WILEY

GSM

Switching, Services and Protocols

Jörg Eberspächer
and Hans-Jörg Vögel

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

GSM

Switching, Services and Protocols
Second Edition

GSM

Switching, Services and Protocols
Second Edition

Jörg Eberspächer
Technische Universität München, Germany

Hans-Jörg Vögel
The Fantastic Corporation, Switzerland

and

Christian Bettstetter
Technische Universität München, Germany

JOHN WILEY & SONS, LTD
Chichester · New York · Weinheim · Brisbane · Singapore · Toronto

Originally published in the German language by B. G. Teubner GmbH as "Jörg Eberspächer/Hans-Jörg Vögel/Christian Bettstetter: GSM Global System for Mobile Communication. 3. Auflage (3rd edition)".

Baffins Lane, Chichester,
West Sussex, PO19 1UD, England
National 01243 779777
International (+44) 1243 779777

e-mail (for orders and customer service enquiries): cs-books@wiley.co.uk

Visit our Home Page on http://www.wiley.co.uk or http://www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons, Inc., 605 Third Avenue,
New York, NY 10158-0012, USA

WILEY-VCH Verlag GmbH
Pappelallee 3, D-69469 Weinheim, Germany

John Wiley & Sons Australia, Ltd, 33 Park Road, Milton,
Queensland 4064, Australia

John Wiley & Sons (Canada) Ltd, 22 Worcester Road
Rexdale, Ontario, M9W 1L1, Canada

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01,
Jin Xing Distripark, Singapore 129809

Library of Congress Cataloging-in-Publication Data

Eberspächer, I. (Jörg)
[GSM, Global System for Mobile Communication. English]
GSM switching, services, and protocols / Jörg Eberspächer, Hans-Jörg Vögel,
Christian Bettstetter.– 2nd ed.
p. cm.
Includes bibliographical references and index.
Prey. ed.: GSM switching, services, and protocol. 1999.
ISBN 0-471-49903-X (alk. paper)
1. Global system for mobile communications. I. Vögel, Hans-Jörg. II. Bettstetter,
Christian. III Title.

TK5103.483 .E2413 1999
621.382′2–dc2l

00-054550

Use the Internet and eliminate mail time and postage costs http://cip.loc.gov/cip

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN 0471 49903 X

Typeset by Deerpark Publishing Services Ltd, Shannon, Ireland
Printed and bound in Great Britain by Biddles Ltd, Guildford, U.K.

This book is printed on acid-free paper responsibly manufactured from sustainable forestry, in which at least two trees are planted for each one used for paper production.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

Contents

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

Preface for Second Edition

"GSM – the story goes on" is the new title of the last chapter of this book – and GSM is indeed an ongoing success story. Since the release of the first edition of this book (2 years ago), the number of GSM subscribers has grown from 100 to 380 million worldwide. Nobody expected such an enormous number when the first GSM networks started their operation in 1991! In some countries the number of cellular phones is already higher than the number of fixed phones.

Not only are the subscriber numbers experiencing a tremendous growth, but the technological evolution of GSM is also continuing. Many new services and applications have been developed and standardized during the last few years and are now being implemented in GSM networks and terminals.

Substantial progress has been achieved, for example, by improving the voice services. Enhanced speech codecs, such as the *Enhanced Full-Rate* (EFR) and the *Adaptive Multi-Rate* (AMR) codecs, provide better speech quality. Moreover, services for group communication have been developed, which are especially useful for closed user groups. Service platforms (e.g. CAMEL and the *SIM Application Toolkit*) allow network operators to quickly introduce new services.

In addition to speech communication, the mobile data traffic is growing. Several billion text messages are being exchanged between mobile users each month with the GSM *Short Message Service* (SMS). Indeed, the field for GSM data applications and products is huge: news services, mobile payment with cellular phones, telemetry, fleet management, location-based information services, and automatic emergency call systems are just some examples of the broad range of services that became possible with GSM.

In the future, mobile access to the Internet will be of particular importance. The *Wireless Application Protocol* (WAP) has been developed to create an "information Web" for cellular phones. WAP applications, such as stock broking and online auctions, enjoy an increasing popularity. The introduction of the *General Packet Radio Service* (GPRS) – with its packet switched transmission technology at the air interface – enables more efficient, faster, and easier access to the worldwide Internet. GPRS will contribute to the soft migration from GSM toward third generation mobile systems (UMTS, IMT-2000). The world of mobile communications remains exciting!

This second edition of our book gave us the opportunity to include the new GSM technologies. They are treated in Chapters 11 and 12. Chapter 11 is completely new and explains in detail the *General Packet Radio Service* (GPRS). Chapter 12 gives an overview of services recently introduced in GSM Phase 2+. It covers new speech and data services, supplementary services, location services, service platforms, WAP, *Advanced Speech Call*

Items (ASCI), and gives an outlook toward UMTS. Some other chapters have been updated and slightly modified.

We are grateful to Professor Gottfried R. Luderer and Christoph Schmelz for the proofreading of some chapters as well as to Sarah Hinton and the other people from Wiley for the good cooperation.

Last but not least, we would like to thank our readers for many comments and suggestions that have reached us. Their feedback greatly helped us to refine and enhance the book and to correct some errors. We are looking forward to staying in contact with you!

Munich, March 2001

Jörg Eberspächer
joerg.eberspaecher@ei.tum.de

Hans-Jörg Vögel
h.voegel@fantastic.com

Christian Bettstetter
christian.bettstetter@ei.tum.de

PS: Please visit our book's Web page at http://www.lkn.ei.tum.de/gsm_buch with comments, news, and errata.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

Preface

GSM is much more than the acronym of Global System for Mobile Communication; it stands for an extraordinarily successful stage of development in modern information technology. GSM means a new dimension for more than 50 million users – and there are more and more every day – a dimension of personal communication. Today GSM is deployed in more than 100 countries and by over 220 network operators, many of them outside Europe. The mobile telephone has advanced from status symbol to useful appliance, not only in business but also in private everyday life. Its principal use is for wireless telephony, but GSM data communication is increasingly gaining importance.

This modern digital system for mobile communication is based on a set of standards, which were worked out in Europe and can now be considered truly global. Many of the new standardization initiatives of GSM Phase 2+ are in fact coming from outside of Europe. Depending on locally available frequency bands, different GSM air interfaces are defined (e.g. for 900 MHz, 1800 MHz, and 1900 MHz). However, architecture and protocols, in particular for user–network signaling and global roaming are identical in all networks. Thus, GSM enables worldwide development, manufacturing and marketing of innovative products, that stand up well under competition.

GSM also stands for complexity. Whether in the terminals or the exchange equipment, whether in hardware or software, GSM technology is extraordinarily involved and extensive; certainly the most complex communication systems by themselves comprise the standards published by the European Telecommunication Standards Institute (ETSI).

This book arose from an effort to explain and illustrate the essential technical principles of GSM in spite of this complexity, and to show the interrelations between the different subfunctions in a better way than is possible in the framework of standards. Points of crystallization were provided by our course "Communication Networks 2" at the Munich University of Technology as well as our GSM lab course, which requires the students to prepare by studying an extensive GSM manuscript. This lab course is also part of the English graduate program in "communications engineering" at our university which is leading to an MSc degree. The foundation of this book is, however, in the ETSI standards themselves (besides some scientific publications), which were, on one hand, "boiled down" in this book and, on the other hand, augmented by explanations and interpretations.

The book is intended for all those who want to acquire a deeper knowledge of the complex GSM system without losing their way in the detail and wording of the standards. Addressed are the students of electrical engineering, computer science, and information technology at universities and technical institutes, those in industry or network operations

who use and apply the technology, but also researchers who want to gain insight into the architecture and functional operation of the GSM system.

In accordance with the publisher and editors, our book presents the entire architecture of GSM with concentration on the communication protocols, the exchange technology, and the realization of services. The most important principles of the GSM transmission technology are also included in order to give a rounded treatment. Those who are involved with the implementation of GSM systems should find the book to be a useful start and they should find adequate guidance on the standards. The study of the standards is also recommended when there are doubts about the latest issues of the ETSI standards, for with this book we had to consider the standards to be "frozen" in their state as of summer 1997.

The authors especially thank Professor Martin Bossert (Ulm University) for many helpful hints and clarifying discussions. We are very grateful to Professor Gottfried R. Luderer (Arizona State University, Tempe, AZ) for the translation of the German version of the book as well as for the critical technical review of the manuscript and numerous proposals for improvement. It was his strong commitment and determined translation work, which made this book possible. We also give our cordial thanks to the people at Wiley for initiating this book and for the smooth cooperation. Their support in every phase of the project was critical to its speedy production and publication.

The authors are grateful in advance for any kind of response to this book. Readers should address us (wireless or over guided media), preferably via email.

Munich, July 1998

Jörg Eberspächer
Joerg.Eberspaecher@ei.tum.de
Hans-Jörg Vögel
Hans-Joerg.Voegel@ei.tum.de

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

1 Introduction

1.1 Digital, Mobile, Global: Evolution of Networks

Communication everywhere, with everybody, and at any time – we have come much closer to this goal during the last few years. Digitalization of communication systems, enormous progress in microelectronics, computers, and software technology, inventions of efficient algorithms and procedures for compression, security, and processing of all kinds of signals, as well as the development of flexible communication protocols have been important prerequisites for this progress. Today, technologies are available that enable the realization of high-performance and cost-effective communication systems for many application areas.

In the field of fixed networks – where the end systems (user equipment) are connected to the network over a line (two-wire copper line, coaxial cable, glass fiber) – new network technologies (such as xDSL and cable modem) have been introduced, providing broadband access to the Internet.

The largest technological and organizational challenge is, however, the support of subscriber mobility. It can be distinguished between two kinds of mobility: terminal mobility and personal mobility.

In the case of terminal mobility, the subscriber is connected to the network in a wireless way – via radio or light waves – and can move with his or her terminal freely, even during a communication connection. The degree of mobility depends on the type of mobile radio network. The requirements for a cordless in-house telephone are much less critical than for a mobile telephone that can be used in a car or train. If mobility is to be supported across the whole network (or country) or even beyond the network (or national) boundaries, additional switching technology and administrative functions are required, to enable the subscribers to communicate in wireless mode outside of their home areas.

Such extended network functions are also needed to realize personal mobility and universal reachability. This is understood to comprise the possibility of location-independent use of all kinds of telecommunication services – including and especially in fixed networks. The user identifies himself or herself (the person), e.g. by using a chip card, at the place where he or she is currently staying and has access to the network. There, the same communication services can be used as at home, limited only by the properties of the

local network or terminal used. A worldwide unique and uniform addressing is an important requirement.

