300 m, which cannot be obtained in small cells. Therefore, it is easy to see that the 5 MHz WCDMA can provide multipath diversity in small cells, which is not possible with IS-95.

2. Also, for a certain time delay position there are usually many paths nearly equal in length along which the radio signal travels. For example, paths with a length difference of half a wavelength (at 2GHz this is approximately 7 cm) arrive at virtually the same instant when compared to the duration of a single chip, which is 78 m at 3.84 Mcps. As a result, signal cancellation, called fast fading, takes place as the receiver moves across even short distances. Signal cancellation is best understood as a summation of several weighted phasors that describe the phase shift and attenuation along a certain path at a certain time instant.

Figure 3.4 shows an exemplary fast fading pattern as would be discerned for the arriving signal energy at a particular delay position as the receiver moves. We see that the received signal power can drop considerably (by 25-35 dB) when phase cancellation of multipath reflections occurs. These fading dips make error-free reception of data bits very difficult, and countermeasures are needed in WCDMA. The countermeasures against fading in WCDMA are shown below.

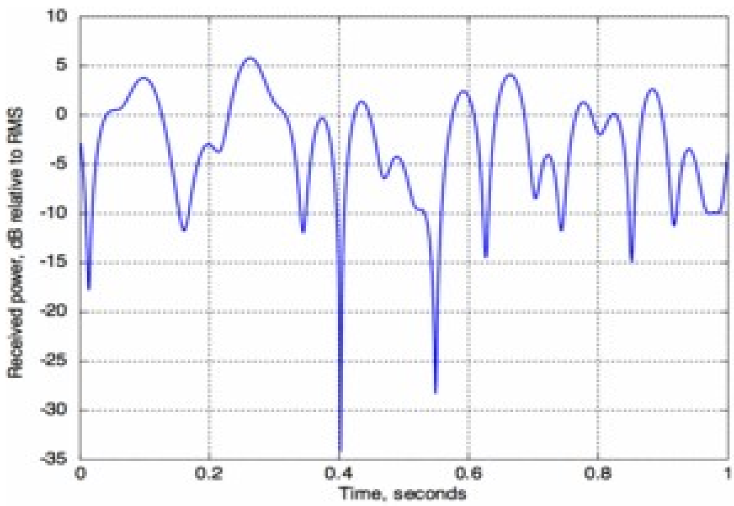

Figure 3.4. Fast Rayleigh fading as caused by multipath propagation

1. The delay dispersive energy is combined by utilizing multiple Rake fingers (correlation receivers) allocated to those delay positions on which significant energy arrives.

2. Fast power control and the inherent diversity reception of the Rake receiver are used to mitigate the problem of fading signal power.

3. Strong coding and interleaving and retransmission protocols are used to add redundancy and time diversity to the signal and thus help the receiver in recovering the user bits across fades.

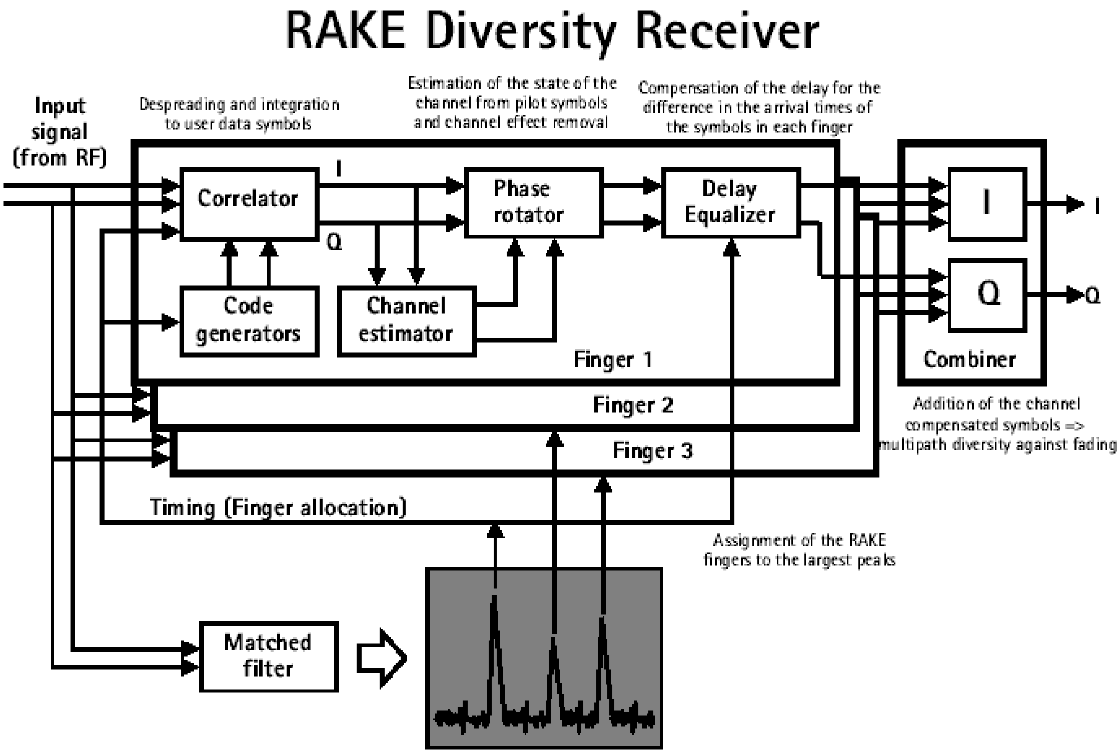

Figure 3.5. Block diagram of the CDMA Rake receiver

Figure 3.5 shows a block diagram of a Rake receiver. Digitized input samples are received from the RF front-end circuitry in the form of I and Q branches (i.e. in complex low-pass number format). Code generators and correlator perfomr the despreading and integration to user data symbols. The channel estimator uses the pilot symbols for estimating the channel state which will then be removed by the phase rotator from the received symbols. The delay is compensated for the difference in the arrival times of the symbols in each finger. The Rake combiner then sums the channel-compensated symbols, thereby providing multipath diversity against fading. Also shown is a matched filter used for determining and updating the current multipath delay profile of the channel. This measured and possibly averaged multipath delay profile is then used to assign the Rake fingers to the largest peaks.

In typical implementation of the Rake receiver, processing at the chip rate (correlator, code generator, matched filter) is done in ASICs, whereas symbol-level processing (channel estimator, phase rotator, combiner) is implemented by a DSP. Although there are several differences between the WCDMA Rake receiver in the mobile and the base station, all the basics principles presented here are the same.

Finally, we note that multiple receive antennas can be accommodated in the same ways as multiple paths received from a single antenna: by just adding additional Rake fingers to the antennas, we can then receive all the energy from multiple paths and antennas. From the Rake receiver´s perspective, there is essentially no difference between these two forms of diversity reception.

3.1.4 Power Control

Tight and fast power control is perhaps the most important aspect in WCDMA, in particular on the uplink. Without it, a single overpowered mobile could block a whole cell. Figure 3.6 depicts the problem and the solution in the form of closed loop transmission power control.

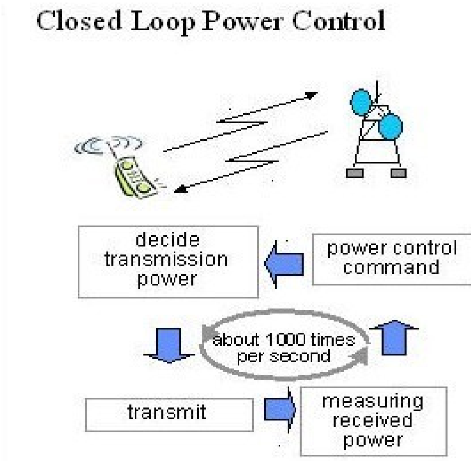

Figure 3.6. Closed loop power control

Let us suppose two mobile stations MS1 and MS2 operate within the same frequency, separable at the base station only by their respective spreading codes. It may happen that MS1 at the cell edge suffers a path loss, for example 70 dB below that MS2, which is near the base station. If there were no mechanism for MS1 and MS2 to be powered-controlled to the same level at the base station, MS2 could easily mask MS1 and thus block a large part of the cell, giving rise to the so-called near-far problem of CDMA. The optimum strategy in the sense of maximizing capacity is to equalize the received power per bit of all mobile stations at all times.

While one can conceive open loop power control mechanisms that attempt to make a rough estimate of path loss by means of a downlink beacon signal, such a method would be far too inaccurate. The prime reason for this is that the fast fading is essentially uncorrelated between uplink and downlink, due to the large frequency separation of the uplink and downlink bands of the WCDMA FDD mode. Open loop power control is, however, used in WCDMA, but only to provide a coarse initial power setting of the mobile station at the beginning of a connection.

The solution to power control in WCDMA is fast closed loop power control, also shown in Figure 3.6. In closed loop power control in the uplink, the base station performs frequent estimates of the received Signal-to-Interference Ratio (SIR) and compares it to a target SIR. If the measured SIR is higher than the target SIR, the base station will command the mobile station to lower the power; if it is too low it will command the mobile station to increase its power. This measure-command-react cycle is executed at a rate of 1500 times per second (1.5 kHz) for each mobile station and thus operates faster than any significant change of path loss could possibly happen and, indeed, even faster than the speed of fast Rayleigh fading for low to moderate mobile speeds. Thus, closed loop power control will prevent any power imbalance among all the uplink signals received at the base station.

The same closed loop power control technique is also used on the downlink, thought here the motivation is different: on the downlink there is no near-far problem due to the one-to-many scenario. All the signals within one cell originate from the one base station to all mobiles. It is, however, desirable to provide a marginal amount of additional power to mobile stations at the cell edge, as they suffer from increased other-cell interference. Also on the downlink a method of enhancing weak signals caused by Rayleigh fading with additional power is needed at low speeds when other error-correcting methods based on interleaving and error correcting codes do not yet work effectively.

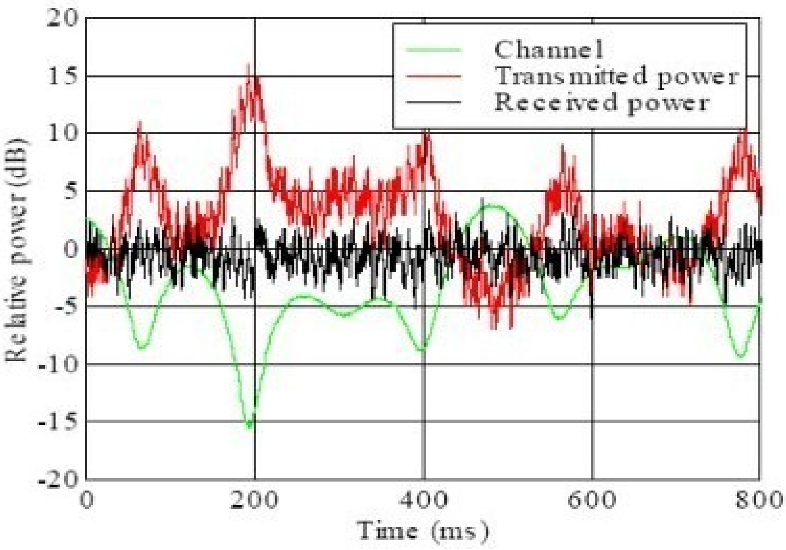

Figure 3.7. Closed-loop power control compensates a fading channel

Figure 3.7 shows how uplink closed loop power control works on a fading channel at low speed. Closed loop power control commands the mobile station to use a transmit power proportional to the inverse of the received in order to compensate a fading channel. Provided the mobile station has headroom to ramp the power up, only very little residual fading is left and the channel becomes essentially non-fading channel as seen from the base station receiver.

While this fading removal is highly desirable from the receiver point of view, it comes at the expense of increased average transmit power at the transmitting end. This means that a mobile station in a deep fade, i.e. using a large transmission power, will cause increased interference to other cells. Figure 3.7 illustrates this point.

Other item related control loop connected with it: outer loop power control. Outer loop power control adjusts the target SIR setpoint in the base station according to the needs of the individual radio link and aims at a constant quality, usually defined as a certain target bit error rate (BER) or block error rate (BLER). Why should there be a need for changing the target SIR setpoint? The required SIR (there exists a proportional E_b/N_0 requirement) for, say, BLER = 1% depends on the mobile speed and the multipath profile. Now, if one were to set the target SIR setpoint for the worst case, i.e. high mobile speeds, one would waste much capacity for those connections at low speeds. Thus, the best strategy is to let the target SIR setpoint float around the minimum value that just fulfills the required target quality. The target SIR setpoint will change over time as the speed and propagation environment changes.

Outer loop control is typically implemented by having the base station tag each uplink user data frame with a frame reliability indicator, such as a cyclic redundancy check (CRC) check result obtained during decoding of that particular user data frame. Should the frame quality indicator indicate to the Radio Network Controller (RNC) that the transmission quality is decreasing, the RNC in turn will command the base station to increase the target SIR setpoint by a certain amount. The reason for having outer loop control reside in the RNC is that this function should be performed after a possible handover combining. Soft handover will be presented in the next section.

3.1.5 Different kinds of Handovers

3.1.5.1 Softer Handover

Strictly speaking softer handover is not really a handover. In this case the User Equipment (UE) combines more than one radio link to improve the reception quality. On the other hand the Node B (Base Transceiver Station (BTS)) combines the data from more than one cell to obtain good quality data from the UE. The maximum number of Radio Links that a UE can simultaneously support as 8. In practice this would be limited to 4 as it is very difficult to make the receiver with 8 fingers.

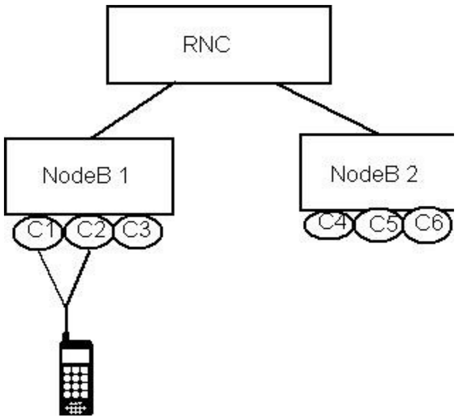

Figure 3.8. Softer handover

Generally speaking when Radio Resource Control (RRC) connection is established, it would always be established on one cell. The network initiates Intra-Frequency measurements to check if there are any other cells the UE can connect simultaneously to improve the quality of the data being transferred between the RNC and the UE. If a suitable cell is found then Active Set Update[1] procedure is initiated. Using this Active Set Update message, the network adds or deletes more than one radio link to the UE. The only requirement is that from the start till the end of this Active Set Update procedure, one Radio Link should remain common. Softer handover typically occurs in about 5-15 % of connections.

1. Active Set is defined as the set of Node-Bs the UE is simultaneously connected to (i.e., the UTRA cells currently assigning a downlink DPCH to the UE constitute the active set).

3.1.5.2 Soft Handover

Soft Handover is the same as softer handover but in this case the cells belong to more than one node B. In this case the combining is done in the RNC. It is possible to simultaneously have soft and softer handovers.

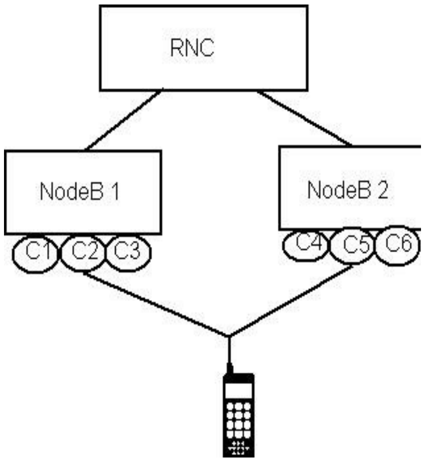

Figure 3.9. Soft handover

A more complicated soft handover would include a cell that belongs to a Node B in different RNC. In this case an Iur connection is established with the drift RNC (RNC 2) and the data would be transferred to the Serving RNC (RNC 1) via Iur connection.

One of the very important requirements for the soft/softer handover is that the frames from different cells should be within 50ms of each other or this would not work. The last thing one needs to remember is that the soft/softer handover is initiated from the RNC and the core network is not involved in this procedure. Soft handovers occurs in about 20-40 % of connections.

3.1.5.3 Hard Handover

Hard handover means that all the old radio links in the UE are removed before the new radio links are established. Hard handover can be seamless or non-seamless. Seamless hard handover means that the handover is not perceptible to the user.

There are two types of hard handover.

- Inter-frequency hard handovers that can be used, for example, to hand a mobile over from one WCDMA frequency carrier to another. One application for this is high capacity base stations with several carriers.

- Inter-system hard handovers that take place between the WCDMA FDD system and another system, such as WCDMA TDD or GSM.

3.2 Radio Access Network Architecture

3.2.1 Introduction to the System Architecture

The UMTS system consists of a number of logical network elements that each has a defined functionality. In the standards, network elements are defined at the logical level, but this quite often results in a similar physical implementation, especially since there are a number of open interfaces (for an interface to be "open", the requirement is that it has been defined to such a detailed level that the equipment at the endpoints can be from two different manufacturers). The network elements can be grouped based on similar functionality, or based on which sub-network they belong to.

Functionally these elements are grouped into the Radio Access Network (RAN, UMTS Terrestrial RAN - UTRAN) and the Core Network (CN). The UTRAN handles all radio-related functionality. Whereas, the CN is responsible for switching and routing calls and data connections to external networks. The system is completed by the User Equipment (UE) or 3G terminal, which interfaces with the user and the radio interface. The high-level architecture is shown in Figure 3.10.

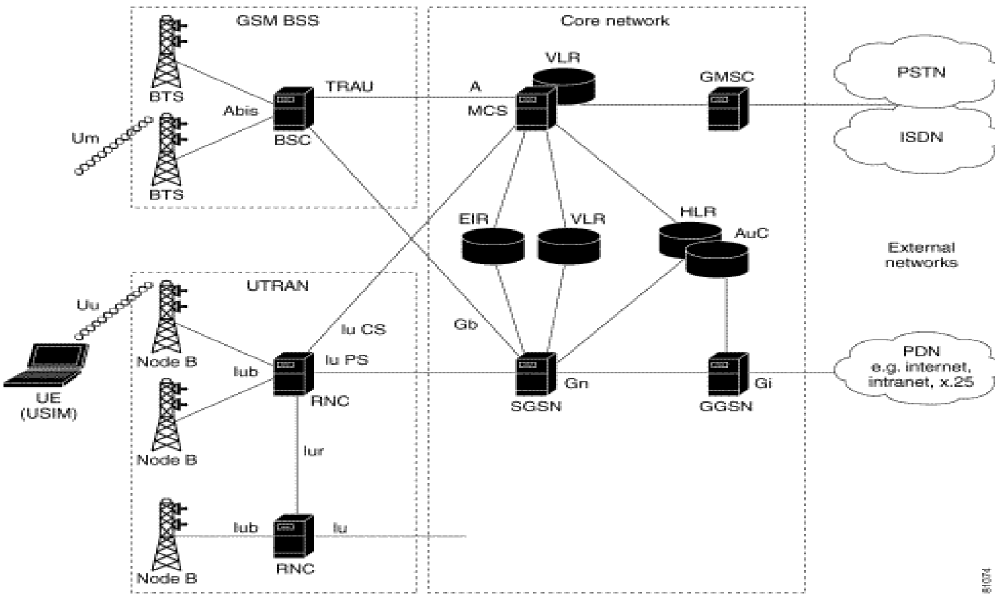

Figura 3.10. UMTS System Architecture

From a specification and standardisation point of view, both UE and UTRAN consist of completely new protocols, the designs of which are based on the needs of the new WCDMA radio technology. On the contrary, the definition of CN is adopted from GSM. This gives the system with new radio technology a global base of known and rugged CN technology that accelerates and facilities its introduction, and enables such competitive advantages as global roaming.

Another way to group UMTS network elements is to divide them into sub-networks. The UMTS is modular in the sense that it is possible to have several network elements of the same type.

In principle, the minimum requirement for a fully featured and operational network is to have at least one logical network element of each type. The possibility of having several entities of the same type allows the division of the UMTS into sub-networks that are operational either on their own or together with other sub-networks, and that are distinguished from each other with unique identities. Such a sub-network is called a UMTS Public Land Mobile Network (PLMN). Tipically, one PLMN is operated by a single operator, and is connected to other PLMNs as well as to other types of network, such as ISDN, PSTN, the internet, and so on. Figure 3.10 shows elements in a PLMN and, in order to illustrate the connections, also external networks.

3.2.2 User Equipment (UE)

The UE consists of ME (Mobile Equipment) and USIM (UMTS subscriber Identity Module). UE domain can be subdivided into ME domain and USIM domain (see Figure 3.11).

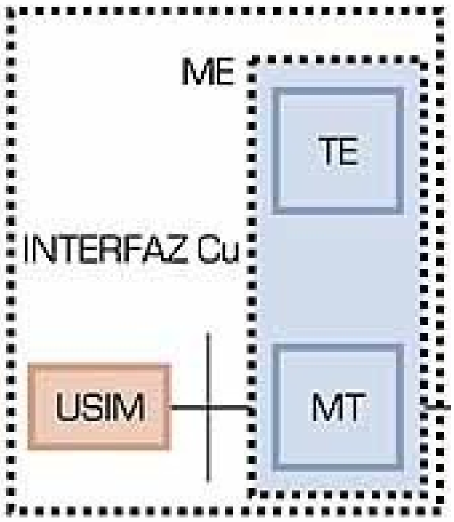

Figure 3.11. User Equipment

The USIM is a smartcard that holds the subscriber identity, performs authentication algorithms, and stores authentication and encryption keys and some subscription information that is needed at the terminal.

The ME is the radio terminal used for radio communication over the Uu interface. It consists of Terminal Equipment (TE), that contains the end-to-end applications functionality (e.g. microphone, loudspeaker); and Mobile Termination (MT), that contains the radio transmission functionality.

On the other hand, Cu interface is the electrical interface between the USIM smartcard and the ME. The interface follows a standard format for smartcards.

The Uu is the interface through which the UE accesses the fixed part of the system and, therefore, is probably the most important open interface in UMTS. There are likely to be many more UE manufacturers than manufacturers of fixed network elements.

3.2.3 UTRAN Architecture

The UTRAN architecture is depicted in Figure 3.12.

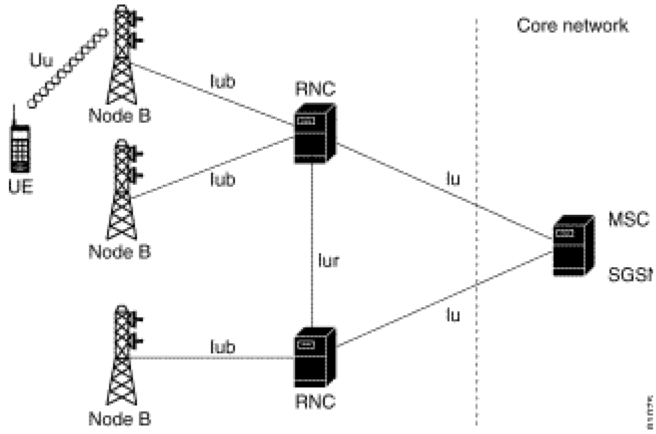

Figure 3.12. UTRAN architecture

UTRAN consists of one or more Radio Network Sub-systems (RNSs). An RNS is a sub-network within UTRAN and consists of one RNC and one or more Node Bs. RNCs may be connected to each other via an Iur interface. RNCs and Node Bs are connected with an Iub Interface. During Release 7, work study on the support of small RNSs was done, meaning use of co-located RNC and Node B functionalities in a flat architecture, and that was found feasible without mandatory specification changes.

3.2.3.1 The Node B

The Node B converts the data flow between the Iub and Uu interfaces. It also participates in radio resource management. It logically corresponds to GSM Base Station but the term "Node B" was initially adopted as a temporary term during the standardization process and then never changed.

Traditionally, the Node Bs have minimum functionality, and are controlled by an RNC. However, this is changing with the emergence of High Speed Downlink Packet Access (HSDPA), where some logic (e.g. retransmission) is handled on the Node B for lower response times.

The utilization of WCDMA technology allows cells belonging to the same or different Node Bs and even controlled by different RNC to overlap and still use the same frequency (the effect is utilized in soft handovers).

Since WCDMA often operates at higher frequencies than GSM, the cell range is considerably smaller compared to GSM cells, and, unlike in GSM, the cells' size is not constant (a phenomenon known as "cell breathing"). This requires a larger number of Node Bs and careful planning in 3G networks. Power requirements on Node Bs and UE are much lower.

A full setup used to contains: a cabinet, an antenna mast and actual antenna. An equipment cabinet contain, for instance, power amplifiers, digital signal processors, back-up batteries and air conditioner equipments. A Node B can serve several cells, also called sectors, depending on the configuration and also on the type of antenna. The most common configuration includes omni cell (360°), 3 sectors (3x120°) or 6 sectors (3 sectors 120° wide overlapping with 3 sectors of different frequency).

3.2.3.2 The Radio Network Controller (RNC)

The RNC is the network element responsible for the control of the radio resources of UTRAN. It interfaces the CN (normally to one MSC and one SGSN) and also terminates the Radio Resource Control (RRC) protocol that defines the messages and procedures between the mobile and UTRAN. It logically corresponds to the GSM BSC.

The RNC controlling one Node B (i.e. terminating the Iub interface towards the Node B) is indicated as the Controlling RNC (CRNC) of the Node B. The CRNC is responsible for the load and congestion control of its own cells, and also executes the admission control and code allocation for new radio links to be established in those cells.

In case one mobile-UTRAN connection uses resources from more than one RNS, the RNCs involved have two separate logical roles:

- Serving RNC (SRNC). The SRNC for one mobile is the RNC that terminates both the Iu link for the transport of user data and the corresponding RAN application part (RANAP) signalling to/from the CN (this connection is referred to as the RANAP connection). The SRNC also terminates the RRC1 Signalling, i.e. the signalling protocol between the UE and UTRAN. It performs the L2 processing of the data to/from the radio interface. Basic Radio Resource Management operations, such as the mapping of Radio Access Bearer (RAB) parameters into air interface transport channel parameters, the handover decision, and outer loop power control, are executed in the SRNC. The SRNC may also be the CRNC of some Node B used by the mobile for connection with UTRAN. One UE connected to UTRAN has one and only one SRNC.

- Drift RNC (DRNC). The DRNC is any RNC, other than the SRNC, that controls cells used by the mobile. If needed, the DRNC may perform macrodiversity combining and splitting. The DRNC does not perform L2 processing of the user plane data, but routes the data transparently between the Iub and Iur interfaces, except when the UE is using a common or shared transport channel. One UE may have zero, one or more DRNCs.

Note that one physical RNC normally contains all the CRNC, SRNC and DRNC functionality.

3.2.3.3 UTRAN Interfaces

The following open interfaces are the UTRAN interfaces:

- Iub interface. The Iub connects a Node B and a RNC. UMTS is the first commercial mobile telephony system where the Controller-Base Station interface is standardised as a fully open interface. Like the other open interfaces, open Iub is expected to motivate further competition between manufacturers in this area. It is likely that new manufacturers concentrating exclusively on Node Bs will enter the market.

- Iur interface. The open Iur interface allows soft handover between RNCs from different manufacturers and, therefore, complements the open Iu interface.

- Iu interface. This connects UTRAN to the CN and is similar to the corresponding interfaces in GSM, A (CS) and Gb (PS), the open Iu interface gives UMTS operators the possibility of acquiring UTRAN and CN from different manufacturers. The enabled competition in this area has been one of the success factors of GSM.

3.2.4 GSM Core Network architecture

Next the main elements of the GSM CN (Figure 3.13) will be briefly described:

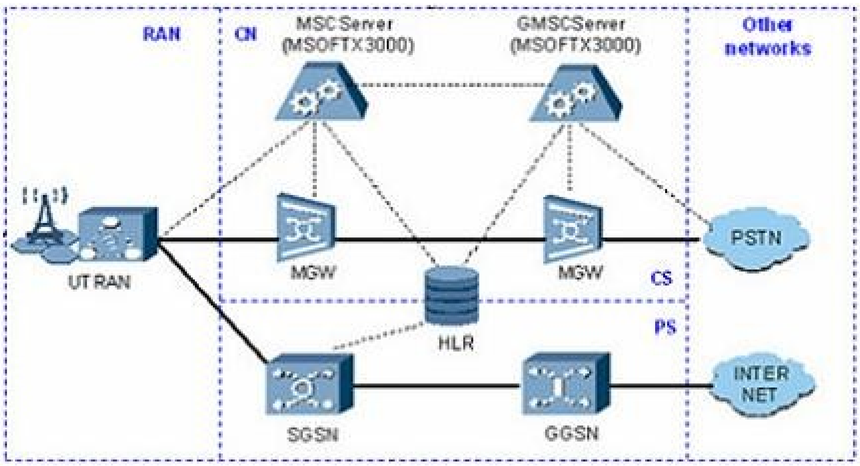

Figure 3.13. GSM Core Network

- Home Location Register (HLR) is a database located in the user´s home system that stores the master copy of the user´s service profile. The service profile consists of, for example, information on allowed services, forbidden roaming areas, and Supplementary Service information such as status of call forwarding and the call forwarding number. It is created when a new user subscribes to the system, and remains stored as long as the subscription is active. For the purpose of routing incoming transactions to the UE (e.g. calls or short messages), the HLR also stores the UE location on the level of MSC/VLR and/or SGSN, i.e. on the level of serving systems.

- Mobile Services Switching Centre/Visitor Location Register (MSC/VLR) is the switch

(MSC) and database (VLR) that serves the UE in its current location for Circuit-Switched (CS) services. The MSC function is used to switch the CS transactions, and the VLR function holds a copy of the visiting user´s service profile, as well as more precise information on the UE´s location within the serving system. The part of the network that is accessed via the MSC/VLR is often referred to as the CS domain.

- Gateway MSC (GMSC) is the switch at the point where UMTS PLMN is connected to external CS networks. All incoming and outgoing CS connections go through GMSC.

- Serving General Packet Radio Service (GPRS) Support Node (SGSN) functionality is similar to that of MSC/VLR but is typically used for Packet-Switched (PS) services. The part of the network that is accessed via the SGSN is often referred to as the PS domain.

- Gateway GPRS Support Node (GGSN) functionality is close to that of GMSC but is in relation to PS services.

3.2.5 External networks

The external networks can be divided into two groups:

1. CS networks. These provide circuit-switched connections, like the existing telephony service. ISDN and PSTN are examples of CS networks.

2. PS networks. These provide connections for packet data services. The internet is one example of a PS network.

3.2.6 General Protocol Model for UTRAN Terrestrial Interfaces

Protocol structures in UTRAN terrestrial interfaces are designed according to the same general protocol model. This model is shown in Figure 3.14. The structure is based on the principle that the layers and planes are logically independent of each other and, if needed, parts of the protocol structure may be changed in the future while other parts remain intact.

3.2.6.1 Horizontal Layers

The protocol structure consists of two main layers, the Radio Network Layer and the Transport Network Layer. All UTRAN-related issues are visible only in the Radio Network Layer, and the Transport Network Layer represents standard transport technology that is selected to be used for UTRAN but without any UTRAN-specific changes.

3.2.6.2 Vertical Planes

Control Plane is used for all UMTS-specific control signalling. It includes the Application

Protocol (i.e. RANAP in Iu, Radio Network System Application Part (RNSAP) in Iur and Node B Application Part (NBAP) in Iub), and the Signalling Bearer for transporting the Application Protocol messages.

The Application Protocol is used, among other things, for setting up bearers to the UE (i.e. the RAB (Radio Access Bearer) in Iu and subsequently the Radio Link in Iur and Iub). In the three-plane structure the bearer parameters in the Application Protocol are not directly tied to the User Plane technology, but rather are general bearer parameters.

The signalling Bearer for the Application Protocol may or may not be of the same type as the Signalling Bearer for the ALCAP (Access Link Control Application Part). It is always set up by operation and maintenance (O&M) actions.

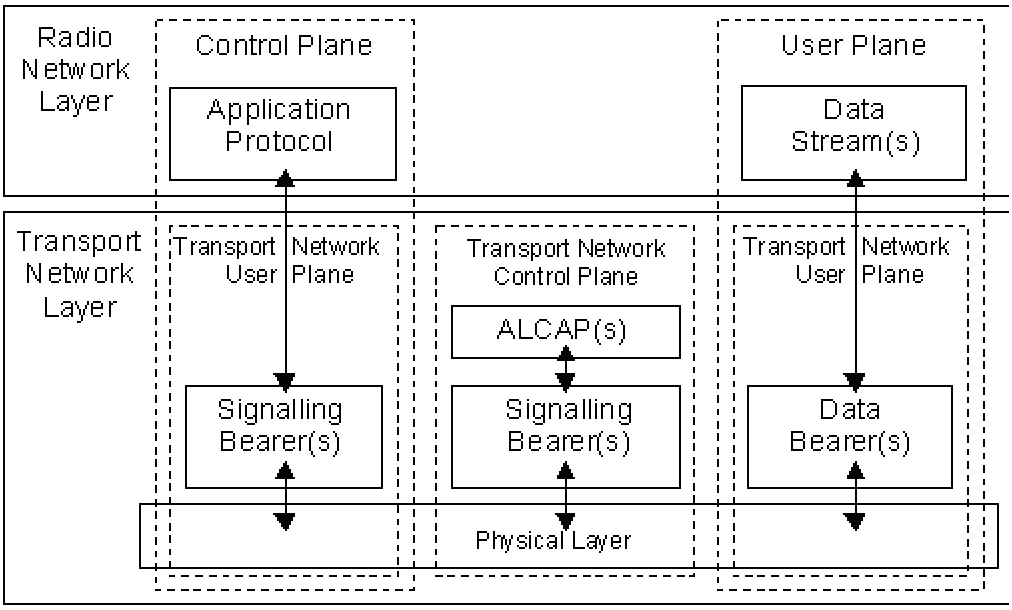

Figure 3.14. General protocol model for UTRAN terrestrial interfaces

User Plane. All information sent and received by the user, such as the coded voice in a voice call or the packets in an internet connection, are transported via the User Plane. The User Plane includes the Data Stream(s), and the Data Bearer(s) for the Data Stream(s). Each Data Stream is characterised by one or more frame protocols specified for that interface..

Transport Network Control Plane. It is used for all control signalling within the Transport Layer. It does not include any Radio Network Layer information. It includes the ALCAP protocol that is needed to set up the transport bearers (Data Bearer) for the User Plane. It also includes the Signalling Bearer needed for the ALCAP.

The Transport Network Control Plane is a plane that acts between the Control Plane and the User Plane. The introduction of the Transport Network Control Plane makes it possible for the Application Protocol in the Radio Network Control Plane to be completely independent of the technology selected for the Data Bearer in the User Plane

When the Transport Network Control Plane is used, the transport bearers for the Data Bearer in the User Plane are set up in the following fashion. First, there is a signalling transaction by the Application Protocol in the Control Plane, which triggers the setup of the Data Bearer by the ALCAP protocol that is specific for the User Plane technology.

The independence of the Control Plane and the User Plane assumes that an ALCAP signalling transaction takes place. It should be noted that ALCAP might not be used for all types of Data Bearers. If there is no ALCAP signalling transaction, then the Transport Network Control Plane is not needed at all. This is the case when it is enough simply to select the user plane resources e.g. selecting end-point addresses for IP transport or selecting a preconfigured Data Bearer. It should also be noted that the ALCAP protocol(s) in the Transport Network Control Plane is/are not used for setting up the Signalling Bearer for the Application Protocol or for the ALCAP during real-time operation.