In the digital mobile communication system GSM (*Global System for Mobile Communication*), which is the subject of this book, terminal mobility is the predominant issue. Wireless communication has become possible with GSM in any town, any country, and even on any continent.

GSM technology contains the essential "intelligent" functions for the support of personal mobility, especially with regard to user identification and authentication, and for the localization and administration of mobile users. Here it is often overlooked that in mobile communication networks by far the largest part of the communication occurs over the fixed network part, which interconnects the radio stations (base stations). Therefore it is no surprise that in the course of further development and evolution of the telecommunication networks, a lot of thought is given to the convergence of fixed and mobile networks.

Today, GSM is used mainly for speech communication, but its use for mobile data communication is growing steadily. The GSM *Short Message Service* (SMS) is a great success story: several billion text messages are being exchanged between mobile users each month. The driving factor for new (and higher bandwidth) data services is the wireless access to the Internet. The key technologies that have been introduced in GSM, the *General Packet Radio Service* (GPRS) and the *Wireless Application Protocol* (WAP), are also explained in this book.

The next generation of mobile communications is known as *Universal Mobile Telecommunication System* (UMTS) in Europe and as *International Mobile Telecommunication System 2000* (IMT-2000) worldwide. The standardization has already progressed quite far, such that the first networks are expected to start operation in 2002. Despite the differences to GSM (in particular with regard to transmission technique and capacity), it is a clear goal of this future network technology to keep the newly introduced GSM technologies and make them essential components of UMTS/IMT-2000.

1.2 Classification of Mobile Communication Systems

This book deals almost exclusively with GSM; however, GSM is only one of many facets of modern mobile communication. Figure 1.1 shows the whole spectrum of today's and – as far as can be seen – future mobile communication systems.

For the bidirectional – and hence genuine – communication systems, the simplest variant is the cordless telephone with very limited mobility (in Europe especially the DECT standard). This technology is also employed for the expansion of digital PBXs with mobile extensions. A related concept is *Radio in the Local Loop* (RLL) or *Wireless Local Loop* (WLL). Both concepts require only limited mobility.

Local Area Networks (LANs) have also been augmented with mobility functions: *Wireless LANs* have been standardized and are now offered by several companies. WLANs offer IP-based, wireless data communication with very high bit rates but limited mobility. IEEE 802.11 systems transmit up to 11 Mbit/s, and HIPERLAN will offer up to 25 Mbit/s. Both systems form pico-cellular networks. They are installed, for example, in office environ-

ments and airports, as supplement or alternative to wired LANs, and they are also considered to be a good supplement to UMTS access technologies. The efforts to "mobilize" the Internet are also worth mentioning in this context. A new routing protocol called Mobile IP [48,49] has been developed, which allows a mobile computer to change its point of attachment to the Internet. A further strong innovation impulse for mobile data and multimedia communication is the development of wireless Mobile ATM systems based on the exchange technology *Asynchronous Transfer Mode* (ATM).

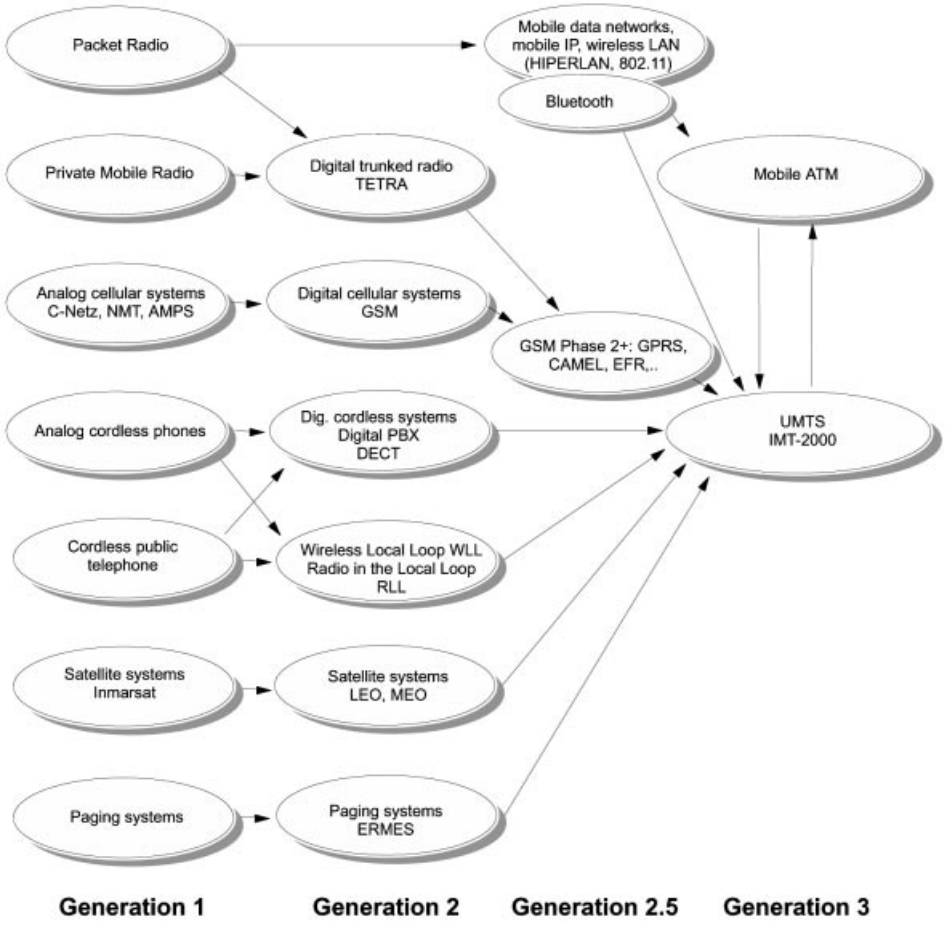

Figure 1.1: Overview of contemporary and future mobile communication systems

Another emerging class of wireless networks is used for short-range communication. Bluetooth, for example, replaces cables by enabling direct wireless information exchange between electronic devices (e.g. between cellular phones, *Personal Digital Assistants* (PDAs), computers, and peripherals). These networks are also called *Body Area Networks* or *Personal Area Networks*. Unlike the mobile technologies mentioned above, they are not based on a fixed network infrastructure (e.g. base stations). The possibility of building up

such networks in a spontaneous and fast way gave them the name *ad hoc networks*. WLAN technologies also include the capability for peer-to-peer ad hoc communication (besides the classical client-to-base station transmission modus).

GSM belongs to the class of cellular networks, which are used predominantly for public mass communication. They had an early success with analog systems like the *Advance Mobile Phone System* (AMPS) in America, the *Nordic Mobile Telephone* (NMT) in Scandinavia, or the *C-Netz* in Germany. Founded on the digital system GSM (with its variants for 900 MHz, 1800 MHz, and 1900 MHz), a market with millions of subscribers worldwide was generated, and it represents an important economic force. A strongly contributing factor to this rapid development of markets and technologies has been the deregulation of the telecommunication markets, which allowed the establishment of new network operators.

Another competing or supplementing technology is satellite communication based on *Low Earth Orbiting* (LEO) or *Medium Earth Orbiting* (MEO) satellites, which also offers global, and in the long term even broadband, communication services. Trunked radio systems – in digital form with the European standard *Trans European Trunked Radio* (TETRA) – are used for business applications like fleet control. They offer private services that are only accessible by closed user groups.

Besides bidirectional communication systems, there also exists a variety of unidirectional systems, where subscribers can only receive but not send data. With unidirectional message systems (paging systems) users may receive short text messages. A couple of years ago, paging systems were very popular, since they offered a cost-effective reachability with wide-area coverage. Today, the SMS in GSM has replaced the function of paging systems. Some billion SMS messages are being exchanged between mobile GSM users each month. Digital broadcast systems, such as *Digital Audio Broadcast* (DAB) and *Digital Video Broadcast* (DVB), are very interesting for wireless transmission of radio and television stations as well as for audio- and video-on-demand and broadband transmission of Internet pages.

The path to the future universal telecommunication networks (UMTS/IMT-2000) has been opened with the realization of the personal communication services, *Universal Personal Telecommunication* (UPT), based on intelligent networks. During the last few years, the huge success of GSM as well as the exploding number of Internet users gave the design and development of third generation mobile systems a new orientation: One of the most important goals in the evolution from GSM to UMTS is to offer an efficient and powerful mobile access to the Internet.

GSM and its enhancements, however, will remain for many years the technological base for mobile communication, and it continues to open up new application areas. At the moment, the area of mobile e-commerce (e.g. mobile payment with cellular phones, mobile banking) is particularly attractive. Also text-based news services, locating, fleet management, telemetry applications, and automatic emergency call systems are of great interest. The techniques and procedures presented in this book are the foundation for such innovative applications.

1.3 Some GSM History and Statistics

In 1982 the development of a pan-European standard for digital cellular mobile radio was started by the *Groupe Spécial Mobile* of the CEPT (Conférence Européenne des Administrations des Postes et des Télécommunications). Initially, the acronym GSM was derived from the name of this group. After the founding of the European standardization institute ETSI (*European Telecommunication Standards Institute*), the GSM group became a Technical Committee of ETSI in 1989. After the rapid worldwide proliferation of GSM networks, the name has been reinterpreted as *Global System for Mobile Communication*.

After a series of incompatible analog networks had been introduced in parallel in Europe, e.g. *Total Access Communication System* (TACS) in the UK, *NMT* in Scandinavia, and the *C-Netz* in Germany, work on the definition of a Europe-wide standard for digital mobile radio was started in the late 1980s. The GSM was founded, which developed a set of technical recommendations and presented them to ETSI for approval. These proposals were produced by the *Special Mobile Group* (SMG) in working groups called *Sub Technical Committees* (STCs), with the following division of tasks: service aspects (SMG 01), radio aspects (SMG 02), network aspects (SMG 03), data services (SMG 04), and network operation and maintenance (SMG 06). Further working groups were mobile station testing (SMG 07), IC card aspects (SGM 09), security (SGM 10), speech aspects (SMG 11), and system architecture (SMG 12) [18]. SGM 05 dealt with future networks and was responsible for the initial standardization phase of the next generation of the European mobile radio system, the UMTS. Later, SMG 05 was closed, and UMTS became an independent project and *Technical Body* of ETSI. In the meantime, the *Third Generation Partnership Project* (3GPP) has been founded in cooperation with other standardization committees worldwide. Its goal is the composition of the *Technical Specifications* for UMTS. Finally, in July 2000, ETSI announced the closure of the SMG which has been responsible for setting GSM standards for the last 18 years. Their remaining and further work has been transferred to groups inside and outside ETSI; most of the ongoing work has been handed over to the 3GPP.