The Signalling Bearer for the ALCAP may or may not be of the same type as that for the Application Protocol. The UMTS specifications assume that the Signalling Bearer for ALCAP is always set up by O&M actions,and do not specify this in detail.

Transport Network User Plane. The Data Bearer(s) in the User Plane and the Signalling Bearer(s) for the Application Protocol also belong to the Transport Network User Plane. As described in the previous section, the Data Bearers in the Transport Network User Plane are directly controlled by the Transport Network Control Plane during real-time operation, but the control actions required for setting up the Signalling Bearer(s) for the Application Protocol are considered O&M actions.

Chapter 4

The path towards Fourth Generation

4.1 High-Speed Downlink Packet Access (HSDPA)

4.1.1 Introduction to HSDPA

This section presents High-Speed Downlink Packet Access (HSDPA) for Wideband Code Division Multiple Access (WCDMA), the key new feature included in the Release 5 specifications. The HSDPA concept has been designed to increase downlink packet data throughput by means of fast physical layer (L1) retransmission and transmission combining, as well as fast link adaptation controlled by the Node B.

Various methods for packet data transmission in WCDMA downlink already exist in Release 99. The three different channels in Release 99/Release 4 WCDMA specifications that can be used for downlink packet data are: Dedicated Channel (DCH), Downlink-shared Channel (DSCH) and Forward Access Channel (FACH).

The DCH can be used basically for any type of service, and it has a fixed spreading factor (SF) in the downlink. It is always a bi-directional channel with both uplink and downlink connections. Because of the feedback channel, fast power control and soft handover can be used. These features improve their radio performance, and consequently less interference is generated than with common channels. No breaks occur due to mobility since soft handovers is used. On the other hand, setting up a dedicated channel takes more time than accessing common channels. Then the DCH set-up time should preferably be minimised to improve the response times and to avoid the potential TCP (Transport Control Protocol) timeout risk during DCH set-up. The dedicated channel can have bit rates from a few kbps up to 384 kbps in the first mobiles, and up to 2 Mbps according to 3GPP. The bit rate can be changed during transmission.

The DSCH, in contrast to DCH (or FACH), has a dynamically varying SF informed on a 10 ms frame-by-frame basis with the Transport Format Combination Indicator (TFCI) signalling carried on the associated DCH. The DSCH code resources can be shared between several users and the channel may employ either single-code or multi-code transmission. The DSCH may be fast power-controlled with the associated DCH but does not support soft handover. The associated DCH can be in soft handover, e.g. speech is provided on DCH if present with packet data. However, 3GPP recognised that HSDPA was such a major step that motivation for DSCH was no longer there, so it was agreed to remove DSCH from the 3GPP specifications from Release 5 onwards.

The FACH, carried on the secondary common control physical channel (S-CCPCH), can be used to downlink packet data as well. The FACH is operated normally on its own, and it is sent with a fixed SF and typically at rather high-power level to reach all users in the cell owing to the lack of physical layer feedback in the uplink. There is no fast power control or soft handover for FACH. FACH cannot be used in cases in which a simultaneous speech and packet data service is required.

The HSDPA is operated similar to DSCH together with DCH, which carries the services with tighter delay constraints, such as AMR speech. To implement the HSDPA feature, three new channels are introduced in the physical layer specifications [1]:

- High Speed Downlink-shared Channel (HS-DSCH) carries the user data in the downlink direction, with the peak rate reaching up the 10 Mbps region with 16 quadrature amplitude modulation (QAM).
- High-Speed Shared Control Channel (HS-SCCH) carries the necessary physical layer control information to enable decoding of the data on HS-DSCH and to perform the possible physical layer combining of the data sent on HS-DSCH in case of retransmission or an erroneous packet.
- Uplink High-Speed Dedicated Physical Control Channel (HS-DPCCH) carries the necessary control information in the uplink, namely ARQ[1] acknowledgments (both positive and negative ones) and downlink quality feedback information.

A comparison of the basic properties and components of HS-DSCH and DCH is conducted in Table 4.1

1.ARQ is an error control method for data transmission which uses acknowledgments and timeouts to achieve reliable data transmission. An acknowledgment is a message sent by the receiver to the transmitter to indicate that it has correctly received a data frame or packet. A timeout is a reasonable point in time after the sender sends the frame/packet; if the sender does not receive an acknowledgment before the timeout, it usually re-transmits the frame/packet until it receives an acknowledgment or exceeds a predefined number of re-transmissions.

Feature	DCH	HS-DSCH
Soft handover	Yes	No
Fast power control	Yes	No
AMC	No	Yes
Multi-code operation	Yes	Yes, extended
Fast L1 Hybrid ARQ (HARQ)	No	Yes
BTS scheduling	No	Yes

Table 4.1. Comparison of fundamental properties of DCH and HS-DSCH

With HSDPA, two of the most fundamental features of WCDMA, variable SF and fast power control, are disabled and replaced by means of adaptive modulation and coding (AMC), extensive multi-code operation and a fast and spectrally efficient retransmission strategy. In the downlink, WCDMA power control dynamics is in the order of 20 dB, compared with the uplink power control dynamics of 70 dB. The downlink dynamics is limited by the intra-cell interference (interference between users on parallel code channels) and by the Node B implementation. This means that, for a user close to the Node B, the power control cannot reduce power maximally; on the other hand, reducing the power to beyond 20 dB dynamics would have only marginal impact on the capacity. With HSDPA, this property is now utilised by the link adaptation function and AMC to select a coding and modulation combination that requires higher E_c/I_{or}, which is available for the user close to Node B (or with good interference/channel conditions in a short-term sense). This leads to additional user throughput, basically for free. To enable a large dynamic range of the HSDPA link adaptation and to maintain a good spectral efficiency, a user may simultaneously utilise up to 15 multi-codes in parallel. The use of more robust coding, fast HARQ[2] and multi-code operation removes the need for variable SF.

To allow the system to benefit from the short-term variations, the scheduling decisions are done in the Node B. The idea in HSDPA is to enable a scheduling such that, if desired, most of the cell capacity may be allocated to one user for a very short time, when conditions are favorable. In the optimum scenario, the scheduling is able to track the fast fading of the users.

As a conclusion, the results presented in Reference [2] compared the HSDPA cell packet data throughput against Release 99 DSCH performance as presented, and the conclusions drawn were that HSDPA increased the cell throughput up to 100 % compared with the Release 99.

---

2. A variation of ARQ is Hybrid ARQ (HARQ) which has better performance, particularly over wireless channels, at the cost of increased implementation complexity.

4.1.2 HSDPA Terminal Capability and Achievable Data Rates

The HSDPA feature is optional for terminals in Release 5 with a total of 12 different categories of terminals (from physical layer point of view) with resulting maximum data rates ranging between 0.9 to 14.4 Mbps. The HSDPA capability is otherwise independent from Release 99-based capabilities; but, if HS-DSCH has been configured for the terminal then DCH capability in the downlink is limited to the value given by the terminal. A terminal can indicate 32, 64, 128 or 384 kbps DCH capability.

The terminal capability classes are shown in Table 4.2.

Category	Maximum number of received codes	Minimum TTI	Maximum number of received bits/TTI	Maximum H-ARQ buffer size	16QAM compatibility	Maximum throughput (Mbit/s)
Category 1	5	3	7298	19,200	Yes	1.2
Category 2	5	3	7298	28,800	Yes	1.2
Category 3	5	2	7298	28,800	Yes	1.8
Category 4	5	2	7298	38,400	Yes	1.8
Category 5	5	1	7298	57,600	Yes	3.6
Category 6	5	1	7298	67,200	Yes	3.6
Category 7	10	1	14,411	115,200	Yes	7.2
Category 8	10	1	14,411	134,400	Yes	7.2
Category 9	15	1	20,251	172,800	Yes	10.2
Category 10	15	1	27,952	172,800	Yes	14.0
Category 11	5	2	3630	14,400	No	0.9
Category 12	5	1	3630	28,800	No	1.8

Table 4.2. HSDPA terminal capability categories

The first 10 HSDPA terminal capability categories need to support 16 QAM, but the last two, categories 11 and 12, support only QPSK modulation. The differences between classes lie in the maximum number of parallel codes that must be supported and wether the reception in every 2 ms TTI (Transmission Time Interval) is required. The highest HSDPA class supports 10 Mbps. Besides the values indicated in Table 4.2, there is the soft buffer capability with two principles used for determining the value for soft buffer capability. The specifications indicate the absolute values, which should be understood in the way that a higher value means support for incremental redundancy at maximum data rate and a lower value permits only soft combining at full rate.

Category number 10 is intended to allow the theoretical maximum data rate of 14.4 Mbps, permitting basically the data rate that is achievable with rate 1/3 turbo coding and significant puncturing resulting in the code rate close to 1. For category 9, the maximum turbo-encoding block size (from Release 99) resulting in the 10.2 Mbps peak user data rate value with four turbo-encoding blocks. It should be noted that, for HSDPA operation, the terminal will not report individual values, only the category.

Currently, in the market place, different UE vendors are offering devices that typically support either 1.8 Mbps, 3.6 Mbps or even 7.2 Mbps that was available in mid 2007.

Besides the parameters part of the UE capability, the terminal data rate can be largely varied by changing the coding rate as well. Table 4.3 shows the achievable data rates when keeping the number of codes constant (15) and changing the coding rate as well as the modulation. Table 4.3 shows some example bit rates without overhead considerations for different transport format and resource combinations (TFRCs).

TFRC	Modulation	Effective code rate	Max. Throughput (Mbps)
1	QPSK	1/4	1.8
2	QPSK	2/4	3.6
3	QPSK	3/4	5.3
4	16QAM	2/4	7.2
5	16QAM	3/4	10.7

Table 4.3. Theoretical bit rates with 15 multi-codes for different TFRCs

These theoretical data rates can be allocated for a single user or divided between several users. This way, the network can match the allocated power/code resources to the terminal capabilities and data requirements of the active terminals. In contrast to Release 99 operation, it is worth noting that the data rate negotiated with the core network is typically smaller than the peak data rate used in the air interface. Thus, even if the maximum data rate negotiated with the core network were, for example, 1Mbps or 2 Mbps, the physical layer would use (if conditions permit) a peak data rate of, for example, 3.6 Mbps.

4.1.3 Mobility with HSDPA

The mobility procedures for HSDPA users are affected by the fact that transmission of the HS-PDSCH (High Speed Physical downlink shared channel) and the HS-SCCH to a user belongs to only one of the radio links assigned to the UE, namely the serving HS-DSCH cell. UTRAN determines the serving HS-DSCH cell for an HSDPA-capable UE, just as it is the UTRAN that selects the cells in a certain user´s active set for DCH transmission/reception. Synchronised change of the serving HS-DSCH cell is supported between UTRAN and the UE, so that start and stop of transmission and reception of the HS-PDSCH and the HS-SCCH are done at a certain time dictated by the UTRAN. This allows implementation of HSDPA with full mobility and coverage to exploit fully the advantages of this scheme over Release 99 channels.

Next, different possibilities of handovers will be briefly discuss.

4.1.3.1 Intra-Node B HS-DSCH to HS-DSCH Handover

Once the SRNC decides to make an intra-Node B handover from a source HS-DSCH cell to a new target HS-DSCH cell under the same Node B as ilustrated on Figure 4.1, the SRNC sends a synchronised radio link reconfiguration prepare message to the Node B, as well as a radio resource control (RRC) physical channel reconfiguration

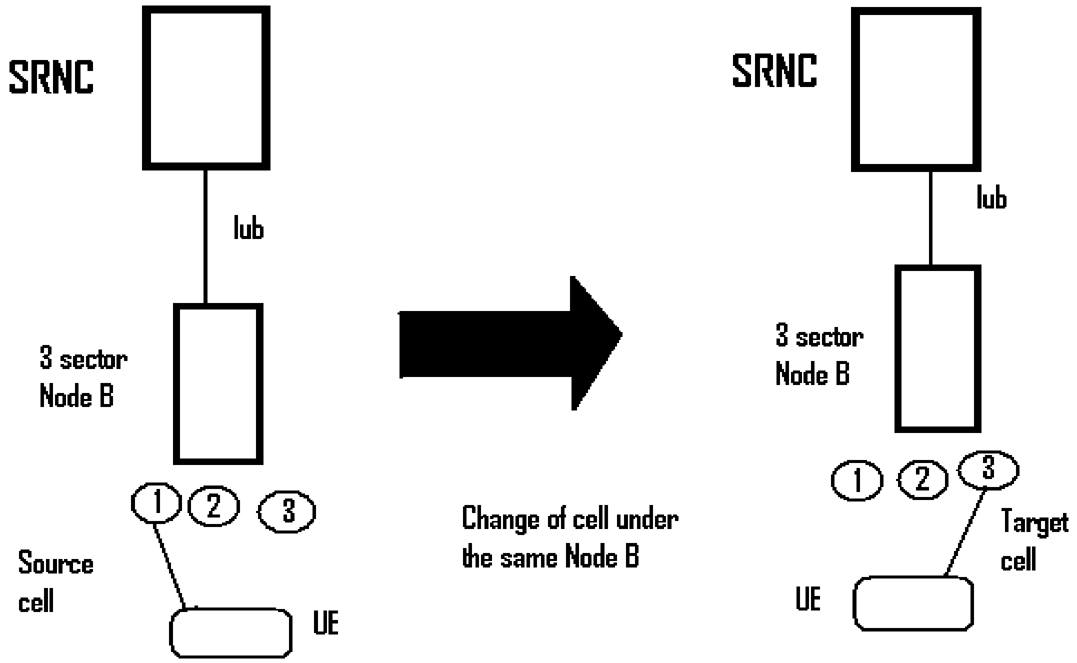

Figure 4.1. Example of intra-Node B HS-DSCH to HS-DSCH handover

At a specified time index where the handover from the source cell to the new target cell is carried out, the source cell stops transmitting to the user, and the MAC-hs packet scheduler in the target cell is thereafter allowed to control transmission to the user. Similarly, the terminal starts to listen to the HS-SCCH from the new target cell, i.e. the new serving HS-DSCH cell. This also implies that the CQI (Channel Quality Indicator) reports from the user are measured from the channel quality corresponding to the new target cell. It is typically recommended that the MAC-hs in the target cell does not start transmitting to the user until has received the first CQI report that is measured from the target cell.

Prior to the HS-DSCH handover from the source cell to the new target cell, there are likely to be several protocol data units (PDUs) buffered in the source cell´s MAC-hs for the user, both PDUs that have never been transmitted to the user and pending PDUs in the HARQ manager that are either awaiting ACK/NACK on the uplink HS-DPCCH or PDUs that are waiting to be transmitted to the user. Assuming that the Node B supports MAC-hs preservation, all the PDUs for the user are moved from the MAC-hs in the source cell to the MAC-hs in the target cell during the HS-DSCH handover. This means that the status of the HARQ manager is also preserved without triggering any higher layer retransmission, such as RLC retransmission during intra-Node B HS-DSCH to HS-DSCH handover. In the case that Node B does not support the MAC-hs preservation, the handling of the PDU not completed is the same as in inter-Node B handover case.

4.1.3.2 Inter-Node-Node B HS-DSCH to HS-DSCH Handover

Inter-Node B HS-DSCH to HS-DSCH handover is also supported by the 3GPP specifications, where the serving HS-DSCH source cell is under one Node B while the new target cell is under another Node B, and potentially also under another RNC, as illustrated in Figure 4.2.

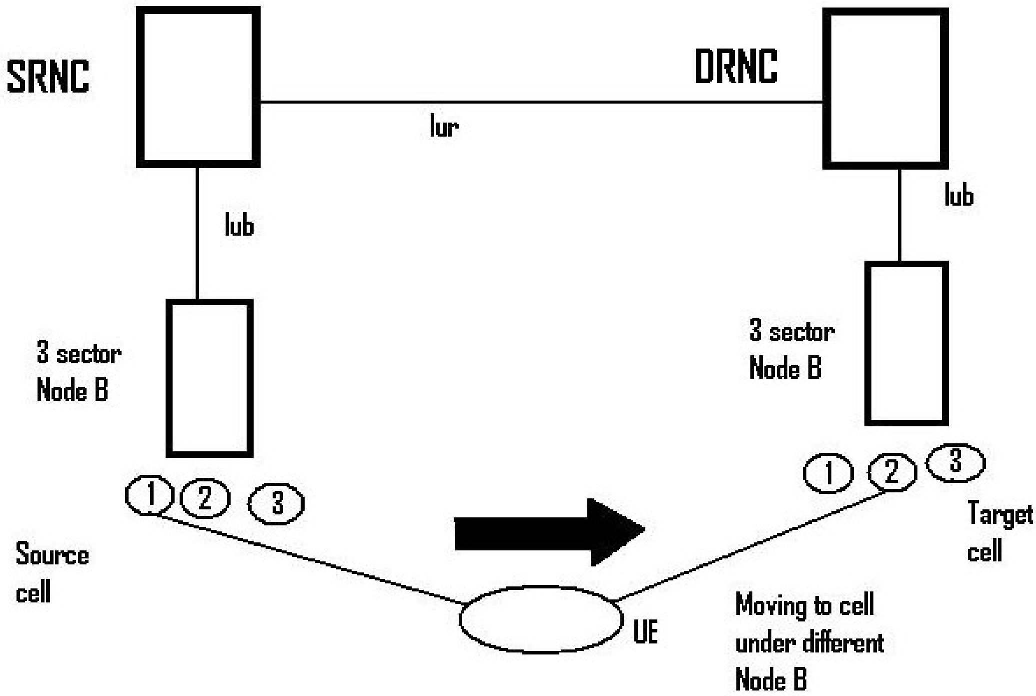

Figure 4.2. Example of inter-Node B HS-DSCH to HS-DSCH handover

Once the SRNC decides to initiate such a handover, a synchronised radio link reconfiguration prepare message is sent to the drifting RNC and the Node B that controls the target cell, as well as an RRC physical channel reconfiguration message to the user. At the time where the cell change is implemented, the MAC-hs for the user in the source cell is reset, which basically means that all buffered PDUs for the user are deleted, including the pending PDUs in the HARQ manager. At the same time index, the flow control unit in the MAC-hs in the target cell starts to request PDUs from the MAC-d in the SRNC, so that it can start to transmit data on the HS-DSCH to the user.

As the PDUs that were buffered in the source cell prior to the handover are deleted, these PDUs must be recovered by higher layer retransmissions, such as RLC retransmissions. When the RLC protocol realises that the PDUs it has originally forwarded to the source cell are not acknowledged, it will initiate retransmissions, which basically implies forwarding the same PDUs to the new target cell that was deleted in the source cell. In order to reduce the potential PDU transmission delays during this recovery phase, the RLC protocol at the user can be configured to send an RLC status report to the UTRAN at the first time instant after serving HS-DSCH cell has been changed [3]. This implies that the RLC protocol in the RNC can immediately start to forward the PDUs that were deleted in the source cell prior to the HS-DSCH cell change.

4.1.3.3 HS-DSCH to DCH Handover

Handover from an HS-DSCH to DCH may potentially be needed for HSDPA users that are moving from a cell with HSDPA to a cell without HSDPA (Release 99-compliant only cell),as illustrated in Figure 4.3.

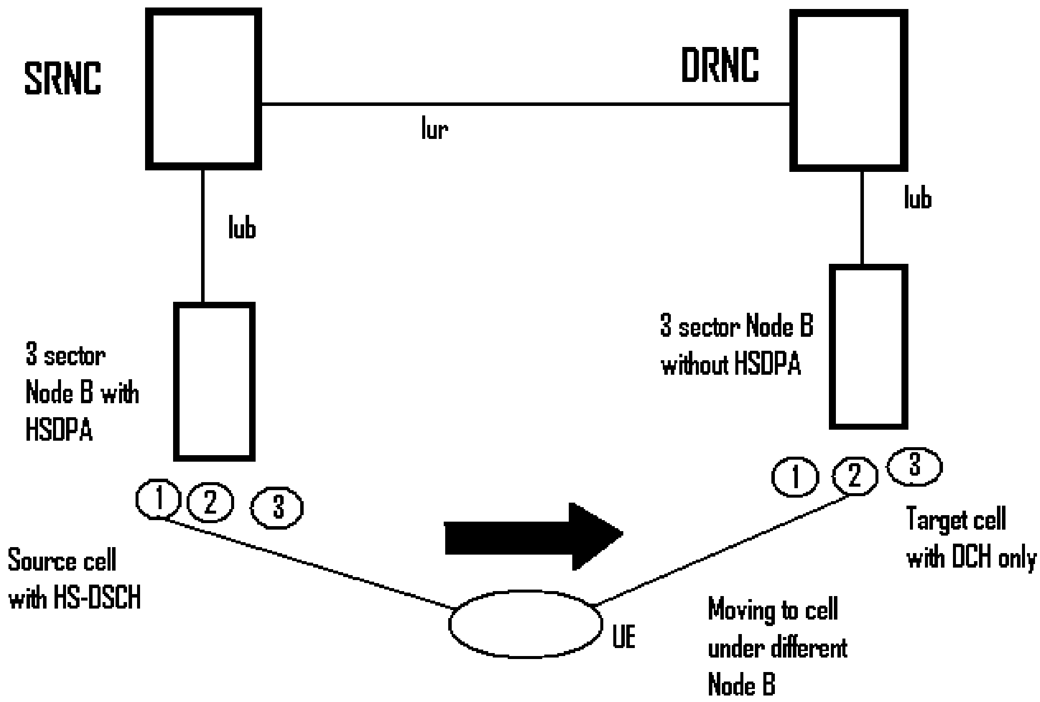

Figure 4.3. Example of HS-DSCH to DCH handover

Once the SRNC decides to initiate such a handover, a synchronised radio link reconfiguration prepare message is sent to the Node B involved, as well as an RRC physical channel reconfiguration message to the user. Similarly, for the inter-Node B HS-DSCH to HS-DSCH handover, the HS-DSCH to DCH handover results in a reset of the PDUs in MAC-hs in the source cell, which subsequently requires recovery via higher layer retransmission, such as RLC retransmissions.

The Release 5 specifications also support implementation of handover from DCH to HS-DSCH. This handover type may, for instance, be used if a user is moving from a non-HSDPA-capable cell into an HSDPA-capable cell, or to optimise the load balance in between HSDPA and DCH use in a cell.

4.1.4 HSDPA Performance

In this section, different performance aspects related to HSDPA are discussed. Since the two most basic features of WCDMA, fast power control and variable SF, have been disabled, a performance evaluation of HSDPA involves considerations that differ somewhat from the general WCDMA analysis. Next, some information of HSDPA system performance are given.

4.1.4.1 Factors Governing Performance

The performance of HSDPA depends on a number of factors, including the following:

- Channel conditions. Time dispersion, cell environment, terminal velocity as well as the ratio of experienced own-cell interference with the other cell interference (I_{or}/I_{oc}). Compared with the DCHs, the average I_{or}/I_{oc} ratio at the cell edge is reduced for HSDPA owing to the lack of soft handover gain. Macrocell network measurements indicate typical values down to -5 dB compared with approximately -2 to 0 dB for DCH.

- Terminal performance. Basic detector performance and HSDPA capability level, including supported peak data rates and number of multi-codes.

- Nature and accuracy of radio resource management. Power and code resources allocated to the HSDPA channel and accuracy/philosophy of Signal-to-Interference (SIR) power ratio estimation and packet-scheduling algorithms.

For a terminal with high detection performance, some experienced SIR would potentially map into a higher throughput performance experienced directly by the HSDPA user.

4.1.4.2 Spectral Efficiency, Code Efficiency and Dynamic Range

In WCDMA, both spectral efficiency and code efficiency are important optimisation criteria to accommodate code-limited and power-limited system states. In this respect, HSDPA provides some important improvements over Release 99 DCH and DSCH:

- Spectral efficiency is improved at lower SIR ranges (medium to long distance from Node B) by introducing more efficient coding and fast HARQ with redundancy combining. HARQ combines each packet retransmission with earlier transmissions, such that no transmissions are wasted. Further, extensive multi-code operations offer high spectral efficiency, similar to variable SF but with higher resolution. At very good SIR conditions (vicinity of Node B), HSDPA offers higher peak data rates and, thus, better channel utilisation and spectral efficiency.

- Code efficiency is obtained by offering more user bits per symbol and, thus, more data per channelisation code. This is obtained through higher-order modulation and reduced coding. Further, the use of time multiplexing and shared channels generally leads to better code utilisation for bursty traffic.

The principle of HSDPA is to adapt to the current channel conditions by selecting the most suitable modulation and coding scheme, leading to the highest throughput level.

In reality, the available data rate range may be slightly limited in both ends due to reasons of packet header overhead and practical detection limitations. The maximum peak data rate is, thus, often described to be on the order of 11-12 Mbps. The key measure for describing the link performance is the narrowband signal-to-interference-and-noise ratio (SINR) as experienced by the UE detector (e.g. the received E_S/N_O). In hostile environments, the availability of high SINR is limited, which reduces the link and cell throughput capabilities.

4.2 High-Speed Uplink Packet Access (HSUPA)

4.2.1 Introduction to HSUPA

This section presents High-Speed Uplink Packet Access (HSUPA) for Wideband Code Division Multiple Access (WCDMA), the key new feature included in the Release 6 specifications. The HSUPA solution has been designed to deliver similar benefits for the uplink as did the High-Speed Downlink Packet Access (HSDPA) in Release 5 for the downlink. The technologies applied with HSUPA are to improve uplink packet data performance by means of fast physical layer (L1) retransmission and transmission combining, as well as fast Node B (Base Transceiver Station (BTS)) controlled scheduling.

In Release 99, various methods exist for packet data transmission in WCDMA uplink. The three different channels in Release 99 and Release 4 WCDMA specifications that can be used for uplink packet data are:

- Dedicated Channel (DCH)
- Common Packet Channel (CPCH)
- Random Access Channel (RACH)

From Release 5 forward, however, only DCH and RACH are retained, as 3GPP concluded in June 2006 with the removal of a set of features (including CPCH) not implemented nor in the plans for introduction in the market place by operators and equipment manufacturers.

The DCH can be used basically for any type of service, and it has a dynamically variable spreading factor (SF) in the uplink, with an adjustment period of 10 to 40 ms. The momentary data rate can thus vary every interleaving period, which is between 10 and 40 ms. The issue of code space occupancy is not a real concern in the uplink direction, as each user has a user specific scrambling code and, thus, can use the full core tree if needed. Rather, the uplink DCH consumes both noise raise budget and network resources according to the peak data rate configured for the connection. The theoretical data rate with Release 99 runs up to 2 Mbps, but in practise the devices and networks have implemented typically 384 kbps as the maximum uplink capability. Higher numbers would easily mean reserving the whole cell/sector capacity for a single user regardless of the actual data rate being used.

On the other hand, the key of the HSUPA concept is to increase uplink packet data throughput with methods similar to HSDPA, base station scheduling and fast physical layer (L1) retransmission combining. While the telecoms industry is widely using the term HSUPA, it is not used in 3GPP specifications. In the specifications the term Enhanced DCH (E-DCH) is applied for the transport channel carrying the user data with HSUPA. A comparison of the basic properties and components of E-DCH and DCH is conducted in Table 4.4

Feature	DCH	E-DCH
Variable SF	Yes	Yes
Fast power control	Yes	Yes
Adaptive modulation	No	No
Multi-code operation	Yes (in specs, not used)	Yes, extended
Fast L1 HARQ	No	Yes
Fast BTS scheduler	No	Yes
Soft handover	Yes	Yes

Table 4.4. Comparison of fundamental properties of DCH and E-DCH

The general functionality of HSUPA is the following. The Node B estimates the data rate transmission needs of each active HSUPA user based on the device-specific feedback. The scheduler in Node B then provides instruction to devices on the uplink data rate to be used at a fast pace depending on the feedback received, the scheduling algorithm and the user priorisation scheme. Further, the retransmission are initiated by the Nodo B feedback.

Whereas HSDPA no longer uses power control, the same does not hold with HSUPA. HSUPA retains the uplink power control with a 70 dB or more dynamic range (exact range depends on the power class and terminal minimum power level). Thus, with HSUPA the signal never arrives at too high a symbol energy level, which is the case with HSDPA, and thus a justification for the use of higher-order modulation with HSDPA. Thus, the key thing for increased data rate is extensive multi-code operation together with base-station-based scheduling and retransmission handling. The Release 99 uplink feature of variable SF is retained; the range of Sfs is only slightly changed.

As the control of the scheduling is now in the base station, i.e. the receiving side of the radio link, there is adedd delay in the operation. This is in contrast to the HSDPA operation, where the base station scheduler resides in the transmitting side of the radio link. Thus, tracking the fast fading of the user for scheduling the uplink is not necessarily that beneficial. Rather, the key idea is to enable the scheduling to track the instantaneous transmission needs and capabilities of each device and then allocate such a data rate when really needed by the device. The fast scheduling allows dynamics sharing not only of the interference budget, but also of network resources, like baseband processing capacity and Iub transmission resources.

The physical layer retransmission combining is similar to HSDPA: now, it is just the base station that stores the received data packets in soft memory and, if decoding has failed, it combines the new transmission attempt with the old one. The key functionalities between HSDPA and HSUPA are compared in Table 4.5

Feature	HSUPA	HSDPA
Variable SF	Yes	No
Fast power control	Yes	No
Adaptive modulation	No	Yes
Scheduling	Multipoint to Point	Point to Multipoint
Fast L1 HARQ	Yes	Yes
Non-scheduled transmissions	Yes	No
Soft handover	Yes	No

Table 4.5. Comparison of fundamental properties of HSDPA and HSUPA

4.2.2 HSUPA Feasibility

After having completed HSDPA specifications, 3GPP started a feasibility study to investigate how the methods known from the HSDPA feature could be applied to the uplink direction and what would be the resulting benefits from doing so. The key difference from HSDPA was the finding that, now, there was no capacity benefit from the higher-order modulation due to the previously mentioned power control possibility with large dynamic range and the resulting higher E_b/N_0 from the use of the higher-order modulation. The use of higher-order modulation was considered more from the point of view of whether one could obtain a transmission with better envelope properties when comparing the use of multicode binary phase shift keying (BPSK) operation with a single code 8 PSK case. The Node B-based scheduling and physical layer retransmission were found beneficial. The results presented in the feasibility study report [4] showed a 50-70 % increase in the uplink packet data throughput with HSUPA from Release 99 DCH operation. Thus, in March 2004, 3GPP decided to close the study and to start the actual specification work, which was finalized then for the end of 2004.

4.2.3 HSUPA Physical Layer Structure

The transport channel, E-DCH, is sent in the uplink together with the Release 99 DCH, and at least the control part of DCH (Dedicated Physical Control Channel (DPCCH)) is always present to carry the pilot bits and downlink power control commands in the uplink direction. The presence of the Dedicated Physical Data Channel (DPDCH) for user data depends whether there is e.g. AMR speech call operated in parallel to uplink packet data transmission. The following new physical channels are introduced in the 3GPP Release 6 specifications to enable HSUPA operation [4]:

- Enhanced DPDCH (E-DPDCH) carries the user data in the uplink direction and reaches the physical layer peak rate of 5.76 Mbps when up to four parallel code channels are in use.
- Enhanced DPCCH (E-DPCCH) carries in the uplink the E-DPDCH-related rate information, retransmission information and the information to be used by the base station for scheduling control.
- E-DCH Hybrid ARQ (HARQ) Indicator Channel (E-HICH) carries information in the downlink direction on whether a particular base station has received the uplink packet correctly or not.
- E-DCH Absolute Grant Channel (E-AGCH) and E-DCH Relative Grant Channel (E-RGCH) carry the Node B scheduling control information to control the uplink transmission rate.

Next figure (Figure 4.4) depicts briefly the Physical Layer Structure .

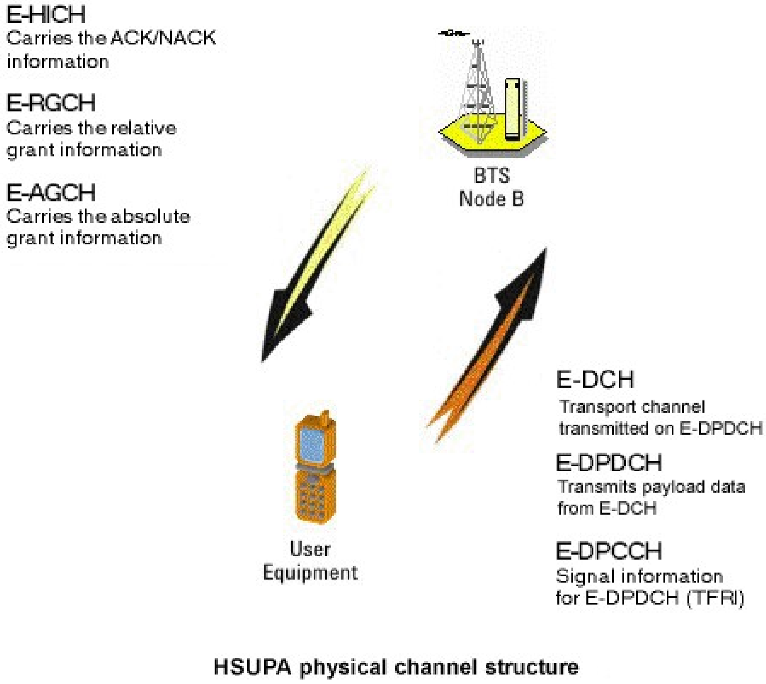

Figure 4.4. HSUPA Physical Layer Structure

4.2.4 HSUPA Physical Layer Operation Procedure

The HSUPA Node B scheduler operation can be described as follows:

- The scheduler in the Node B measures, for example, the noise level at the base station receiver to decide whether additional traffic can be allocated or whether some users should have smaller data rates.

- The scheduler also monitors the uplink feedback, the happy bits, on E-DPCCHs from different users sent in every TTI. This tells which users could transmit at a higher data rate both from the buffer status and the transmission power availability point of view. There is also more detailed information in the MAC-e header on the buffer occupancy and uplink power headroom availability. The latter informs how much reserve transmission power the terminal still has, and the former gives the Node B scheduler the information on whether the UE would actually benefit from having a higher data rate or whether it could be downgraded to a lower one.

- Depending on possible user priorities given from the RNC, the scheduler chooses a particular user or users for data rate adjustment. The respective relative or absolute rate commands are then send on the E-RGCH or E-AGCH.

Thus, in the uplink direction, the possible user data rate restriction needs to be informed to the base station from RNC so that the maximum data rate for the service subscribed is not exceeded. The RNC may also give different priorities for the different services for the same users, based on the MAC flow identifiers.

In addition to the scheduled traffic there may be also non-scheduled transmissions, such as Signaling Radio Bearer (SRB)or, for example, a Voice-over-IP connection. Both types of data have a limited delay or delay variance budget and low data rate. Thus, scheduling them would not add much value and would, in the worst case, just degrade the system operation due, for example, to delayed measurement reports; so, such services are given a permanent grant by the RNC that the Node B scheduler cannot influence. Thus, the non-scheduled transmission operates similar to the Release 99 DCH, but only taking advantage of the physical layer retransmission procedure.

The HSUPA physical layer retransmissions operate as follows

- Depending on the TTI, four or eight HARQ processes are in use.

- The terminal will send a packet in line with the allowed data rate

- After the packet has been transmitted, E-HICH is monitored from all the cells in the active set with E-DCH activated. A maximum of four cells can have E-DCH allocated from the maximum active set of six cells.