After the official start of the GSM networks during the summer of 1992 (Table 1.1), the number of subscribers has increased rapidly, such that during the fall of 1993 already far more than one million subscribers made calls in GSM networks, more than 80% of them in Germany. On a global scale, the GSM standard also received very fast recognition, as evident from the fact that at the end of 1993 several commercial GSM networks started operation outside Europe, in Australia, Hong Kong, and New Zealand. Afterward, GSM has also been introduced in Brunei, Cameroon, Iran, South Africa, Syria, Thailand, USA and United Arab Emirates. Whereas the majority of the GSM networks operate in the 900 MHz band (GSM900), there are also networks operating in the 1800 MHz band (GSM1800) – *Personal Communication Network* (PCN), *Digital Communication System* (DCS1800) – and in the United States in the 1900 MHz band (GSM1900) – *Personal Communication System* (PCS). These networks use almost completely identical technology and architecture; they differ essentially only in the radio frequencies used and the pertinent high-frequency technology, such that synergy effects can be taken advantage of, and the mobile exchanges can be constructed with standard components.

In parallel to the standardization efforts of ETSI, already in 1987 the then existing prospec-

Table 1.1: Time history – milestones in the evolution of GSM

Year	Event
1982	Groupe Spécial Mobile established by the CEPT.
1987	Essential elements of wireless transmission are specified, based on prototype evaluation (1986). Memorandum of Understanding (MoU) Association founded in September with 13 members from 12 countries.
1989	GSM becomes an ETSI Technical Committee (TC).
1990	The Phase 1 GSM900 specifications (designed 1987–1990) are frozen. Adaptation to DCS1800 commences.
1991	First GSM networks launched. The DCS1800 specifications are finalized.
1992	Most European GSM networks turn commercial by offering voice communication services. Some 13 networks in 7 countries are “on air” by the end of the year.
1993	First roaming agreements in effect. By the end of 1993, 32 networks in 18 countries are operational.
1994	Data transmission capabilities launched. The number of networks rises to 69 in 43 different countries by the end of 1994.
1995	MoU counts 156 members from 86 countries. After the GSM standardization Phase 2 including adaptations and modifications for the PCS1900 (Personal Communication System) is passed, the first PCS1900 Network is launched in the USA. Facsimile, data and SMS roaming starts. Video signals are transmitted via GSM for demonstration purposes. An estimated 50 000 GSM base stations are in use all over the world.
1996	January: 120 networks in 71 countries operational. June: 133 networks in 81 countries operational.
1997	July: 200 GSM networks from 109 countries operational, amounting to 44 million subscribers worldwide.
1998	January: 268 GSM networks with 70 million subscribers worldwide. End of 1998: 320 GSM networks in 118 countries with 135 million subscribers worldwide.
1999	*Wireless Application Protocol* (WAP). End of 1999: 130 countries, 260 million subscribers.
2000	August: 362 million users. *General Packet Radio Service* (GPRS).

tive GSM network operators and the national administrations joined in a group whose members signed a common *Memorandum of Understanding* (MoU). The MoU Association was supposed to form a base for allowing the transnational operation of mobile stations using internationally standardized interfaces. In August 2000, the GSM MoU had 394 members which operated GSM networks in 150 countries (see Figure 1.2).

Figure 1.2 illustrates the impressive growth in the number of GSM networks and GSM subscribers. In 1997, 6 years after the commercial start of the first GSM networks, GSM

had 68 million users and thus a share of approx. 28% of the worldwide mobile market. In the following year, the subscriber number almost doubled, and it doubled again by the beginning of 2000. At the time of writing, in September 2000, there were about 380 million subscribers in all three frequency bands (900 MHz, 1800 MHz, 1900 MHz). In total, there were 373 networks in 142 countries in operation. The share of GSM in the worldwide radio communication market has thus grown up to 60% (of 635 million users) and is still rising. If we consider only digital systems, GSM is even more successful; its market share was over 68% in the middle of 2000. The largest market is Europe with 64% of all subscribers, followed by the Asian Pacific region with 28%. Moreover, China and many African and South-American countries are operating GSM networks, which opens up a market with substantial growth possibilities. It is expected that in the year 2003 over 600 million people will be using GSM. Relevant numbers can be obtained from the Web page of the GSM Association at *http://www.gsmworld.com*.

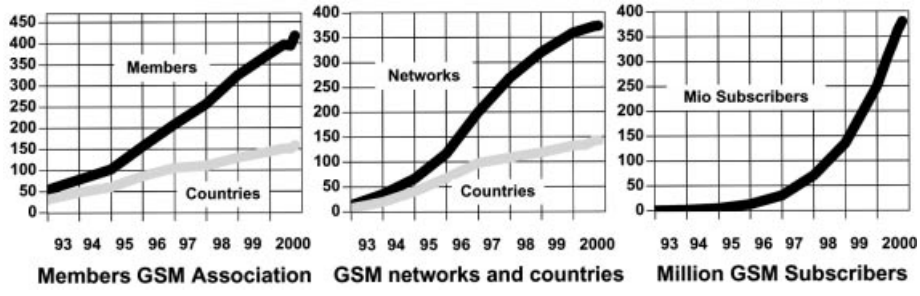

Figure 1.2: GSM network and subscriber statistics. Source: GSM Association, EMC World Cellular Database

All of these networks have implemented Phase 1 of the GSM standard, or the later defined PCN/PCS version of it. In many places, additional services and service characteristics of GSM Phase 2 have also been realized. Phase 1 is essentially the basis for this book, but we will also go into important developments of Phase 2 and Phase 2+.

1.4 Overview of the Book

The remainder of this book is as follows. In Chapter 2, we give an introduction to radio channel characteristics and the cellular principle. The understanding of duplex and multiple access schemes serves as the basis for understanding GSM technology. Chapter 3 introduces the GSM system architecture and addressing. It explains the basic structure and elements of a GSM system and their interfaces as well as the identifiers of users, equipment, and system areas. The GSM services are covered in Chapter 4. Next, Chapter 5 deals with the physical layer at the air interface (How is speech and data transmitted over the radio channel?). Among other things, it describes GSM modulation, multiple access, duplexing, frequency hopping, the logical channels, and synchronization. In Chapter 6, we discuss GSM coding (source coding, speech processing, and channel coding) and mechanisms for authentication and encryption. Chapter 7 covers the entire protocol architecture of GSM (payload transport and signaling). For example, communication protocols for radio

resource management, mobility management, connection management at the air interface are explained. Chapter 8 describes in detail three main principles that are needed for roaming and switching: location registration and update (i.e. How does the network keep track of the user and find him or her when there is an incoming call?), connection establishment and termination, and handover. In Chapter 9 we give an overview of data communication and networking, and Chapter 10 deals with some aspect of network operation. Finally, Chapters 11 and 12 present the latest developments in GSM technology. Chapter 11 explains in detail GPRS which can be used for wireless Internet access. Chapter 12 gives an overview of some more services recently introduced in GSM Phase 2+. It covers new speech services, high-rate data services, supplementary services for speech and location services, service platforms, WAP, and *Advanced Speech Call Items* (ASCI). We conclude this book with an outlook to UMTS.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

2 The Mobile Radio Channel and the Cellular Principle

Many measures, functions and protocols in digital mobile radio networks are based on the properties of the radio channel and its specific qualities in contrast to information transmission through guided media. For the understanding of digital mobile radio networks it is therefore absolutely necessary to know a few related basic principles. For this reason, the most important fundamentals of the radio channel and of cellular and transmission technology will be presented and briefly explained in the following. For a more detailed treatment, see the extensive literature [4,42,50,64].

2.1 Characteristics of the Mobile Radio Channel

The electromagnetic wave of the radio signal propagates under ideal conditions in free space in a radial-symmetric pattern, i.e. the received power P_{Ef}, decreases with the square of the distance L from the transmitter:

$$P_{Ef} \sim \frac{1}{L^2}$$

These idealized conditions do not apply in terrestrial mobile radio. The signal is scattered and reflected, for example, at natural obstacles like mountains, vegetation, or water surfaces. The direct and reflected signal components are then superimposed at the receiver. This multipath propagation can already be explained quite well with a simple two-path model (Figure 2.1). With this model, one can show that the received power decreases much

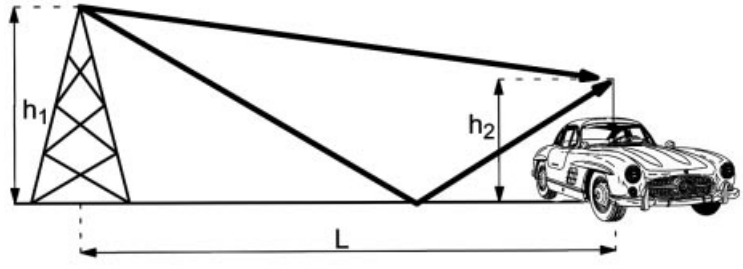

Figure 2.1: Simplified two-path model of radio propagation

more than with the square of the distance from the transmitter. We can approximate the received power by considering the direct path and only one reflected path (two-path propagation) [42]:

$$P_E = P_0 \frac{4}{(4\pi L/\lambda)^2} \left(\frac{2\pi h_1 h_2}{\lambda L} \right)^2 = P_0 \left(\frac{h_1 h_2}{L^2} \right)^2$$

and we obtain, under the simplified assumptions of the two-path propagation model, from Figure 2.1, a propagation loss of 40 dB per decade:

$$\alpha_E = \frac{P_{E2}}{P_{E1}} = \left(\frac{L_1}{L_2} \right)^4, \qquad \alpha_E = 40 \log \left(\frac{L_1}{L_2} \right) \text{ in dB}$$

In reality, the propagation loss depends on the propagation coefficient γ, which is determined by environmental conditions:

$$P_E \sim L^{-\gamma}, \qquad 2 \leq \gamma \leq 5$$

In addition, propagation losses are also frequency dependent, i.e. in a simplified way, propagation attenuation increases disproportionately with the frequency.

However, multipath propagation not only incurs a disproportionately high path propagation loss. The different signal components reaching the receiver have traveled different distances by virtue of dispersion, infraction, and multiple reflections, hence they show different phase shifts. On the one hand, there is the advantage of multipath propagation, that a partial signal can be received even if there is no direct path, i.e. there is no line of sight between mobile and base station. On the other hand, there is a serious disadvantage: the superpositions of the individual signal components having different phase shifts with regard to the direct path can lead, in the worst cases, to cancellations, i.e. the received signal level shows severe disruptions. This phenomenon is called fading. In contrast to this fast fading caused by multipath propagation, there is slow fading caused by shadowing. Along the way traveled by a mobile station, multipath fading can cause significant variations of the received signal level (Figure 2.2). Periodically occurring signal breaks at a distance of about half a wavelength are typically 30–40 dB. The smaller the transmission bandwidth of the mobile radio system, the stronger the signal breaks – at a bandwidth of about 200 kHz per channel this effect is still very visible [8].