- If any of the cells indicates a positive acknowledgment (ACK), then the terminal will proceed for a new packet, otherwise retransmission occurs.

The HSUPA operation procedure has strictly specified timing values for the terminal operation as well as for the ACK/negative acknowledgment (NACK) timing response from the base station, i.e. the whole procedure is synchronised starting from the initial transmission by the UE and ending with the positive ACK reception from the Node B covering the potential retransmission in between. This removes the need for signalling separately any HARQ process numbers, as the transmission timing always tells for which HARQ channel the retransmission or HARQ feedback is for.

The timing and number of HARQ channels depends now only on the TTI in use; there are no HARQ channels to configure, unlike with HSDPA. With a 10 ms TTI the number of HARQ channels is four and with a 2 ms TTI the number of channels is eight.

4.2.5 HSUPA Terminal Capability

The HSUPA feature is optional for terminals in Release 6 with six different categories of terminals allowed by the standard with resulting maximum physical layer data rates ranging between 0.72 Mbps and 5.76 Mbps. If a terminal supports HSUPA, then it is also mandatory to support HSDPA. Thus, it is possible in Release 6 to have three kinds of device, with the main classification as:

- DCH-only device

- DCH and HSDPA-capable device

- DCH, HSDPA- and HSUPA-capable device

For the HSDPA and HSUPA, a terminal can obviously choose the category defining the maximum data rate supported, and especially with HSUPA whether the optional 2 ms TTI is supported or not.

4.2.6 HSUPA Performance

The two fundamental HSUPA features, physical layer packet combining and Node B scheduling will be briefly explained in the following paragraphs.

4.2.6.1 Physical Layer Retransmission Combining

The physical layer retransmission allow a smaller E_b/N_0 for the initial transmission. The BLER for the first transmission is also a trade-off between delay, capacity and resulting baseband resource. With too low a BLER there are too few retransmissions and, for example, only 1 % of the packets are transmitted more than once, leaving the benefits of fast retransmission rather marginal. Thus, in the order of 10 % of the packets should fail for the first transmission. This provides a bit more than 1 dB reduction for the transmission power needed for a given data rate when comparing with a 1 % initial BLER target. Too high a BLER will not eventually boost capacity, it will just occupy more resources in the network. If every packet gets transmitted two times on average, then supporting a 1 Mbps data stream requires resources equal to 2 Mbps and then also the maximum single user data rate reduces as so many packets need to be transmitted.

4.2.6.2 Node B-Based Scheduling

The Node B based scheduling allows faster reaction to the transmission needs from the terminal. This means that the air interface capacity is better utilised, resulting in higher capacity. Also, one can then allocate more high bit-rate users simultaneously, as now they are not all going to be allowed to use the maximum bit rate simultaneously, which means better availability of a high bit-rate uplink service.

4.3 High-Speed Packet Access Evolution (HSPA+)

4.3.1 Introduction to HSPA+

High-Speed Downlink Packet Access (HSDPA) in 3rd Generation Partnership Program Release 5 included major improvements in downlink data rates and capacity. Similar technical solutions were applied to the uplink direction as part of Release 6 with High-Speed Uplink Packet Access (HSUPA). 3GPP Release 7 brings a number of further substantial enhancements to the end-user performance, network capacity and to the network architecture. The 3GPP Release 7 solutions also bring High-Speed Packet Access (HSPA) capabilities closer to 3GPP long-term evolution (LTE) targets. LTE will be specified as part of Release 8 and further push the radio capabilities higher. The 3GPP Release 7 and 8 solutions for HSPA evolution will be worked in parallel with LTE development, and some aspects of the LTE work are also expected to reflect on HSPA evolution. The HSPA evolution is also known as HSPA+.

In the next sections, the most important features of HSPA+ will be explained.

4.3.2 Downlink MIMO for HSPA+

The term MIMO (Multiple Input Multiple Output) is widely used to refer to multi antenna technology. In general, the term MIMO refers to a system having multiple input signals and multiple output signals. In practice, MIMO means the use of multiple antennas at transmitter and receiver side in order to exploit the spatial dimension of the radio channel. MIMO systems significantly enhance the performance of data transmission. Note that different types of performance gains can be discriminated. On the one hand side, diversity gains can be exploited to increase the quality of data transmission. On the other hand side, spatial multiplexing gains can be exploited to increase the throughput of data transmission.

Downlink MIMO has been introduced in the context of HSPA+ to increase throughput and data rate. Baseline is a 2x2 MIMO system, i.e. 2 transmit antennas at the base station side, and 2 receive antennas at the UE side. MIMO for HSPA+ allows (theoretical) downlink peak data rates of 28 Mbps. Note that HSPA+ does not support uplink MIMO.

The process of introducing MIMO in HSPA+ took a long time in 3GPP. A large number of different approaches was evaluated and extensive performance studies were carried out. Finally, a consensus was reached to extend the closed loop transmit diversity scheme of 3GPP release 99 WCDMA (Wideband Code Division Multiple Access) to a full MIMO approach including spatial multiplexing. The approach is called D-TxAA which means Double Transmit Antenna Array. It is only applicable for the High Speed Downlink Shared Channel, the HS-DSCH. Figure 4.5 shows the basic principle of the 2x2 approach.

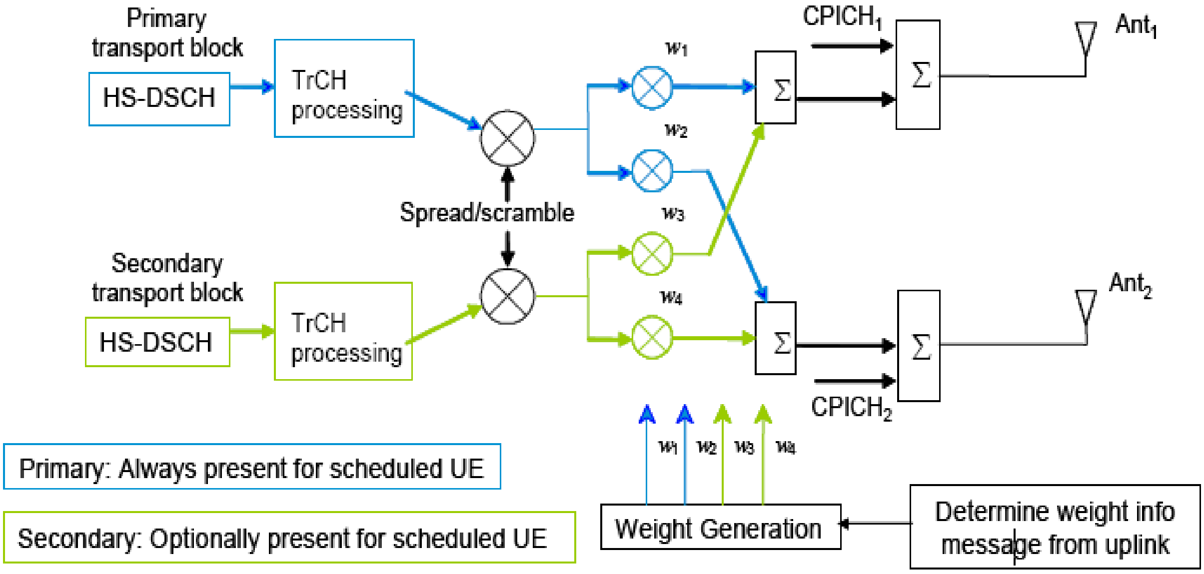

Figure 4.5. MIMO for HSPA+

With D-TxAA, two independent data streams (transport blocks to be more precise) can be transmitted simultaneously over the radio channel over the same WCDMA channelization codes. The two data streams are indicated with blue and green colour in Figure 4.5. Each transport block is processed and channel coded separately. After spreading and scrambling, precoding based on weight factors is applied to optimize the signal for transmission over the mobile radio channel. Four precoding weights w1- w4 are available. The first stream is multiplied with w1 and w2, the second stream is multiplied with w3 and w4.

After multiplication with the weight factors, the two data streams are summed up before transmission on each antenna, so that each antenna transmits a part of each stream. Note that the two different transport blocks can have a different modulation and coding scheme depending on data rate requirements and radio channel condition. The UE has to be able to do channel estimation for the radio channels seen from each transmit antenna, respectively. Thus, the transmit antennas have to transmit a different pilot signal. One of the antennas will transmit the antenna 1 modulation pattern of PCPICH (Primary Common Pilot Channel). The other antenna will transmit either the antenna 2 modulation pattern of P-CPICH, or the antenna 1 modulation pattern of SCPICH.

Also the UE receiver has to know the precoding weights that were applied at the transmitter. Therefore, the base station signals to the UE the precoding weight w2 via the HS-SCCH (High Speed Shared Control Channel). The 2 bit precoding weight indication is used on HS-SCCH to signal one out of four possible w2 values. The other weights applied on HS-DSCH can then be derived from w2. The precoding weight adjustment is done at the sub-frame border.

The 16-Quadrature Amplitude Modulation (QAM) and 64QAM constellations, among other, are shown in Figure 4.6 16QAM can double the bit rate compared with Quadrature Phase Shift

Keying (QPSK) by transmitting 4 bits instead of 2 bits per symbol. 64QAM can increase the peak bit rate by 50 % compared with 16 QAM, since 64QAM transmits 6 bits with a single symbol. On the other hand, the constellation points are closer to each other for the higher-order modulation and the required signal-to-noise ratio for correct reception is higher. The difference in the required signal-to-noise is approximately 6 dB between 16 QAM and QPSK, and between 64QAM and 16QAM. Therefore, downlink 64QAM and uplink 16QAM can be utilised only when the channel conditions are favorable.

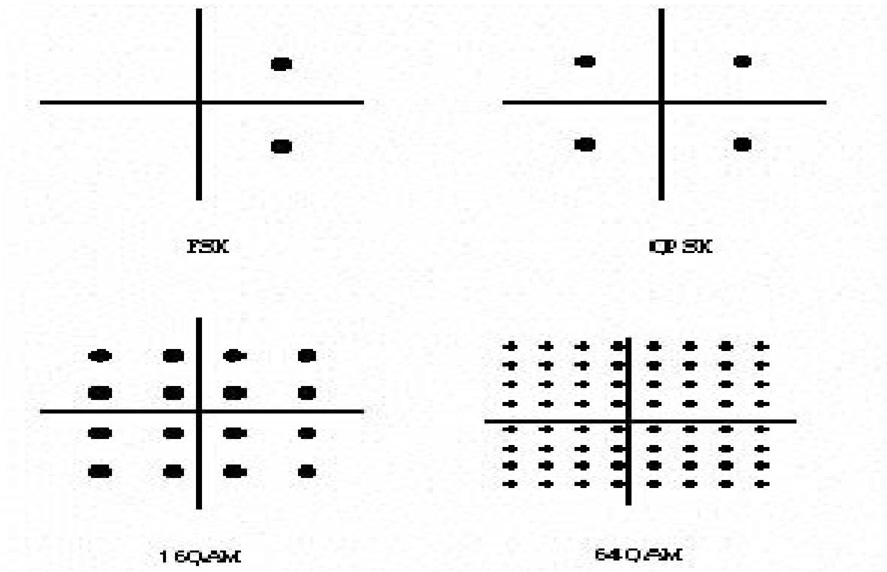

Figure 4.6. Different types of constellations

HSDPA terminal categories are illustrated in Table 4.6 and HSUPA categories in Table 4.7.

The HSDPA categories 13-16 are added in Release 7. Categories 13 and 14 include 64QAM and categories 15 and 16 MIMO. The peak bit rate with 64QAM is 21.1 Mbps and with 16QAM ans MIMO is 28.0 Mbps. The combination of 2x2 MIMO and 64QAM modulation would push the theoretical peak data rate beyond 40 Mbps, but that combination is not included in Release 7. The HSUPA category 7 is added in Release 7 with 16QAM capability pushing the peak rate to 11.5 Mbps.

Category	*Codes*	*Modulation*	*MIMO*	*Coding rate*	*Peak bit rate (Mbps)*	*3GPP release*
12	5	QPSK	-	3/4	1.8	Release 5
5/6	5	16QAM	-	3/4	3.6	Release 5
7/8	10	16QAM	-	3/4	7.2	Release 5
9	15	16QAM	-	3/4	10.1	Release 5
10	15	16QAM	-	Approx 1/1	14.0	Release 5
13	15	64QAM	-	5/6	17.4	Release 7
14	15	64QAM	-	Approx 1/1	21.1	Release 7
15	15	16QAM	2x2	5/6	23.4	Release 7
16	15	16QAM	2x2	Approx 1/1	28.0	Release 7

Table 4.7. Selected HSDPA terminal categories

Category	*TTI (ms)*	*Modulation*	*Coding rate*	*Peak bit rate (Mbps)*	*3GPP release*
3	10	QPSK	3/4	1.4	Release 6
5	10	QPSK	3/4	2.0	Release 6
6	2	QPSK	1/1	5.7	Release 6
7	2	16QAM	1/1	11.5	Release 7

Table 4.8. Selected HSUPA terminal categories

4.3.3 Mobile Power Consumption Reduction with Continuous Packet Connectivity

Technology evolution, in general, helps to decrease mobile power consumption. Also, fast and accurate power control in WCDMA helps to minimise the transmitted power levels. The challenge is still for WCDMA mobile needs to have continuous reception and transmission, which requires the radio-frequency (RF) parts to be running all the time during the voice or data call. 3GPP Release 7 introduces a few improvements to HSDPA/HSUPA that help to reduce the power consumption for bursty data or for low data-rate packet services, like VoIP. These improvements are part of Continuous Packet Connectivity feature in Release 7.

In Release 6 the uplink physical control channel is transmitted even if there is no data channel active. In Release 7 the mobile terminal cuts off the control channel transmission where there is no user-plane data to be sent. This solution is called discontinuous uplink transmission or, more exactly, discontinuous uplink DCPCCH transmission. The discontinuous transmission allows the shutting down of the terminal RF parts when there is no transmission. A similar concept is also introduced in the downlink, enabling the UE to turn off its receiver at times. The terminal will be informed in downlink about the possible times where there may be scheduled data. The terminal can use power-saving mode in the RF receiver chain during other parts of the frame when it has no data to receive. This solution is called downlink discontinuous reception (DRX). Owing to these Release 7 concepts, the operating times with Release 7 can be longer than for Releases 99, 5, and 6 for low data-rates services, like VoIP or services generating bursty traffic. The lower power consumption is also beneficial in the case when the data, like a web page, have been downloaded and the connection is inactive but still in the Cell_DCH state.

4.3.4 Voice-over-IP (VOIP) Capacity Enhancements

Circuit-switched (CS) voice used to be the only way to provide a voice service in cellular networks. The introduction of 3G networks, including WCDMA Release 99, made it possible to run VoIP over cellular networks with reasonable quality, but with lower than CS voice. 3GPP Release 5 and 6 HSPA was originally designed to carry high bit-rate delay-tolerant data. A number of features have been introduced to 3GPP Release 6 and 7 to improve the efficiency of low bit-rate delay-critical applications, like VoIP.

The discontinuous uplink transmission not only reduces the power consumption, but there is also less interference transmitted by the terminal; consequently, a higher system capacity can be achieved. The VoIP capacity gain of discontinuous uplink transmission is approximately 50 % more capacity with Release 7 utilising the discontinuous uplink DPCCH than with the continuous uplink DPCCH of Release 6.

Another reason for higher uplink capacity in Release 7 is packet bundling for uplink. The physical layer overhead can be minimised and turbo coding gain increased by transmitting two or three VoIP packets together as long as the end-to-end delay allows that. For downlink this behaviour is a property of the Node B scheduler and, thus, can be supported already in Release 5 HSDPA. With HSUPA, the Release 6 UE must transmit each VoIP packet immediately and, thus, separately without packet bundling.

The VoIP capacity simulations are summarised in Figure 4.7 in terms of maximum number of simultaneous users per sector per 5 MHz carrier.

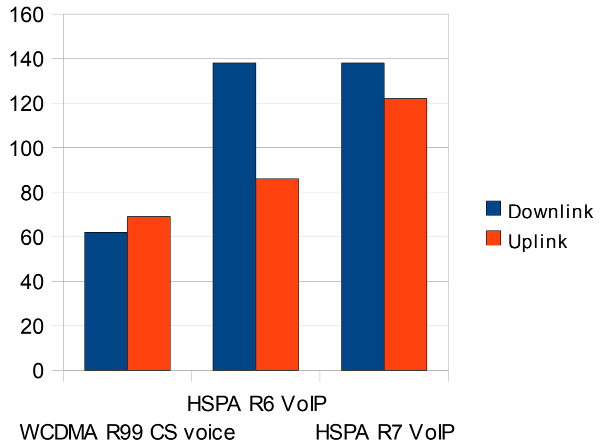

Figure 4.7. Voice capacity evolution with Release 7 VoIP

The CS capacity with Release 99 is estimated to be 60-70 users, whereas VoIP capacity with HSPA Release is up to two times higher, at 120 users. The VoIP capacity assumes that Robust Header Compression is used to compress IP headers [5].

4.3.5 Flat Architecture

3GPP networks will increasingly be used for IP-based packet services. 3GPP Release 6 has four networks elements in the user and control plane: base station, RNC, Serving General Packet Radio System (GPRS) Support Node (SGSN) and Gateway GPRS Support Node. The architecture in Release 8 LTE will have only two network elements: base station in the radio network and Access Gateway in the core network. A flat network architecture is also considered beneficial for HSPA , since reduces the network latency and, thus, improves the overall performance of IP-based services.

The introduction of HSDPA and HSUPA enables the flat architecture, since more functionalities are anyway distributed from RNC to Node B in Release 5 and 6. It is possible to design a flat HSPA radio network architecture where all the RNC functionalities are located in Node B. Such a network architecture can be implemented already based on 3GPP Release 5 or Release 6 by using the existing open interfaces.