Furthermore, the fading dips become flatter as one of the multipath components becomes stronger and more pronounced. Such a dominant signal component arises, for example, in the case of a direct line of sight between mobile and base station, but it can also occur under other conditions. If such a dominant signal component exists, we talk of a Rice channel and Ricean fading, respectively. (S. O. Rice was an American scientist and mathematician.) Otherwise, if all multipath components suffer from approximately equal propagation conditions, we talk of Rayleigh fading. (J. W. Strutt, 3rd Baron Rayleigh, was a British physicist, Nobel prize winner.)

During certain time periods or time slots, the transmission can be heavily impacted because of fading or can be entirely impossible, whereas other time slots may be undisturbed. The results of this effect within the user data are alternating phases, which show either a high or low bit error rate, which is leading to error bursts. The channel thus has

memory in contrast to the statistically independent bit errors in memoryless symmetric binary channels.

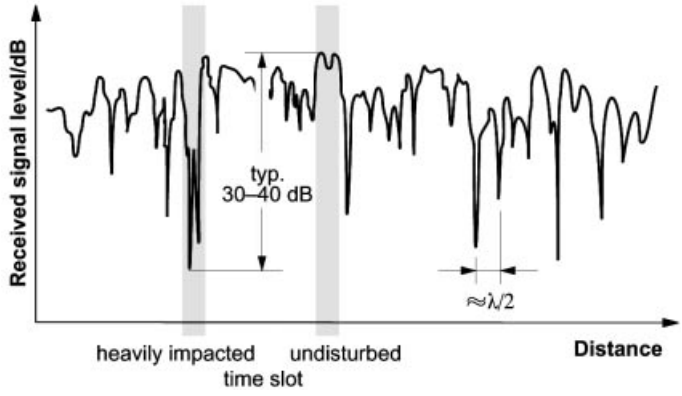

Figure 2.2: Typical signal in a channel with Rayleigh fading

The signal level observed at a specific location is also determined by the phase shift of the multipath signal components. This phase shift depends on the wavelength of the signal, and thus the signal level at a fixed location is also dependent on the transmission frequency. Therefore the fading phenomena in radio communication are also frequency specific. If the bandwidth of the mobile radio channel is small (narrowband signal), then the whole frequency band of this channel is subject to the same propagation conditions, and the mobile radio channel is considered *frequency-nonselective*. Depending on location (Figure 2.2) and the spectral range (Figure 2.3), the received signal level of the channel, however, can vary considerably. On the other hand, if the bandwidth of a channel is large (broadband signal), the individual frequencies suffer from different degrees of fading (Figure 2.3) and this is called a *frequency-selective* channel [15,54]. Signal breaks because of frequency-selective fading along a signal path are much less frequent for a broadband signal than for a narrowband signal, because the fading holes only shift within the band and the received total signal energy remains relatively constant [8].

Besides frequency-selective fading, the different propagation times of the individual multipath components also cause time dispersion on their propagation paths. Therefore, signal distortions can occur due to interference of one symbol with its neighboring symbols (''intersymbol interference''). These distortions depend first on the spread experienced by a pulse on the mobile channel, and second on the duration of the symbol or of the interval between symbols. Typical multipath channel delays have a range from half a microsecond in urban areas to about 16–20 μs in mountainous terrain, i.e. a transmitted pulse generates several echoes which reach the receiver with delays of up to 20 μs. In digital mobile radio systems with typical symbol durations of a few microseconds, this can lead to smearing of individual pulses over several symbol durations.

In contrast to wireline transmission, the mobile radio channel is a very bad transmission medium of highly variable quality. This can go so far that the channel cuts out for short periods (deep fading holes) or that single sections in the data stream are so much interfered

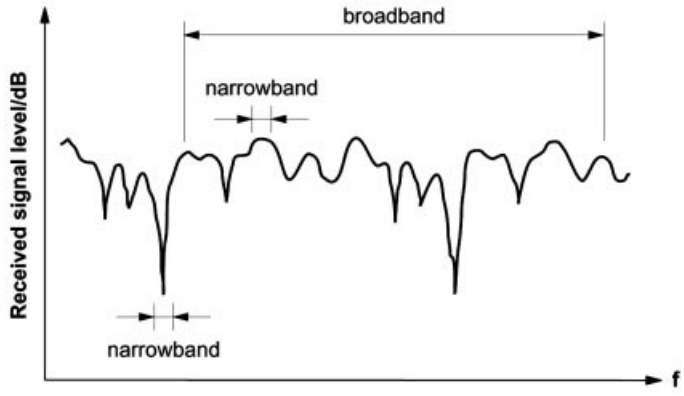

Figure 2.3: Frequency selectivity of a mobile radio channel

with (bit error rate typically 10^{-2} or 10^{-1}), that unprotected transmission without further protection or correction measures is hardly possible. Therefore, mobile information transport requires additional, often very extensive measures, which compensate for the effects of multipath propagation. First, an equalizer is necessary, which attempts to eliminate the signal distortions caused by intersymbol interference. The operational principle of such an equalizer for mobile radio is based on the estimation of the channel pulse response to periodically transmitted, well-known bit patterns, known as the training sequences [4,64]. This allows the determination of the time dispersion of the channel and its compensation. The performance of the equalizer has a significant effect on the quality of the digital transmission. On the other hand, for efficient transmission in digital mobile radio, channel coding measures are indispensable, such as forward error correction with error-correcting codes, which allows reduction of the effective bit error rate to a tolerable value (about 10^{-5} to 10^{-6}). Further important measures are control of the transmitter power and algorithms for the compensation of signal interruptions in fading, which may be of such a short duration that a disconnection of the call would not be appropriate.

2.2 Separation of Directions and Duplex Transmission

The most frequent form of communication is the bidirectional communication which allows simultaneous transmitting and receiving. A system capable of doing this is called full-duplex. One can also achieve full-duplex capability, if sending and receiving do not occur simultaneously but switching between both phases is done so fast that it is not noticed by the user, i.e. both directions can be used quasi-simultaneously. Modern digital mobile radio systems are always full-duplex capable.

Essentially, two basic duplex procedures are employed: *Frequency Division Duplex* (FDD) using different frequency bands in each direction, and *Time Division Duplex* (TDD) which periodically switches the direction of transmission.

2.2.1 Frequency Division Duplex (FDD)

The frequency duplex procedure has been used already in analog mobile radio systems and is also used in digital systems. For the communication between mobile and base station, the available frequency band is split into two partial bands, to enable simultaneous sending and receiving. One partial band is assigned as *uplink* (from mobile to base station) and the other partial band is assigned as *downlink* (from base to mobile station):

- Uplink: transmission band of mobile station = receiving band of base station
- Downlink: receiving band of mobile station = transmission band of base station

To achieve good separation between both directions, the partial bands must be a sufficient frequency distance apart, i.e. the frequency pairs of a connection assigned to uplink and downlink must have this distance band between them. Usually, the same antenna is used for sending and receiving. A duplexing unit is then used for the directional separation, consisting essentially of two narrowband filters with steep flanks (Figure 2.4). These filters, however, cannot be integrated, so pure frequency duplexing is not appropriate for systems with small compact equipment [15].

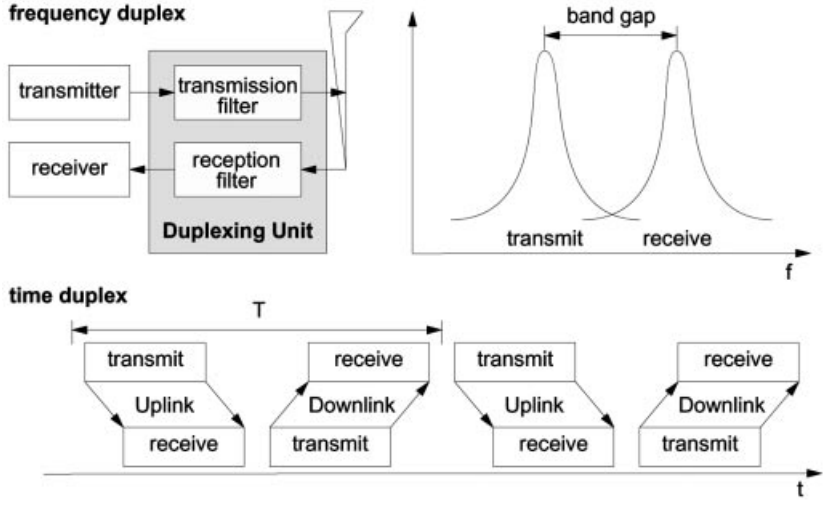

Figure 2.4: Frequency and time duplex (schematic)

2.2.2 Time Division Duplex (TDD)

Time duplexing is therefore a good alternative, especially in digital systems with time division multiple access. Transmitter and receiver operate in this case only quasi-simultaneously at different points in time; i.e. the directional separation is achieved by switching in time between transmission and reception, and thus no duplexing unit is required. Switching occurs frequently enough that the communication appears to be over a quasi-simultaneous full-duplex connection. However, out of the periodic interval T available for the transmission of a time slot only a small part can be used, so that a time duplex system requires more than twice the bit rate of a frequency duplex system.

2.3 Multiple Access Procedures

The radio channel is a communication medium shared by many subscribers in one cell. Mobile stations compete with one another for the frequency resource to transmit their information streams. Without any other measures to control simultaneous access of several users, collisions can occur (multiple access problem). Since collisions are very undesirable for a connection-oriented communication like mobile telephony, the individual subscribers/mobile stations must be assigned dedicated channels on demand. In order to divide the available physical resources of a mobile system, i.e. the frequency bands, into voice channels, special multiple access procedures are used which are presented in the following (Figure 2.5).

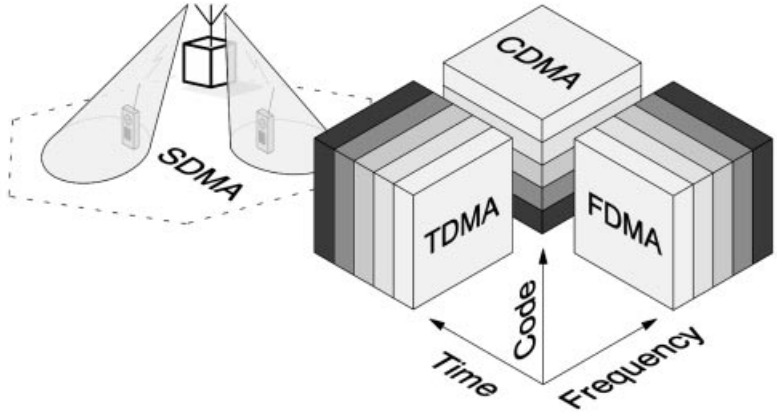

Figure 2.5: Multiple access procedures

2.3.1 Frequency Division Multiple Access (FDMA)

Frequency Division Multiple Access (FDMA) is one of the most common multiple access procedures. The frequency band is divided into channels of equal bandwidth such that each conversation is carried on a different frequency (Figure 2.6). Best suited to analog mobile radio, FDMA systems include the C-Netz in Germany, TACS in the UK, and AMPS in the USA. In the C-Netz, two frequency bands of 4.44 MHz each are subdivided into 222 individual communication channels at 20 kHz bandwidth. The effort in the base station to realize a frequency division multiple access system is very high. Even though the required hardware components are relatively simple, each channel needs its own transceiving unit. Furthermore, the tolerance requirements for the high-frequency networks and the linearity of the amplifiers in the transmitter stages of the base station are quite high, since a large number of channels need to be amplified and transmitted together [15,54]. One also needs a duplexing unit with filters for the transmitter and receiver units to enable full-duplex operation, which makes it nearly impossible to build small, compact mobile stations, since the required narrowband filters can hardly be realized with integrated circuits.

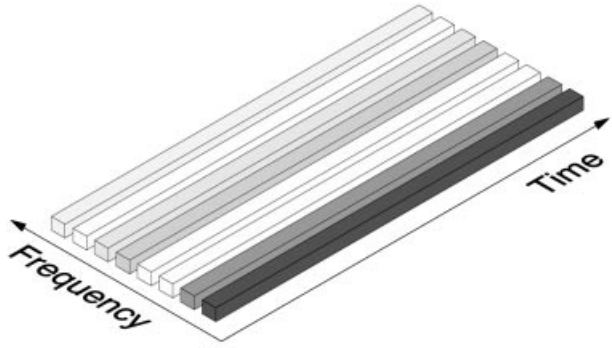

Figure 2.6: Channels of an FDMA system (schematic)

2.3.2 Time Division Multiple Access (TDMA)

Time Division Multiple Access (TDMA) is a more expensive technique, for it needs a highly accurate synchronization between transmitter and receiver. The TDMA technique is used in digital mobile radio systems. The individual mobile stations are cyclically assigned a frequency for exclusive use only for the duration of a time slot. Furthermore, in most cases the whole system bandwidth for a time slot is not assigned to one station, but the system frequency range is subdivided into subbands, and TDMA is used for multiple access to each subband. The subbands are known as carrier frequencies, and the mobile systems using this technique are designated as multicarrier systems (not to be confused with multicarrier modulation). The pan-European digital system GSM employs such a combination of FDMA and TDMA; it is a multicarrier TDMA system. A frequency range of 25 MHz holds 124 single channels (carrier frequencies) of 200 kHz bandwidth each, with each of these frequency channels containing again 8 TDMA conversation channels.

Thus the sequence of time slots assigned to a mobile station represents the physical channels of a TDMA system. In each time slot, the mobile station transmits a data burst. The period assigned to a time slot for a mobile station thus also determines the number of TDMA channels on a carrier frequency. The time slots of one period are combined into a so-called TDMA frame. Figure 2.7 shows five channels in a TDMA system with a period of four time slots and three carrier frequencies.

The TDMA signal transmitted on a carrier frequency in general requires more bandwidth than an FDMA signal, since because of multiple time use, the gross data rate has to be correspondingly higher. For example, GSM systems employ a gross data rate (modulation data rate) of 271 kbit/ s on a subband of 200 kHz, which amounts to 33.9 kbit/ s for each of the eight time slots.

Especially narrowband systems suffer from time- and frequency-selective fading (Figures 2.2 and 2.3) as already mentioned. In addition, there are also frequency-selective co-channel interferences, which can contribute to the deterioration of the transmission quality. In a TDMA system, this leads to the phenomenon that the channel can be very good during one time slot, and very bad during the next time slot when some bursts are strongly interfered with. On the other hand, a TDMA system offers very good opportunities to

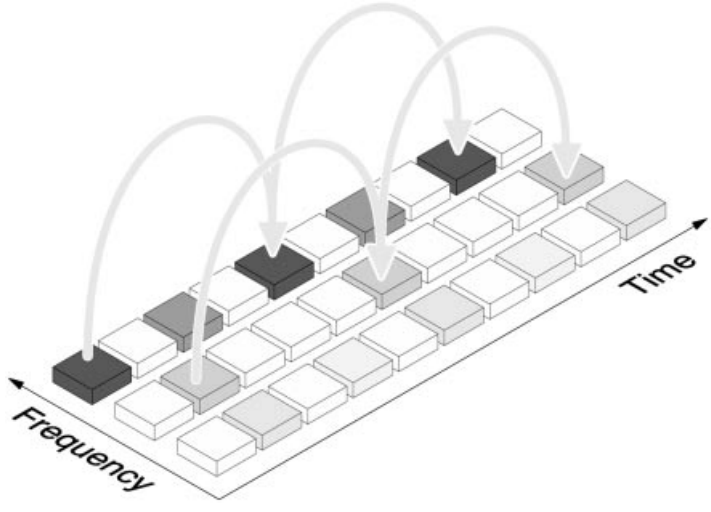

Figure 2.7: TDMA channels on multiple carrier frequencies

attack and drastically reduce such frequency-selective interference by introducing a frequency hopping technique. With this technique, each burst of a TDMA channel is transmitted on a different frequency (Figure 2.8).

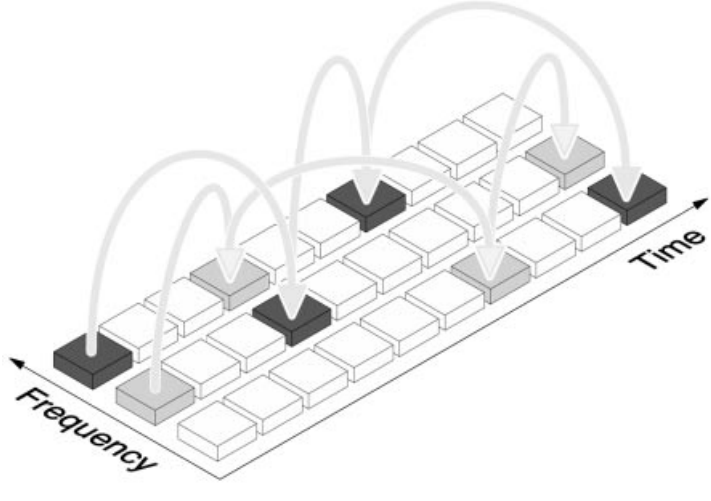

Figure 2.8: TDMA with use of frequency hopping technique

In this technique, selective interference on one frequency at worst hits only every *i*th time slot, if there are *i* frequencies available for hopping. Thus the signal transmitted by a frequency hopping technique uses frequency diversity. Of course, the hopping sequences

must be orthogonal, i.e. one must ascertain that two stations transmitting in the same time slot do not use the same frequency. Since the duration of a hopping period is long compared to the duration of a symbol, this technique is called *slow frequency hopping*. With fast frequency hopping, the hopping period is shorter than a time slot and is of the order of a single symbol duration or even less. This technique then belongs already to the spread spectrum techniques of the family of code division multiple access techniques, *Frequency Hopping CDMA* (FH-CDMA) (see Section 2.3.3).

As mentioned above, for TDM access, a precise synchronization between mobile and base station is necessary. This synchronization becomes even more complex through the mobility of the subscribers, because they can stay at varying distances from the base station and their signals thus incur varying propagation times. First, the basic problem is to determine the exact moment when to transmit. This is typically achieved by using one of the signals as a time reference, like the signal from the base station (downlink, Figure 2.9). On receiving the TDMA frame from the base station, the mobile can synchronize and transmit time slot synchronously with an additional time offset (e.g. three time slots in Figure 2.9).

Another problem is the propagation time of the signals, so far ignored. It also depends on the variable distance of the mobile station from the base. These propagation times are the reason why the signals on the uplink arrive not frame-synchronized at the base, but with variable delays. If these delays are not compensated, collisions of adjacent time slots can occur (Figure 2.9). In principle, the mobile stations must therefore advance the time-offset between reception and transmission, i.e. the start of sending, so much that the signals arrive frame-synchronous at the base station.

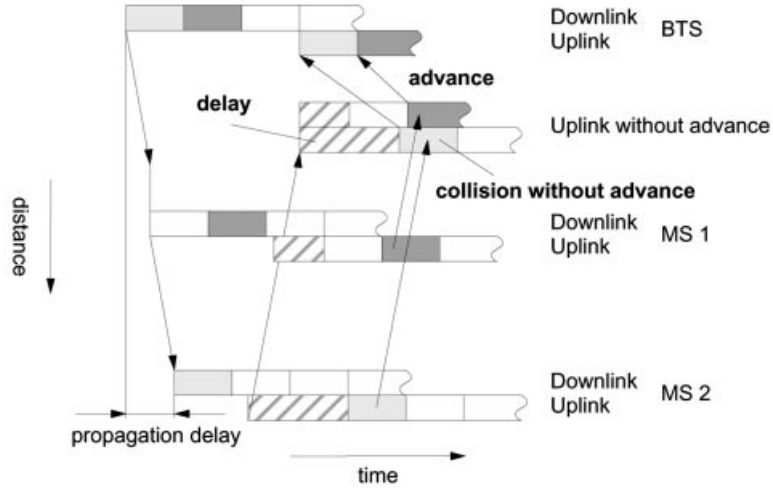

Figure 2.9: Differences in propagation delays and synchronization in TDMA systems

2.3.3 Code Division Multiple Access (CDMA)

Systems with *Code Division Multiple Access* (CDMA) are broadband systems, in which each subscriber uses the whole system bandwidth (similar to TDMA) for the complete duration of the connection (similar to FDMA). Furthermore, usage is not exclusive, i.e. all the subscribers in a cell use the same frequency band simultaneously. To separate the signals, the subscribers are assigned orthogonal codes. The basis of CDMA is a band-spreading or spread spectrum technique. The signal of one subscriber is spread spectrally over a multiple of its original bandwidth. Typically, spreading factors are between 10 and 1000; they generate a broadband signal for transmission from the narrowband signal, and this is less sensitive to frequency-selective interference and disturbances. Furthermore, the spectral power density is decreased by band spreading, and communication is even possible below the noise threshold [15].

2.3.3.1. Direct Sequence CDMA

A common spread-spectrum procedure is the direct sequence technique (Figure 2.10). In it the data sequence is multiplied directly – before modulation – with a spreading sequence to generate the band-spread signal. The bit rate of the spreading signal, the so-called chip rate, is obtained by multiplying the bit rate of the data signal by the spreading factor, which generates the desired broadening of the signal spectrum. Ideally, the spreading sequences are completely orthogonal bit sequences (''codes'') with disappearing cross-correlation functions. Since such completely orthogonal sequences cannot be realized, practical systems use bit sequences from *pseudo noise* (PN) generators to spread the band [15,54]. For despreading, the signal is again multiplied with the spreading sequence at the receiver, which ideally recovers the data sequence in its original form.