The packet architecture evolution in HSPA is designed to be backwards compatible: existing Release 5 and Release 6 terminals can operate with the new architecture.

The hierarchical and flat architectures are illustrated in Figure 4.8. Release 7 includes flat architecture improvements both in the packet core and in the radio network. The so-called direct tunnel solution allows the user plane to bypass SGSN. The SGSN dimensioning will not be affected by the user-plane volume. The flat radio network architecture integrates RNC functionalities as part of the base station. The RNC dimensioning will not be affected by the traffic volume, since no RNC element is required. When the flat radio architecture and direct tunnel SGSN solutions are combined, there are only two network elements in the user plane in the same way as in the LTE architecture.

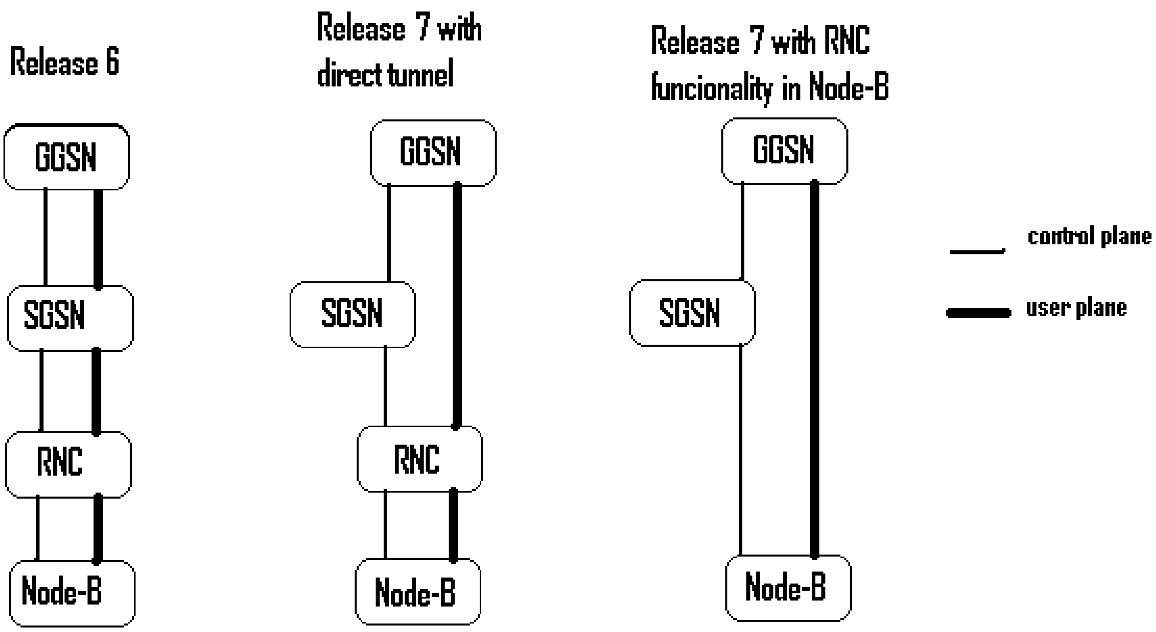

Figure 4.8. Evolution towards flat architecture

4.4 UTRAN Long-Term Evolution (LTE)

4.4.1 Background and introduction to LTE Technology

Long Term Evolution (LTE) describes standardisation work by the Third Generation Partnership Project (3GPP) to define a new high-speed radio access method for mobile communications systems.

LTE is the next step on a clearly-charted roadmap to so-called '4G' mobile systems that starts with today's 2G and 3G networks. Building on the technical foundations of the 3GPP family of cellular systems that embraces GSM, GPRS and EDGE as well as WCDMA and now HSPA (High Speed Packet Access), LTE offers a smooth evolutionary path to higher speeds and lower latency.

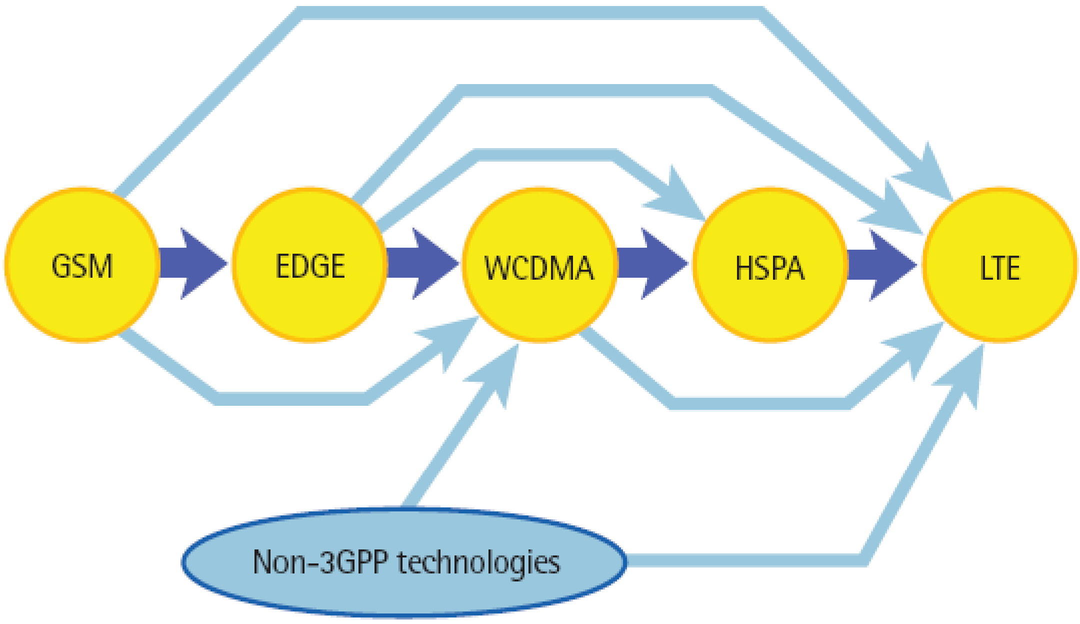

Figure 4.9. Evolution of the mobile technologies

3GPP proposed migrating towards an all-IP core network as early as Release 4, hinting at what would become a prominent feature of later UMTS/HSPA releases and ultimately LTE.

The concept of 'Long Term Evolution' for today's 3G/UMTS standard was discussed in detail in 2004, when a RAN (Radio Access Network) Evolution Workshop in Toronto accepted contributions from more than 40 operators, manufacturers and research institutes (including 3GPP members as well as nonmember organisations). Contributors offered a range of views and proposals on the evolution of the UTRAN (Universal Terrestrial Radio Access Network).

Following the Toronto workshop, in December 2004, 3GPP launched a feasibility study [7], in

March 2005, in order “to develop a framework for the evolution of the 3GPP radioaccess technology towards a high-data-rate, low-latency and packet-optimised radio-access technology”. In other words, the study would map out specifications for a radio access network (RAN) capable of supporting the broadband Internet user experience we already enjoy in today’s fixed networks – with the addition of full mobility to enable exciting new service possibilities.

The key requirements defined for the work are:

- Packet-switched domain optimised;

- Server to User Equipment (UE) round-trip time below 30 ms and access delay below 300 ms;

- Peak rates uplink/downlink 50/100 Mbps;

- Good level of mobility and security ensured;

- Improved terminal power efficiency;

- Frequency allocation flexibility with 1.25/2.5, 5, 10, 15 and 20 MHz allocations; possibility to deploy adjacent to WCDMA;

- Higher capacity compared with the Release 6 HSDPA/HSUPA reference case, in the d ownlink throughput three to four times and in the uplink two to three times reference scenario capacity

Today, specifications for LTE are encapsulated in 3GPP Release 8, the newest set of standards that defines the technical evolution of 3GPP mobile network systems. Release 8 succeeds the previous iteration of 3G standards – Release 7 – that includes specifications for HSPA+, the ‘missing link’ between HSPA and LTE. Defined in 3GPP Releases 7 and 8, HSPA+ allows the introduction of a simpler, ‘flat’, IP-oriented network architecture while bypassing many of the legacy equipment requirements of UMTS/HSPA.

Peak data rates with HSPA+ are 28 Mbit/s on the downlink and 11.5 Mbit/s on the uplink using 2x2 MIMO (Multiple-Input Multiple-Output) antenna techniques and 16QAM (Quadrature Amplitude Modulation). However, HSPA+ can further boost data rates up to 42 Mbit/s on the downlink and 23 Mbit/s on the uplink using 2x2MIMO and 64QAM, a combination that is part of Release 8.

As such, HSPA+ slots neatly between the already impressive performance of HSPA (with its theoretical downlink performance of up to 14.4 Mbit/s) and LTE that promises rates of 300 Mbit/s in the downlink and 75 Mbit/s in the uplink for every 20 MHz of paired spectrum.

From March 2005 until September 2006 3GPP conducted a feasibility study on the Evolved UTRAN (EUTRAN) technology alternatives and made a selection of the multiple access and basic radio access network architecture. 3GPP considered different multiple access options but came rather quickly to the final conclusion: Orthogonal Frequency Division Multiple Access (OFDMA), which was already one proposal considered in the UMTS selection phase, during 1997, but it was not mature enough at that point in time.

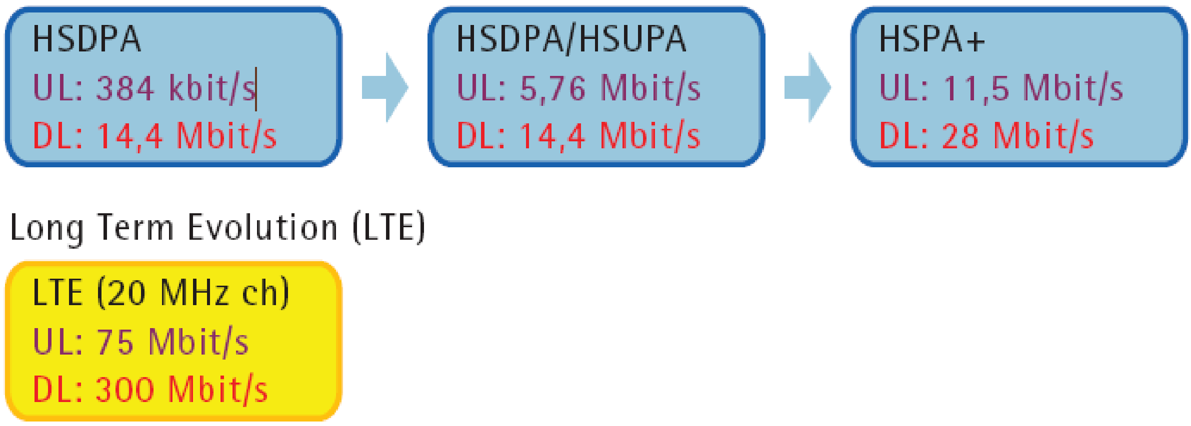

Figure 4.10. Uplink and downlink data rates compared for HSPA and LTE

So, LTE is built on an all-new radio access network based on OFDM (Orthogonal Frequency-Division Multiplexing) technology. Specified in 3GPP Release 8, the air interface for LTE combines OFDMA-based modulation and multiple access scheme for the downlink, together with SC-FDMA (Single Carrier FDMA) for the uplink.

The key developments are:

- Larger bandwidth and bandwidth flexibility. The UMTS spectrum allocation typically did not allow a larger carrier bandwidth than 5 MHz. WCDMA/High-Speed Packet Access (HSPA) shows attractive performance at 5 MHz when equaliser receivers are used. Also, the receiver complexity remains reasonably low. LTE work targets are for higher bit rates using larger bandwidth up to 20 MHz. OFDMA can provide benefits over a Code Division Multiple Access (CDMA)-based system when the bandwidth increases: the OFDMA signal remains orthogonal while CDMA performance suffers due to increased multipath components and the equaliser receiver gets more complex.

- Flat architecture. In the architecture development, more intelligence is being added to the base station, similar to the trend set by HSUPA and HSDPA. The UMTS architecture initially was defined to be hierarchical, where the radio-related functionalities were located in the Radio Network Controller (RNC). In the flat architecture the radio-related functionalities are located in the base station. When the packet scheduling is located in the base station, fast packet scheduling can be applied, including frequency domain scheduling, as shown in Figure 4.11. The frequency domain scheduling is shown to improve the cell capacity up to 50 % in the simulations. The frequency domain scheduling can be done in

OFDMA but not in CDMA, where the interference is always spread over the whole carrier bandwidth.

- Amplifier friendly uplink solution with SC-FDMA. One of the main challenges in OFDMA is the high peak-to-average radio of the transmitted signal, which requires linearity in the transmitter. The linear amplifiers have low efficiency; therefore, OFDMA is not an optimised solution for a mobile uplink where the target is to minimise the terminal power consumption. An LTE uplink uses SC-FDMA, which clearly enables better power-amplifier efficiency. SC-FDMA technology was not available when UMTS multiple access selection was done, but the firs articles were just being published at the time, such as [8].

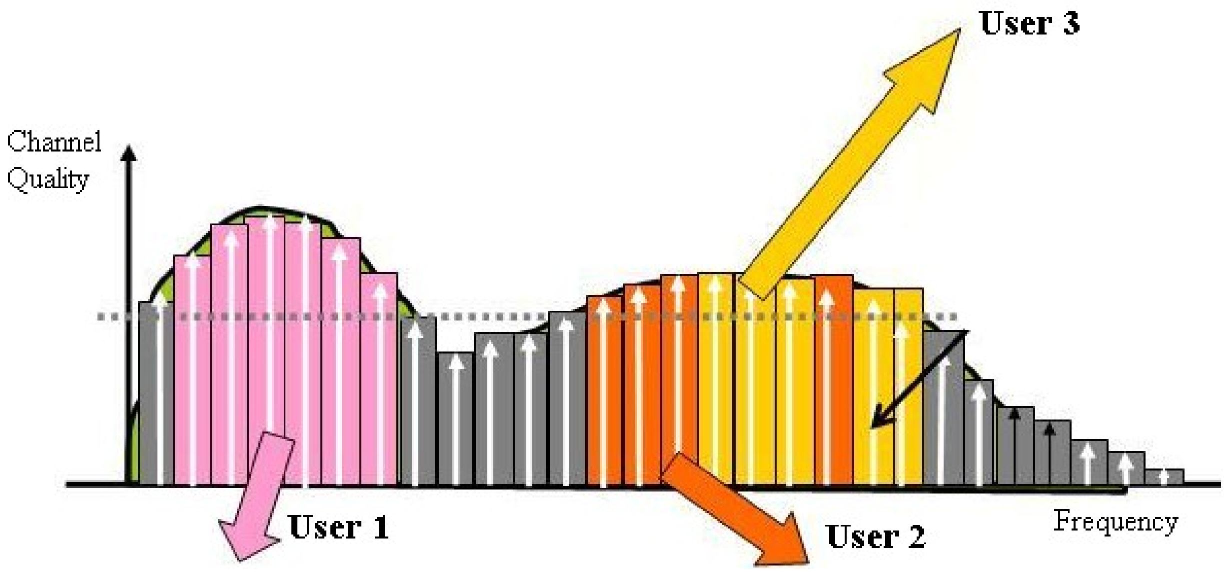

Figure 4.11. Frequency domain scheduling

- Simpler multi-antenna operation. Higher bit rates can generally be obtained by using larger bandwidth and multiple antennas. Multiple input multiple output (MIMO) antenna technologies, emerging over the past few years, are required to achieve the LTE bit-rate targets. MIMO is simpler to implement with OFDMA than with CDMA.

- 3GPP decided to place the radio functionality fully in the base station, as shown the comparison of the architectures in the Figure 4.12. While the only remaining element of the radio access network is eNode B, more elements can be used on the core network side.

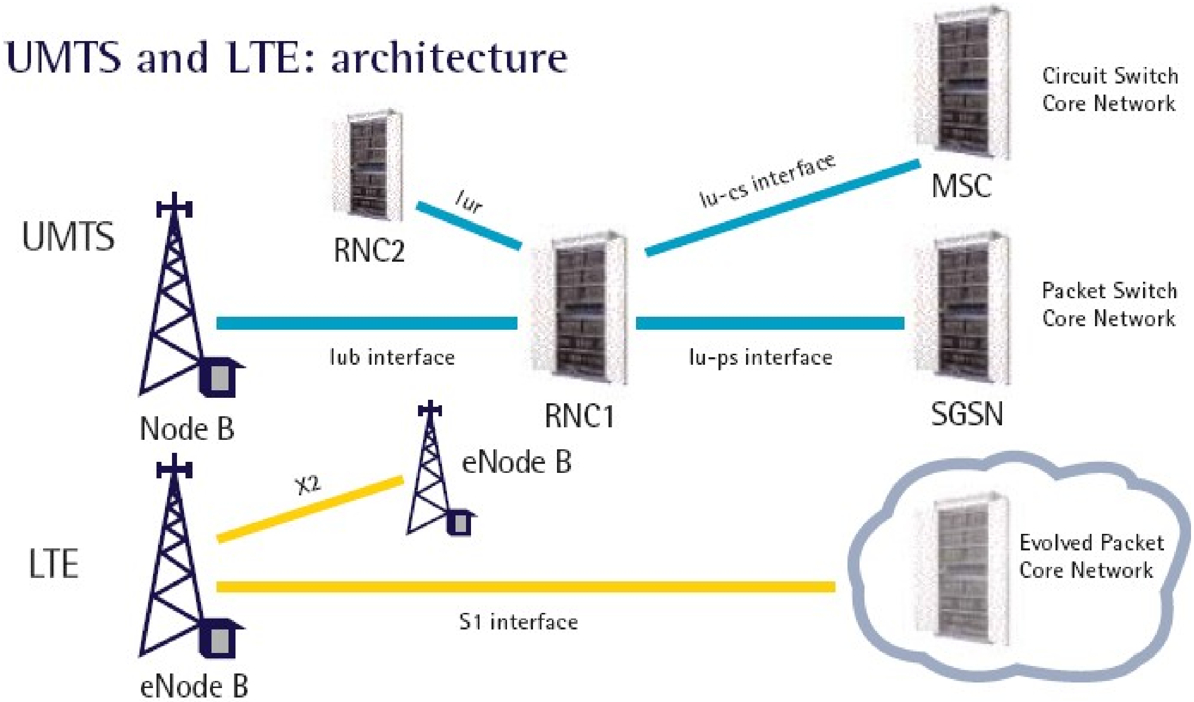

Figure 4.12. Comparison of the UMTS and LTE architectures.