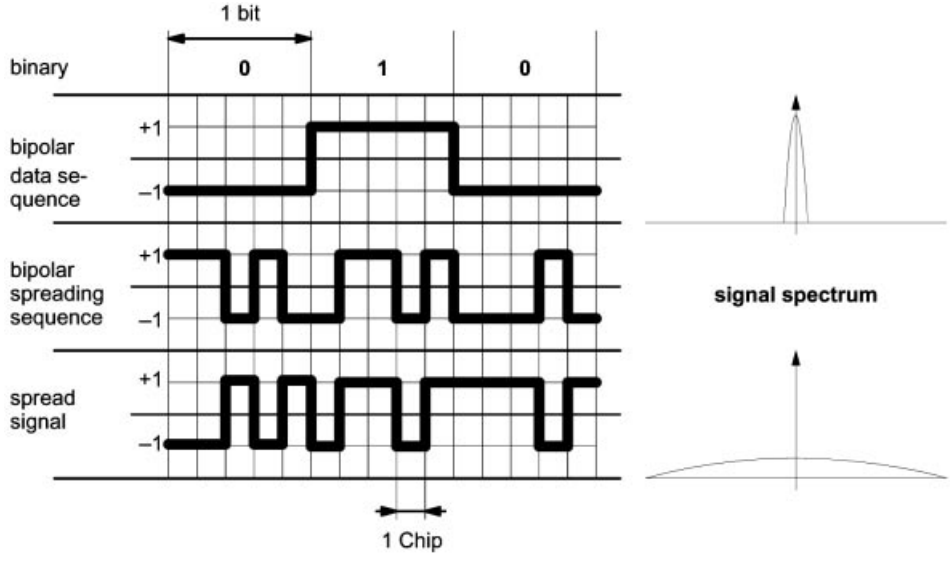

Figure 2.10: Principle of spread spectrum technique for direct sequence CDMA

Thus one can realize a code-based multiple access system. If an orthogonal family of spreading sequences is available, each subscriber can be assigned his or her own unique spreading sequence. Because of the disappearing cross-correlation of the spreading sequences, the signals of the individual subscribers can be separated in spite of being transmitted in the same frequency band at the same time.

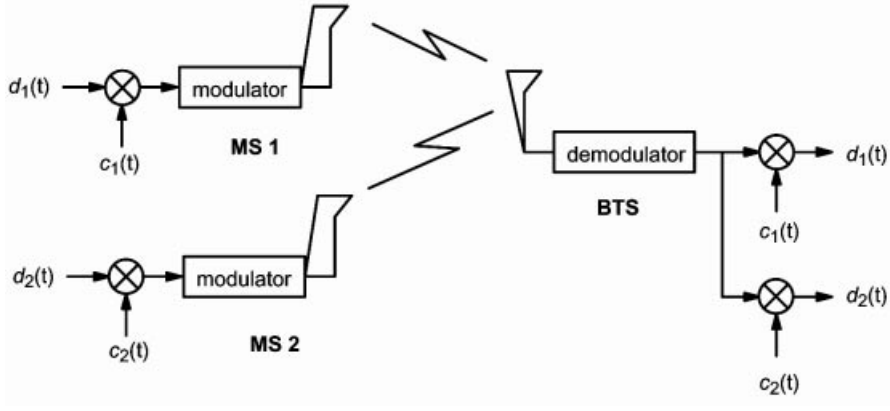

Figure 2.11: Simplified scheme of code division multiple access (uplink)

In a simplified way, this is done by multiplying the received summation signal with the respective code sequence (Figure 2.11):

$$s(t)c_j(t) = c_j(t)\sum_{i=1}^{n} d_i(t)c_i(t) = d_j(t)$$

$$\text{with } c_j(t)c_i(t) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

Thus, if direct sequence spreading is used, the procedure is called *Direct Sequence Code Division Multiple Access* (DS-CDMA).

2.3.3.2. Frequency Hopping CDMA

Another possibility for spreading the band is the use of a fast frequency hopping technique. If one changes the frequency several times during one transmitted data symbol, a similar spreading effect occurs as in case of the direct sequence procedure. If the frequency hopping sequence is again controlled by orthogonal code sequences, another multiple access system can be realized, the Frequency Hopping CDMA (FH-CDMA).

2.3.4 Space Division Multiple Access (SDMA)

An essential property of the mobile radio channel is multipath propagation, which leads to frequency-selective fading phenomena. Furthermore, multipath propagation is the cause of another significant property of the mobile radio channel, the spatial fanning out of signals. This causes the received signal to be a summation signal, which is not only determined by the *Line of Sight* (LOS) connection but also by an undetermined number of individual paths caused by refractions, infractions, and reflections. In principle, the directions of incidence of these multipath components could therefore be distributed arbitrarily at the receiver.

Especially on the uplink from the mobile station to the base station, there is, however, in most cases a main direction of incidence (usually LOS), about which the angles of incidence of the individual signal components are scattered in a relatively narrow range. Frequently, the essential signal portion at the receiver is distributed only over an angle of a few tens of degrees. This is because base stations are installed wherever possible as free-standing units, and there are no interference centers in the immediate neighborhood.

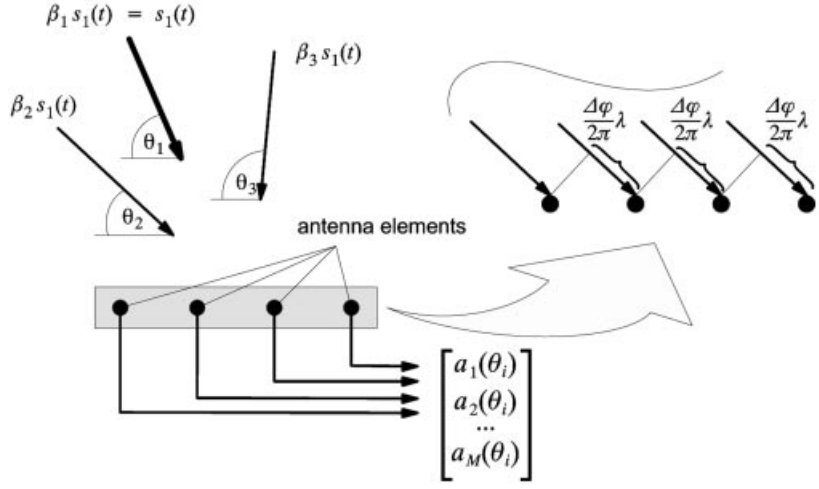

Figure 2.12: Multipath signal at an antenna array

This directional selectivity of the mobile radio channel, which exists in spite of multipath propagation, can be exploited by using array antennas. Antenna arrays generate a directional characteristic by controlling the phases of the signals from the individual antenna elements. This allows the receiver to adjust the antenna selectively to the main direction of incidence of the received signal, and conversely to transmit selectively in one direction. This principle can be illustrated easily with a simple model (Figure 2.12).

The individual multipath components $\beta_i s_1(t)$ of a transmitted signal $s_1(t)$ propagate on different paths such that the multipath components incident at an antenna under the angle θ_i differ in amplitude and phase. If one considers an array antenna with M elements ($M = 4$ in Figure 2.12) and a wave front of a multipath component incident at angle θ_i on

this array antenna, then the received signals at the antenna elements differ mainly in their phase – each shifted by $\Delta\varphi$ (Figure 2.12) – and amplitude.

In this way, the response of the antenna to a signal incident at angle θ_i can be characterized by the complex response vector $\vec{a}(\theta_i)$ which defines amplitude gain and phase of each antenna element relative to the first antenna element ($a_1 = 1$):

$$\vec{a}(\theta_i) = \begin{bmatrix} a_1(\theta_i) \\ a_2(\theta_i) \\ \cdots \\ a_M(\theta_i) \end{bmatrix} = \begin{bmatrix} 1 \\ a_2(\theta_i) \\ \cdots \\ a_M(\theta_i) \end{bmatrix}$$

The N_m multipath components ($N_m = 3$ in Figure 2.12) of a signal $s_1(t)$ generate, depending on the incidence angle θ_i, a received signal vector $\vec{x}_1(t)$ which can be written with the respective antenna response vector and the signal of the ith multipath $\beta_i s_1(t)$ shifted in amplitude and phase against the direct path $s_1(t)$ as

$$\vec{x}_1(t) = \vec{a}(\theta_1)s_1(t) + \sum_{i=2}^{N_m} \vec{a}(\theta_i)\beta_i s_1(t) = \vec{a}_1 s_1(t)$$

In this case, the vector $\vec{a}_1$ is also designated the spatial signature of the signal $s_1(t)$, which remains constant as long as the source of the signal does not move and the propagation conditions do not change [65]. In a multi-access situation, there are typically several sources (N_q); this yields the following result for the total signal at the array antenna: neglecting noise and interferences,

$$\vec{x}(t) = \sum_{j=1}^{N_q} \vec{a}_j s_j(t)$$

From this summation signal, the signals of the individual sources are separated by weighting the received signals of the individual antenna elements with a complex factor (weight vector $\vec{w}_i$), which yields

$$\vec{w}_i^H \vec{a}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

For the weighted summation signal [65] one gets

$$\vec{w}_i^H \vec{x}(t) = \sum_{j=1}^{N_q} \vec{w}_i^H \vec{a}_j s_j(t) = s_i(t)$$

Under ideal conditions, i.e. neglecting noise and interference, the signal $s_i(t)$ of a single source i can be separated from the summation signal of the array antenna by using an appropriate weight vector during signal processing. The determination of the respectively optimal weight vector, however, is a nontrivial and computation-intensive task. Because of the considerable processing effort and also because of the mechanical dimensions of the antenna field, array antennas are predominantly used in base stations.

So far only the receiving direction has been considered. The corresponding principles, however, can also be used for constructing the directional characteristics of the transmitter. Assume symmetric propagation conditions in the sending and receiving directions, and assume the transmitted signals $s_i(t)$ are weighted with the same weight vector $\vec{w}_i$ as the received signal, before they are transmitted through the array antenna; then one obtains the following summation signal radiated by the array antenna:

$$\vec{y}(t) = \sum_{j=1}^{N_q} \vec{w}_j s_j(t)$$

and for the signal received on the ith opposite side, respectively:

$$\hat{s}_i(t) = \vec{a}_i^H \vec{y}(t) = \sum_{j=1}^{N_q} \vec{a}_i^H \vec{w}_j s_j(t) = s_i(t)$$

Thus, by using array antennas, one can separate the simultaneously received signals of spatially separated subscribers by exploiting the directional selectivity of the mobile radio channel. Because of the use of intelligent signal processing and corresponding control algorithms, such systems are also known as systems with intelligent antennas.

The directional characteristics of the array antenna can be controlled adaptively such that a signal is only received or transmitted in exactly the spatial segment where a certain mobile station is currently staying. On the one hand, one can thus reduce co-channel interference in other cells, and on the other hand, the sensitivity against interference can be reduced in the current cell. Furthermore, because of the spatial separation, physical channels in a cell can be reused, and the lobes of the antenna diagram can adaptively follow the movement of mobile stations. In this case, yet another multiple access technique (Figure 2.13) is defined and known as *Space Division Multiple Access* (SDMA).

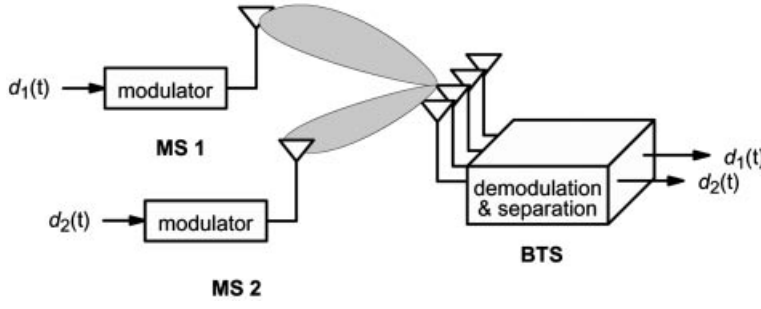

Figure 2.13: Schematic representation of spatial multiple access (uplink)

SDMA systems are currently the subject of intensive research. The SDMA technique can be combined with each of the other multiple access techniques (FDMA, TDMA, CDMA). This enables intracellular spatial channel reuse, which again increases the network capacity [29]. This is especially attractive for existing networks which can use an intelligent implementation of SDMA by selectively upgrading base stations with array antennas, appropriate signal processing, and respective control protocols.

2.4 Cellular Technology

Because of the very limited frequency bands, a mobile radio network has only a relatively small number of speech channels available. For example, the GSM system has an allocation of 25 MHz bandwidth in the 900 MHz frequency range, which amounts to a maximum of 125 frequency channels each with a carrier bandwidth of 200 kHz. Within an eightfold time multiplex for each carrier, a maximum of 1000 channels can be realized. This number is further reduced by guardbands in the frequency spectrum and the overhead required for signaling (Chapter 5). In order to be able to serve several 100 000 or millions of subscribers in spite of this limitation, frequencies must be spatially reused, i.e. deployed repeatedly in a geographic area. In this way, services can be offered with a cost-effective subscriber density and acceptable blocking probability.

2.4.1 Fundamental Definitions

This spatial frequency reuse concept led to the development of cellular technology, which allowed a significant improvement in the economic use of frequencies. The essential characteristics of the cellular network principle are as follows:

- The area to be covered is subdivided into cells (radio zones). For easier manipulation, these cells are modeled in a simplified way as hexagons (Figure 2.14). Most models show the base station in the middle of the cell.
- To each cell i a subset of the frequencies fb_i is assigned from the total set (bundle) assigned to the respective mobile radio network. Two neighboring cells must never use the same frequencies, since this would lead to severe co-channel interference from the adjacent cells.
- Only at distance D (the *frequency reuse distance*) can a frequency from the set fb_i be reused (Figure 2.4), i.e. cells with distance D to cell i are assigned one or all of the frequencies from the set fb_1 belonging to cell i. If D is chosen sufficiently large, the co-channel interference remains small enough not to affect speech quality.
- When a mobile station moves from one cell to another during an ongoing conversation, an automatic channel/frequency change occurs (*handover*), which maintains an active speech connection over cell boundaries.

The spatial repetition of frequencies is done in a regular systematic way, i.e. each cell with the *frequency allocation* fb_i (or one of its frequencies) sees its neighbors with the same frequencies again at a distance D (Figure 2.14). Therefore there exist exactly six such next neighbor cells. Independent of form and size of the cells – not only in the hexagon model – the first ring in the frequency set contains six co-channel cells (see also Figure 2.15).

2.4.2 Signal-to-Noise Ratio

The interference caused by neighboring cells is measured as the signal-to-noise ratio:

$$W = \frac{\text{useful signal}}{\text{disturbing signal}} = \frac{\text{useful signal}}{\text{neighbor cell interference} + \text{noise}}$$

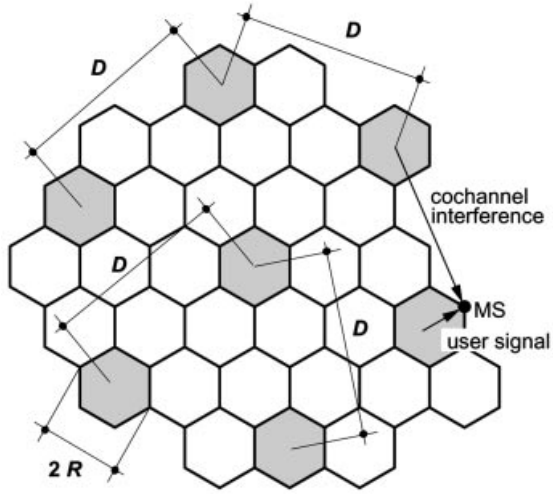

Figure 2.14: Model of a cellular network with frequency reuse

This ratio of the useful signal to the interfering signal is usually measured in decibels (dB) and called the *Signal-to-Noise Ratio* (SNR). The intensity of the interference is essentially a function of co-channel interference depending on the frequency reuse distance D. From the viewpoint of a mobile station, the co-channel interference is caused by base stations at distance D from the current base station. A worst-case estimate for the signal-to-noise ratio W of a mobile station at the border of the covered area at distance R from the base station can be obtained, subject to propagation losses, by assuming that all six neighboring interfering transmitters operate at the same power and are approximately equally far apart (distance D large against cell radius R) [42]:

$$W = \frac{P_0 R^{-\gamma}}{\sum_{i=1}^{6} P_i + N} \approx \frac{P_0 R^{-\gamma}}{\sum_{i=1}^{6} P_0 D^{-\gamma} + N} = \frac{P_0 R^{-\gamma}}{6 P_0 D^{-\gamma} + N}$$

By neglecting the noise N we obtain the following approximation for the *Carrier-to-Interference Ratio C/I* (CIR):

$$W \approx \frac{C}{I} = \frac{R^{-\gamma}}{6 D^{-\gamma}} = \frac{1}{6}\left(\frac{R}{D}\right)^{-\gamma}$$

Therefore the signal-to-noise ratio depends essentially on the ratio of the cell radius R to the frequency reuse distance D. From these considerations it follows that for a desired or needed signal-to-noise ratio W at a given cell radius, one must choose a minimum distance for the frequency reuse, above which the co-channel interference fall below the required threshold.

2.4.3 Formation of Clusters

The regular repetition of frequencies results in a clustering of cells. The clusters generated

in this way can comprise the whole frequency band. In this case all of the frequencies in the available spectrum are used within a cluster. The size of a cluster is characterized by the number of cells per cluster k, which determines the frequency reuse distance D. Figure 2.15 shows some examples of clusters. The numbers designate the respective frequency sets fb_i used within the single cells.

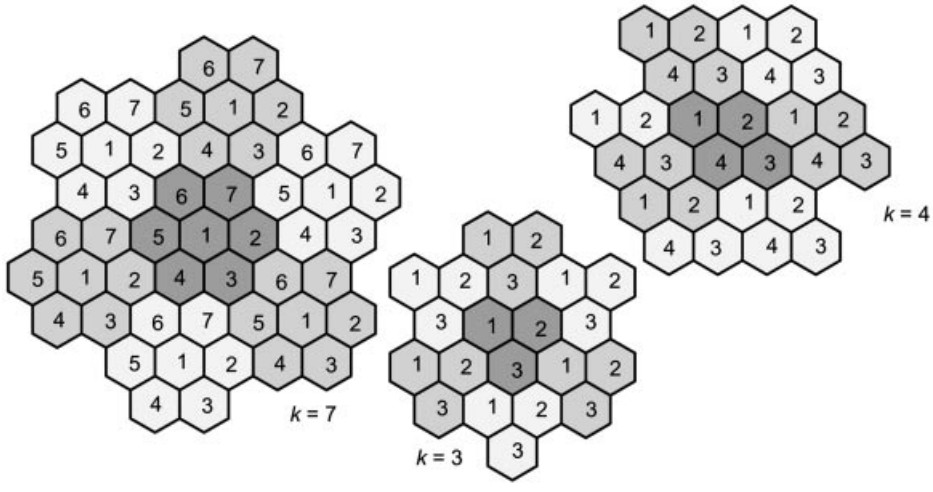

Figure 2.15: Frequency reuse and cluster formation

For each cluster the following holds:

- A cluster can contain all the frequencies of the mobile radio system.
- Within a cluster, no frequency can be reused. The frequencies of a set fb_i may be reused at the earliest in the neighboring cluster.
- The larger a cluster, the larger the frequency reuse distance and the larger the signal-to-noise ratio. However, the larger the values of k, the smaller the number of channels and the number of active subscribers per cell.

The frequency reuse distance D can be derived geometrically from the hexagon model depending on k and the cell radius R:

$$D = R\sqrt{3k}$$

The signal-to-noise ratio W [42] is then

$$W = \frac{R^{-\gamma}}{6D^{-\gamma}} = \frac{R^{-\gamma}}{6\left(R\sqrt{3k}\right)^{-\gamma}} = \frac{1}{6}(3k)^{\gamma/2}$$

According to measurements one can assume that, for good speech understandability, a carrier-to-interference ratio (CIR) of about 18 dB is sufficient. Assuming an approximate propagation coefficient of $\gamma = 4$, this yields the minimum cluster size

$$10\log W \geq 18 \text{ dB}, \qquad W \geq 63.1 \Rightarrow D \approx 4.4R$$

$$\frac{1}{6}(3k)^{\gamma/2} = W \geq 63.1 \Rightarrow k \geq 6.5 \Rightarrow k = 7$$

These values are also confirmed by computer simulations, which have shown that for $W = 18$ dB a reuse distance $D = 4.6R$ is needed [42]. In practically implemented networks, one can find other cluster sizes, e.g. $k = 3$ and $k = 12$. A CIR of 15 dB is considered a conservative value for network engineering.