4.4.2 LTE Multiple Access

As mentioned previously, the 3GPP multiple access is based on the use of OFDMA in the downlink direction and SC-FDMA in the uplink direction. While the basic principles of OFDMA are rather well know from, for example, wireless Local Area Network standards or digital TV standards, such as DVB-T/H, the use of SC-FDMA represents a more recent technology not widely used in any of the existing systems. The following sections will look at the SC-FDMA and OFDMA fundamentals.

4.4.2.1 OFDMA Principles

The principle of the OFDMA is based on the use of narrow, mutually orthogonal sub-carriers. In LTE the sub-carrier spacing is typically 15 kHz regardless of the total transmission bandwidth. Different sub-carriers maintain orthogonality, as at the sampling instant of a single sub-carrier the other sub-carriers have a zero value, as shown in the figure 4.13.

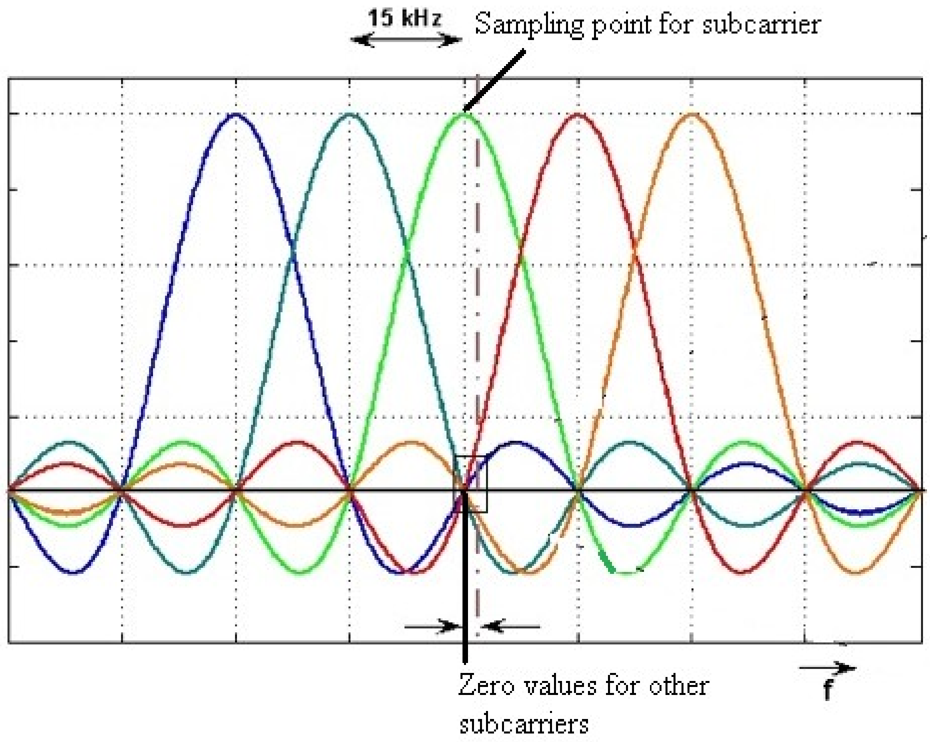

Figure 4.13. Adjacent subcarrier with OFDMA

The actual transmission is then done by transmitting a signal after the Fast Fourier Transform (FFT) block, which is used to change between the time and frequency domain representations of the signal. The transmitted signal now has the following key properties:

- The symbol duration is clearly longer than, for example, with WCDMA and also longer than the channel impulse response; thus, the channel impact is equal to a multiplication by a (complex-valued) scalar.
- There is no inter-symbol interference, as the transmitter uses a guard period (cyclic prefix) longer than the channel impulse response, which is ignored in the receiver and, thus, the effect of the previous symbol is not visible.
- The outcome of an FFT is thus a single signal which is basically a sum of sinusoids and having an amplitude variation that is larger the more sub-carriers have been used as an input to an FFT block.

This kind of signal is ideal from the receiver perspective as one does not need equaliser but only need to compensate the channel amplitude and phase impact on the different sub-carriers. In the receiver side one uses again the FFT to convert back form the frequency domain singles signal to the time domain representation of multiple sub-carriers, as shown in the Figure 4.14.

The channel estimation is done based on the known data symbols that need to be placed periodically on parts of the sub-carriers. The equaliser in Figure 4.14 refers to the estimator to cancel out the complex-valued multiplication caused by the frequency-selective fading of the channel and does not present a great complexity. Also, the FFT or Inverse FFT (IFFT) operations are rather old numerical principles for which computationally efficient algorithms have long been developed. The fundamental functional of an FFT block is to transfer the time-domain signal into the frequency domain, representing basically the frequency components from which the time-domain signal is constructed. In the case of OFDMA the parallel inputs to an FFT could be considered as the frequency-domain components which are then converted to a single time-domain signal, thus carrying in a single long symbol typically up to 512, 1024 or 2048 modulated symbols. Use of FFT length of power of 2 is economical from a computational complexity point of view and, thus, is a generally applied principle. The IFFT/FFT operation is illustrated in Figure 4.15, where parallel sub-carriers with their own modulation and modulation order are going to the IFFT block, which is transforming the input to a single output signal and respectively again the receiver FFT block converts the signal back to parallel Quadrature Amplitude Modulation (QAM) symbols.

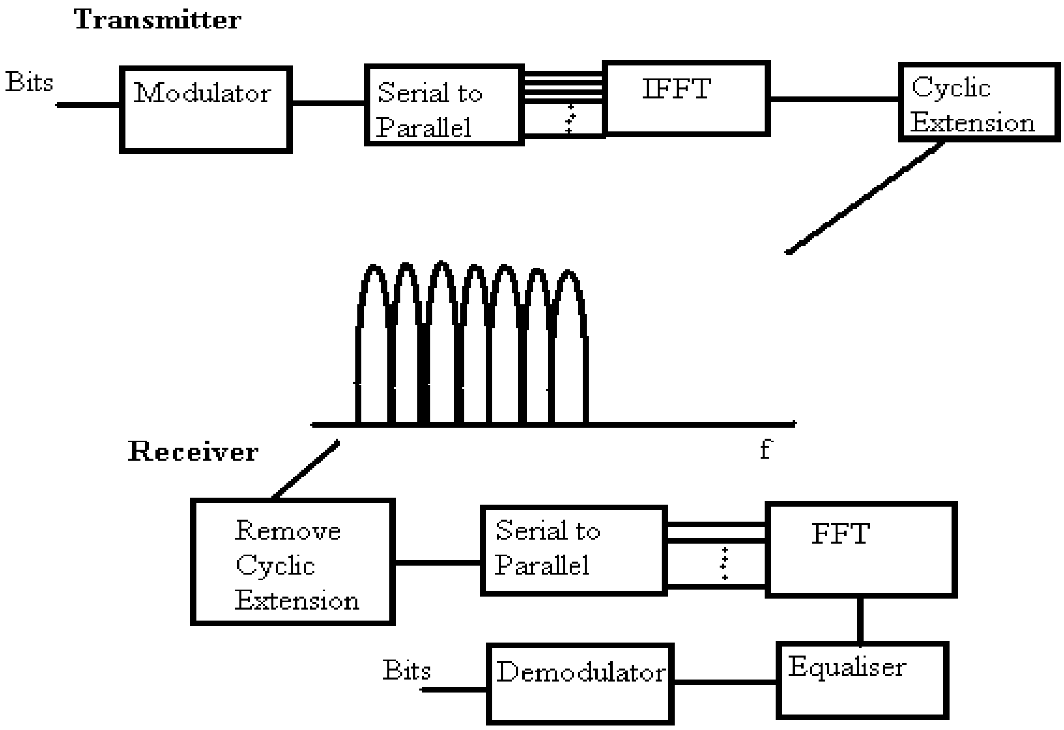

Figure 4.14. OFDMA transmitter and receiver diagrams

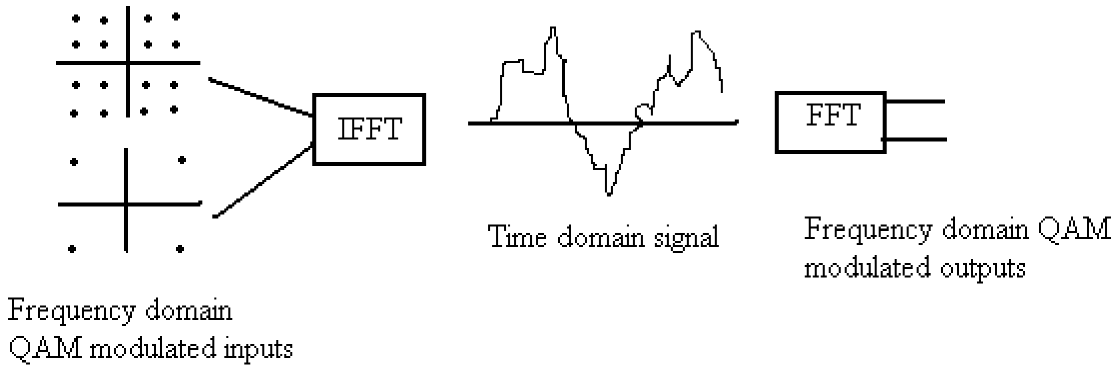

Figure 4.15. IFFT/FFT principle

The modulation of the neighbouring sub-carrier may also be different in terms of the modulation method being used. A practical transmitter needs also additional blocks for the operation, such as windowing operation to satisfy the spectrum mask. While the ideal OFDMA signal has nice spectral properties with steep sidelobes, the resulting non-idealists, and especially clipping the peak amplitudes away, will raise the sidelobes; thus, similar function to the pulse shape filtering of WCDMA is needed.

The use of windowing corresponds to the multiplication of the transmitted signal with the particular shaped window. The practical transmitter, in addition to the windowing functionality, has a cyclic prefix to combat the inter-symbol interference. One could consider the guard interval as a kind of break in the transmission, but actually the waveform is continuous. The transmitter takes part of the waveform to be transmitted and copies part of that to the beginning of the symbol to be transmitted. This part just needs to exceed the channel impulse response duration, as shown in Figure 4.16, and then there is no inter-symbol interference. The receiver will ignore the particular cyclic prefix added. The receiver will actually see only a single symbol, as the channel impulse response is much smaller than the symbol duration, making the channel influence like a finite impulse response filter, and similar multipath components like with the Rake receiver in WCDMA cannot be seen.

The important element of OFDMA is that the transmission can be located in different places in the frequency domain, whereas with WCDMA, by definition, the transmission bandwidth was always independent of the information bandwidth. This allows use of the frequency-domain element in the scheduling, as shown in figure 4.11.

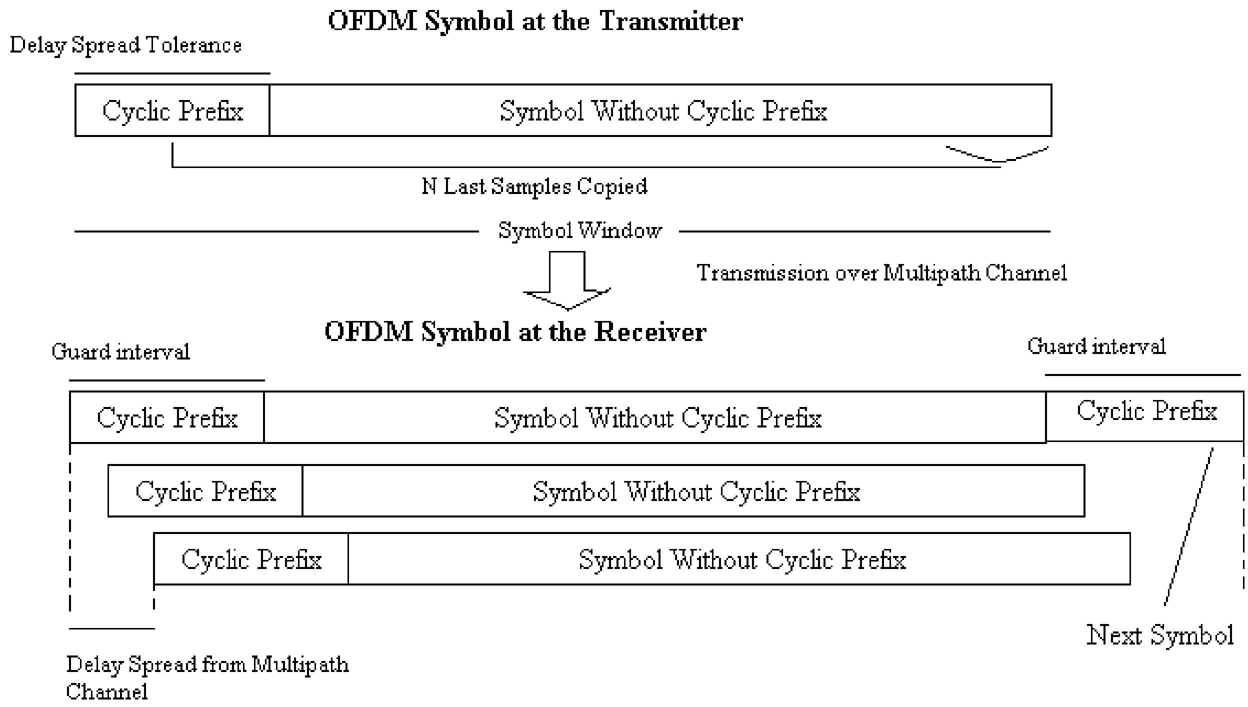

Figure 4.16. Use of cyclic prefix for removing inter-symbol interference

4.4.2.2 SC-FDMA Principles

The SC-FDMA with cyclic prefix has a very simple transmitter structure in the basic form, just a QAM modulator coupled with the addition of the cyclic prefix, as shown Figure 4.17. The signal bandwidth is thus dependent now on the momentary data rate of the modulator in use and only (short) data symbol is being transmitted at the time.

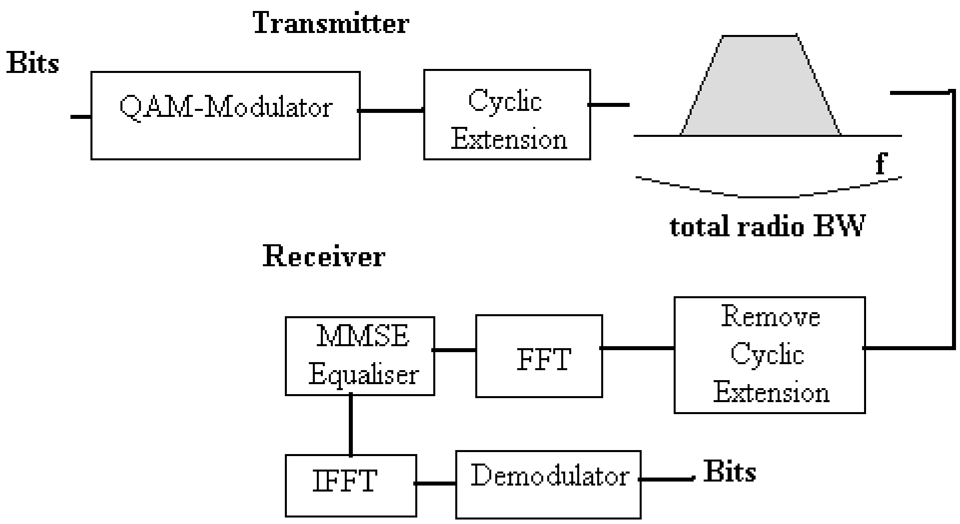

Figure 4.17. Simplified SC-FDMA transmitter and receiver chains

The SC-FDMA with cyclic prefix basically has the following benefits:

- Enabling the low-complexity equaliser receiver by eliminating inter-symbol interference with cyclic prefix.

- But allowing that with low peak to average (PAR) of the signal, as only one information bit is being transmitted at the time; thus, PAR is dominated by the modulation in use.

- Capability to reach a performance similar to OFDMA, assuming an equaliser is being used.

The practical transmitter is likely to take advantage of FFT/IFFT blocks as well to place the transmission in the correct position of the transmit spectrum is case of variable transmission bandwidth, like in LTE. While the maximum transmission bandwidth is up to 20 MHz, the minimum transmission bandwidth is down to 180 kHz, equal to the 12 x 15 kHz sub-carriers in the downlink direction or, rather, one resource block. The different transmitters (terminals) will use the FFT/IFFT pair to place otherwise equal bandwidth transmissions (bandwidth subject to base station scheduler allocations) different uplink frequency blocks by adjusting the "sub-carrier" mapping between FFT and IFFT blocks, as shown in Figure 4.18. This is often referred to in the literature as frequency-domain generalisation of the SC-FDMA transmission.

The receiver side is inherently more complicated, as in OFDMA side (for identical performance) as one needs to combat the multipath interference due to the short symbol duration with the use of an equaliser. The use of a cyclic prefix makes this simpler anyway, and as the equaliser is in the base station side, this does not represent any additional burden for the terminals.

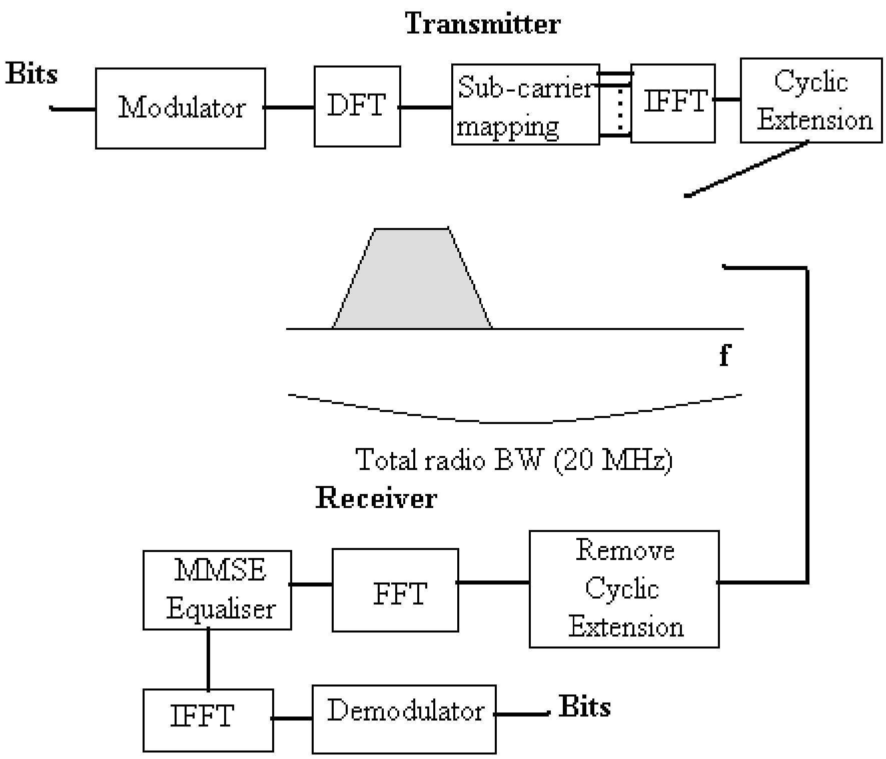

Figure 4.18. SC-FDMA with frequency-domain generation

The use of frequency-domain generation is very practical when the device needs to change the bandwidth and position of the transmission on the uplink frequency allocation. An example is shown in Figure 4.19, with users having identical transmission bandwidth but using the IFFT/FFT blocks in the transmitters to place the resource blocks in the correct place. This is not expected to be in the system constant allocation., but as a response to the base station scheduler commands in the downlink.

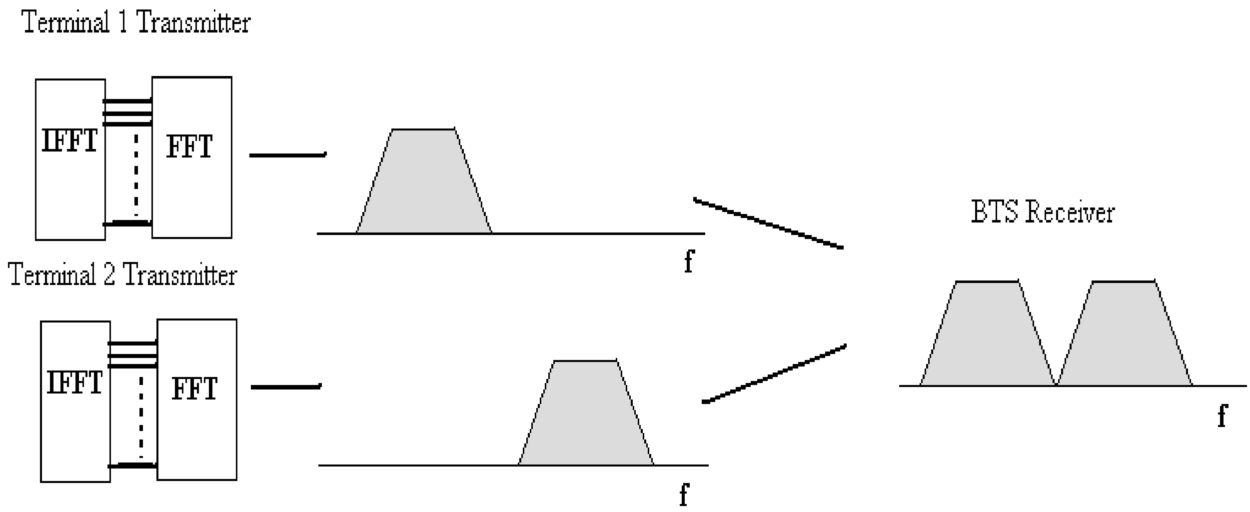

Figure 4.19. FDMA multiplexing of users with FFT/IFFT implemented in the transmitter

4.4.3 Performance

4.4.3.1 Peak Bit Rates

LTE provides high peak bit rates by using large bandwidth up to 20 MHz, high-order 64QAM and multistream MIMO transmission. The downlink peak bit rates can be calculated with Equation (3.1). QPSK modulation carries 2 bits per symbol, 16QAM 4 bits and 64QAM 6 bits. And 2 x 2 MIMO further doubles the peak bit rate. Therefore, QPSK ½ rate coding carries 1 bps/Hz, and 64QAM without any coding and with 2 x 2 MIMO carries 12 bps/Hz. The bandwidth is included in the calculation by taking the corresponding number of sub-carriers for each bandwidth option: 72 per 1.4 MHz and 180 per 3.0 MHz bandwidth. For the bandwidths 5 MHz, 10 MHz and 20 MHz, there are assumed 300, 600 and 1200 sub-carriers respectively. We assume 13 data symbols per 1 ms sub-frame. The achievable peak bit rates are shown in Table 4.8. The highest theoretical data rate is approximately 170 Mbps. If a 4 x 4 MIMO option is applied, then the theoretical peak data rate would double to 340 Mbps.

Peak bit rate [Mbps] = bits/Hz x Number of sub-carriers x Number of symbols per sub-frame/ 1 ms

(3.1)

The uplink peak data rates are shown in Table 4.9: up to 86 Mbps with 64QAM and up to 57 Mbps with 16QAM. The peak rates are lower in uplink than in downlink since single-user MIMO is not specified in uplink. MIMO can be used in uplink as well to increase cell data rates, not single-user peak data rates. The LTE targets of 100 Mbps in downlink and 50 Mbps in uplink are clearly met.

		Peak bit rate per sub-carrier/bandwidth combination				
Modulation coding		**72/1.4 MHz**	**180/3.0 MHz**	**300/5.0 MHz**	**600/10 MHz**	**1200/20 MHz**
QPSK 1/2	Single stream	0.9	2.2	3.6	7.2	14.4
16QAM 1/2	Single stream	1.7	4.3	7.2	14.4	28.8
16QAM 3/4	Single stream	2.6	6.5	10.8	21.6	43.2
64 QAM 3/4	Single stream	3.9	9.7	16.2	32.4	64.8
64 QAM 4/4	Single stream	5.2	13.0	21.6	43.2	86.4
64 QAM 3/4	2 x 2 MIMO	7.8	19.4	32.4	64.8	129.6
64 QAM 4/4	2 x 2 MIMO	10.4	25.9	43.2	86.4	172.8

Table 4.8. Downlink peak bit rates

Modulation coding		Peak bit rate per sub-carrier/bandwidth combination 72/1.4 MHz	180/3.0 MHz	300/5.0 MHz	600/ 10 MHz	1200/20 MHz
QPSK 1/2	Single stream	0.9	2.2	3.6	7.2	14.4
16QAM 1/2	Single stream	1.7	4.3	7.2	14.4	28.8
16QAM 3/4	Single stream	2.6	6.5	10.8	21.6	43.2
16 QAM 4/4	Single stream	3.5	8.6	14.4	28.8	57.6
64 QAM 3/4	Single stream	3.9	9.7	16.2	32.4	64.8
64 QAM 4/4	Single stream	5.2	13.0	21.6	43.2	86.4

Table 4.9. Uplink peak bit rates

4.4.3.2 Spectral Efficiency

The LTE spectral efficiency can be estimated in a number of different ways: using system-level simulation tools, using the Shannon formula combined with the macro-cellular interference distribution or by estimating the expected relative capacity gain over HSPA radio. In this section we present the expected capacity gain over HSPA based on system simulations. The capacity estimates based on the Shannon formula can be found from [9].

The relative spectral efficiency of LTE Release 8 compared with HSPA Release 6 is presented in Figure 4.20. The different bars represent the simulated values from the different companies [10]. The gain of LTE is up to three times compared with HSPA Release 6, which was the original target of LTE.

HSPA+, in Release 7, had already improved the spectral efficiency of the Release 6. The main improvements were mobile equaliser receiver and 2x2 MIMO scheme.

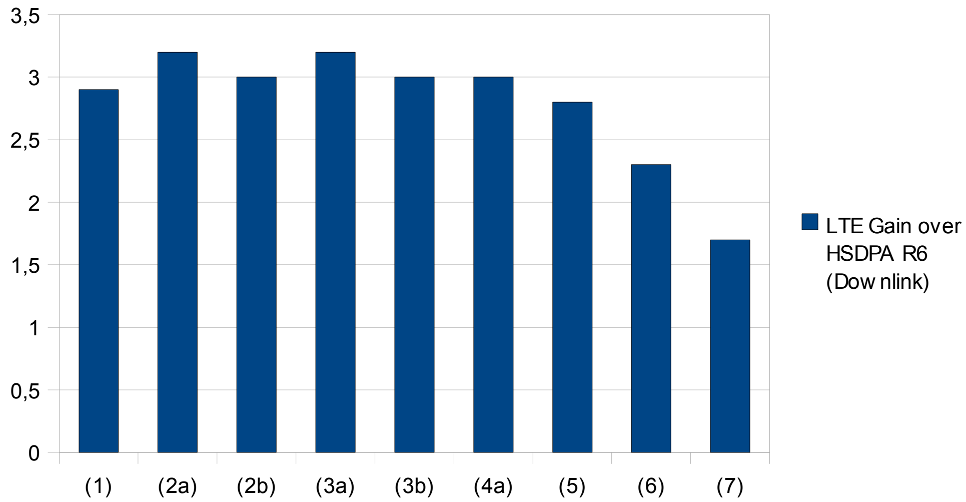

Figure 4.20. Relative spectral efficiency of LTE compared with HSPA R6

There are clear reasons why LTE improves the spectral efficiency compared with HSPA. LTE can exploit such capacity enhancement techniques that are not possible in HSPA or give only limited benefits in HSPA. The OFDMA modulation with frequency-domain scheduling keeps the users orthogonal in LTE. The CDMA transmission in HSPA causes some intra-cell interference in the multipath channel. Part of the intra-cell interference can be removed by the equaliser or by the interference cancellation, but there is still gain provided by OFDMA. The orthogonality is obtained because of long symbols and cyclic prefix in OFDMA. There is no interference between the consecutive symbols as long as the delay spread is shorter than the cyclic prefix. The typical delay spread is a few microseconds, whereas the cyclic prefix is clearly longer 5 or 17 µs in LTE. The HSPA chip is just 0.26 µs and the multipath propagation causes inter-chip interference. On the other hand, the cyclic prefix takes some part of the capacity: 7.5 % in the case of shorter cyclic prefix.

The fast fading is frequency dependent and the typical coherence bandwidth of the signal in macro-cells is in the order of 1 MHz. Within the LTE carrier bandwidth of up to 20 MHz there are some frequencies that are faded and some frequencies that are not faded. In the case of CDMA modulation in HSPA the signal is spread over the whole transmission bandwidth. The spreading gives frequency diversity gain by averaging over the different fading conditions. The spreading, however, is not the most efficient way of transmission. Ideally, the transmission should be done only using those frequencies that are not faded. This is possible in OFDMA in LTE. The transmission can be scheduled by resource blocks, each 180 kHz. The scheduling is based on the mobile channel quality reporting that indicates the most favourables resource blocks. Then the scheduler uses the most favourable resources not only in time domain, but also in frequency domain.

The interference rejection combining or interference cancellation in OFDMA can operate per narrowband sub-carrier. Those algorithms tend to perform better with narrowband sub-carriers, compared with the wideband HSPA carrier.

References

[1] http://www.3GPP.org. 3GPP TS 25.331: RRC Protocol Specification

[2] 3GPP Technical Specification 25.211, Physical Channels and Mapping of Transport Channels onto Physical Channels (FDD), version 5.0.0, March 2002.

[3] 3GPP Technical Specification 25.322, Radio Link Control (RLC), December 2003.

[4] 3GPP Technical Report, TR 25.896, Feasibility Study for Enhanced Uplink for UTRA FDD (Release 6), Version 6.0.0, March 2003.

[5] 3GPP Technical Specifications, TS 25.211, Physical channels and mapping of transport channels onto physical channels (FDD), Version 6.7.0 December 2005.

[6] Holma, H., Kuusela, M., Malkamäki, E., Ranta-aho, K. and Tao C., "VoIP over HSPA with 3GPP Release 7", PIMRC2006, September 2006.

[7] 3GPP Technical Report, TR 25.913, version 2.1.0. "Requirements for Evolved UTRA and UTRAN", June 2005.

[8] Czylwik, A. "Comparison between adaptative OFDM and single carrier modulation with frequency domain equalization", IEEE Vehicular Technology Conference 1997, VTC-97, Phoenix, pp. 863-869.

[9] Mogensen, P., Na, W., Kovacs, I., Frederiksen, F., Pokhariyal, A., Pedersen, K., Kolding, T., Hugl, K. and Kuusela, M., "LTE capacity compared to the Shannon bound". VTC 2007, Spring 2007.

[10] 3GPP Technical Report 25.814, Physical layer aspects of evolved Universal Terrestrial Radio Access (UTRA), version 7.1.0, September 2006.

Chapter 5

Conclusions

One general idea this thesis tries to show is the huge improvements and advances that telecommunication field has experienced in so few time. Thus, it describes briefly the different mobile telecommunication technologies that have existed, from the most primitive with the so-called Zero Generation until the last one which is not deployed yet, Fourth Generation. This way, the reader can realize that, although no many years have passed, this field has advanced in a exponential way. So, after the brief explanation about the telecommunication history, the Third Generation is introduced, what suppose a big step forward with regard to Second Generation. During the thesis, several improvements are presented in the path towards Fourth Generation, each one better than the previous one and its aim is always the same: getting higher data rates, lower power consumption, better throughput, less delay and latency, covering more users and in summary, offering a better service.

During the telecommunication history, different technologies have been deployed, most of them based in one of the three main multiple access techniques: FDMA, TDMA, CDMA. There are many extensions, and hybrid techniques for these methods, such as OFDM, and hybrid TDMA and FDMA systems. An understanding of these three major methods is required for developing of any extensions to these methods. In any case, if someone studies the telecommunication history will realize that the best and most important advances were getting when a new air interface technique was released. For instance, when GSM was launched with its new air interface technique, TDMA, it meant a revolution in the telecommunication market. Now, it seems OFDMA and SC-FDMA are becoming clear as the definitive access technique and the most efficient. Its throughput and features seem awesome and difficult to improve them but probably in several years another research group will discover another air interface with better performances than OFDMA. This field is in continuous evolution.

But although new air interfaces are being launched, 3GPP ensures the harmonious coexistence between the systems because other systems, especially GSM, are still really present and, in the case of GSM, even it continues growing every year with more and more new subscribers.

Anyway some of the new concepts that are explained in this thesis will take long time before being deployed, as in the case of LTE. The first experimental network will be launched during 2011 or 2012. On the other hand, many subscribers are already enjoying the bigger data rates in the downlink direction with HSDPA, which was released in 2002. Some users are also exploiting the new possibilities in the uplink direction with HSUPA, released during 2005. But still, there are not so much users taking advantage of HSPA+, since it was launched during 2007 and few networks count on this new concept.

In summary, in this thesis the author tried to collect as information as was possible, but mainly information about cutting-edge technology because it seems quite probable that mobile phone will become a valuable tool for accessing to the Web thanks to the enhancements of the data rates.

Bibliography

Clark, M. “Wireless Access Networks”. Ed. Wiley.

Niemi, V., Nyberg, K. “UMTS security”. Ed. Wiley.

Schiller, J. “Mobile Communications”. Ed. Addison-Wesley.

www.3gpp.org/

Holma, H., Toskala, A. “WCDMA for UMTS -HSPA evolution and LTE”. Ed. Wiley.

www.itu.int/imt-2000

http://affordablephones.net/HistoryCellular.htm

http://en.wikipedia.org

http://www.umtsworld.com/technology/power.htm

http://epubl.luth.se

http://hosting.udlap.mx/profesores/luisg.guerrero/Cursos/MaterialGeneraldeApoyo/simuladortramas/UMTS/handover.htm

http://www.revistasic.com/revista43/

http://www.3g4g.co.uk/Tutorial/ZG/zg_handover.html

www.revistasic.com/revista43/agorarevista_43.htm

www.cisco.com/.../moblwrls/cmx/mmg_sg/cmxgsm.htm

https://www.ntt-review.jp/archive/ntttechnical.php

http://wireless.agilent.com/rfcomms/refdocs/wcdma/wcdma_gen_bse_hsupa_concepts.php http://www.nec-mobilesolutions.com/infrastructures/solution/hsupa.html

http://www.openbasestation.org/Newsletters/June2006/Explaining%20HSUPA%20-Aeroflex.htm

www.umts-forum.org

www.ericsson.com/ericsson/corpinfo/publications/review/1999_02/49.shtml

www.s3.kth.se/signal/grad/OFDM/URSIOFDM9808.htm

http://www.unibague.edu.co/portal/programas/ingenieria_electronica/el_oraculo_wlan_wpan/telemovil.htm

www.wifinotes.com/3g-3rd-generation.html

www.comsysmobile.com

www.waset.org

www.eurasip.org/Proceedings/Eusipco/Eusipco2005/defevent/

scholar.google.com/

www.dsprelated.com

www.4g-wirelessevolution.tmcnet.com

www.nokia.com