The cellular models mentioned so far are very idealized for illustration and analysis. In reality, cells are neither circular nor hexagonal; rather they possess very irregular forms and sizes because of variable propagation conditions. An example of a possible cellular plan for a real network is shown in Figure 2.16, where one can easily recognize the individual cells with the assigned channels and the frequency reuse. Especially obvious are the different cell sizes, which depend on whether it is an urban, suburban, or rural area. Figure 2.16 gives an impression of the approximate contours of equal signal power around the individual base stations. In spite of this representation, the precise fitting of signal power contours remains an idealization. The cell boundaries are after all blurred and defined by local thresholds, beyond which the neighboring base station's signal is received stronger than the current one.

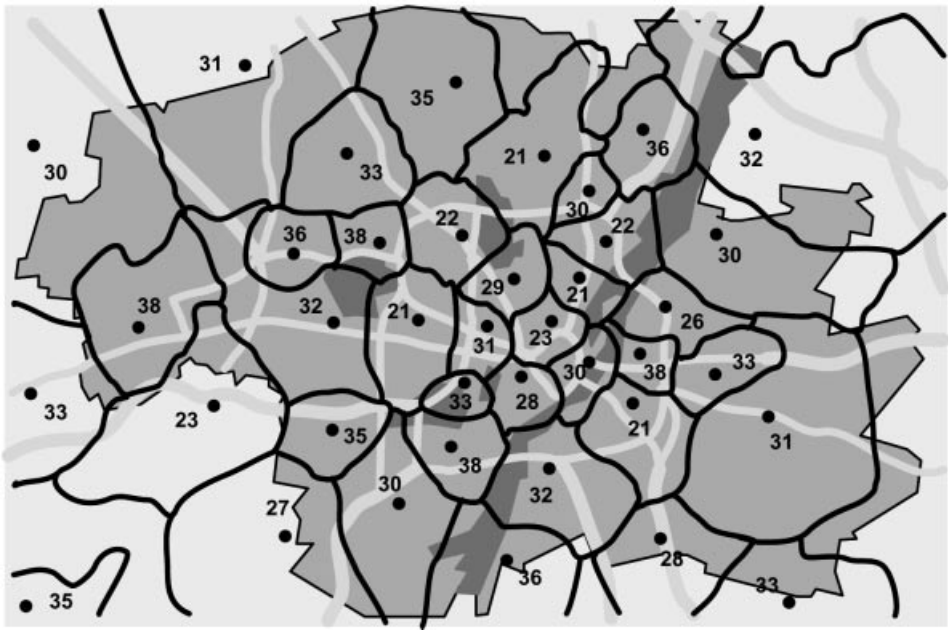

Figure 2.16: Cell structure of a real network

2.4.4 Traffic Capacity and Traffic Engineering

As already mentioned, the number of channels and thus the maximal traffic capacity per cell depends on the cluster size k. The following relation holds:

$$n_F = \frac{B_t}{B_c k}$$

where n_F is the number of frequencies per cell, B_t is the total bandwidth of the system, and B_c is the bandwidth of one channel.

The number of channels per cell in FDMA systems equals the number of frequency channels resulting from the channel and system bandwidth:

$$n = n_F$$

The number of channels per cell in a TDMA system is the number of frequency channels multiplied by the number of time slots per channel (frame size):

$$n = mn_F$$

where m is the number of time slots/frame.

A cell can be modeled as a traffic-theoretical loss system with n servers (channels), assuming a call arrival process with exponentially distributed interarrival times (Poisson process), and another Poisson process as a server process. Arrival and server processes are also called Markov processes, hence such a system is known as an M/M/n loss system [40]. For a given blocking probability B, a cell serves a maximum offered load $A_{\max}$ during the busy hour:

$$A_{\max} = f(B, n) = \lambda_{\max} T_m$$

where $\lambda_{\max}$ is the busy hour call attempts (BHCA) and T_m is the mean call holding time.

The relation between offered load A and blocking probability B with the total number of channels n is given by the Erlang blocking formula (see [40,56] for more details and traffic tables):

$$B = \frac{A^n/n!}{\sum_{i=0}^{n} A^i/i!}$$

However, these approximations are valid only for macrocellular environments, in which the number of users per cell is sufficiently large with regard to the number of available channels, such that the call arrival rate may be considered as approximately constant. For micro- and picocellular systems these assumptions usually no longer hold. Here, the traffic-theoretical dimensioning must be done with Engset models, since the number of participants does not differ very much from the number of available channels. This results in a call arrival rate that is no longer constant. The probability that all channels are busy results from the number of users M per cell and the offer a of a free source at:

$$P_n = \frac{\binom{M}{n} a^n}{\sum_{i=0}^{n} \binom{M}{i} a^i}$$

In this case, the probability that a call arrives when no free channels are available (blocking probability) is

$$P_B = \frac{\binom{M-1}{n} a^n}{\sum_{i=0}^{n} \binom{M-1}{i} a^i}$$

For $M \rightarrow \infty$, the Engset blocking formula becomes the Erlang blocking formula.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

3 System Architecture and Addressing

3.1 General Description

GSM networks are structured hierarchically (Figure 3.1). They substantially consist of at least one administrative region, which is assigned to a *Mobile Switching Center* (MSC). Each administrative region is made up of at least one *Location Area* (LA). Sometimes the LA is also called the visited area. An LA consists of several cell groups. Each cell group is assigned to a *Base Station Controller* (BSC). Therefore for each LA there exists at least one BSC, but cells of one BSC may belong to different LAs.

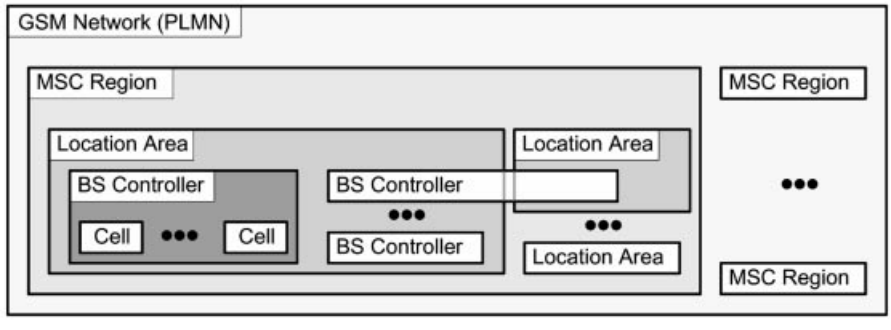

Figure 3.1: GSM system hierarchy

The exact partitioning of the service area into cells and their organization or administration with regard to LAs, BSCs, and MSCs is, however, not uniquely determined and is left to the respective network operator who thus has many possibilities for optimization. Figure 3.2 shows the system architecture of a *GSM Public Land Mobile Network* (PLMN) with essential components. The hierarchical construction of the GSM infrastructure becomes evident again. The cell is formed by the radio area coverage of a *Base Transceiver Station* (BTS). Several base stations together are controlled by one BSC. The combined traffic of the mobile stations in their respective cells is routed through a switch, the *Mobile Switching Center* (MSC). Calls originating from or terminating in the fixed network (e.g. the *Integrated Services Digital Network*, ISDN [7]) are handled by a dedicated *Gateway*

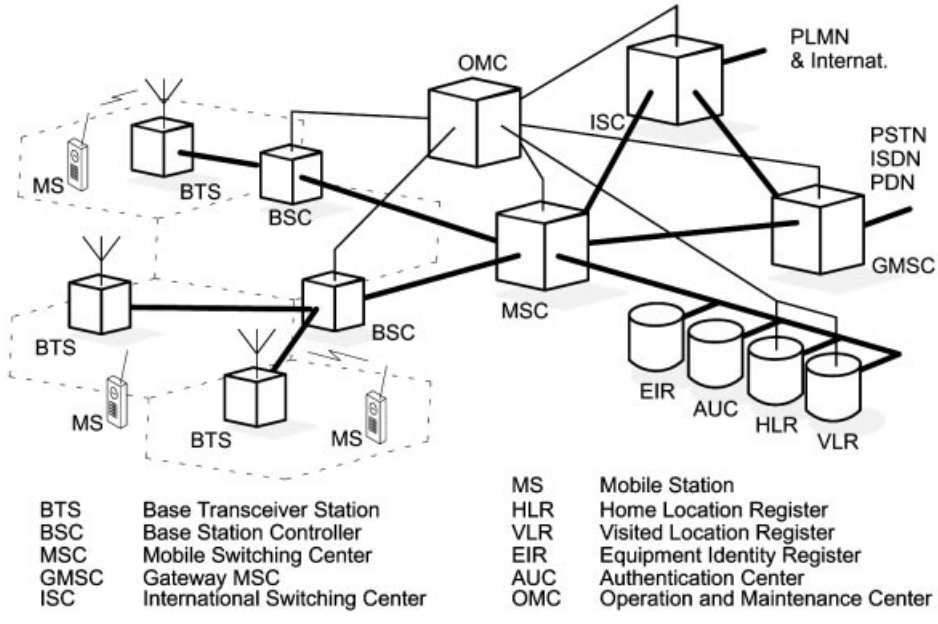

Figure 3.2: GSM system architecture with essential components

Mobile Switching Center (GMSC). Operation and maintenance are organized from a central place, the *Operation and Maintenance Center* (OMC). Several databases are available for call control and network management:

- *Home Location Register* (HLR)
- *Visited Location Register* (VLR)
- *Authentication Center* (AUC)
- *Equipment Identity Register* (EIR)

For all subscribers registered with a network operator, permanent data (such as the user's service profile) as well as temporary data (such as the user's current location) are stored in the HLR. In case of a call to a user, the HLR is always first queried, to determine the user's current location. A VLR is responsible for a group of LAs and stores the data of subscribers who are currently in its area of responsibility. This includes parts of the permanent subscriber data which have been transmitted from the HLR to the VLR for faster access. But the VLR may also assign and store local data such as a temporary identification. The AUC generates and stores security-related data such as keys used for authentication and encryption, whereas the EIR registers equipment data rather than subscriber data.

3.2 Addresses and Identifiers

GSM distinguishes explicitly between user and equipment and deals with them separately. According to this concept, which was introduced with digital mobile networks, mobile equipment and users each receive their own internationally unique identifiers. The user

identity is associated with a mobile station by means of a personal chip card, the *Subscriber Identity Module* (SIM). This SIM usually comes in the form of a chip card, which is transferable between mobile stations. It allows to distinguish between equipment mobility and subscriber mobility. The subscriber can register to the locally available network with his or her SIM card on different mobile stations, or the SIM card could be used as a normal telephone card in the fixed telephone network. However, he or she cannot receive calls on fixed network ports, but further development of the fixed networks as well as convergence of fixed and mobile networks could make this possible, too. In that case, a mobile subscriber could register at an arbitrary ISDN telephone and would be able to receive calls.

In addition, GSM distinguishes between subscriber identity and telephone number. This leaves some scope for development of future services when each subscriber may be called personally, independent of reachability or type of connection (mobile or fixed). Besides the personal identifier, each GSM subscriber is assigned one or several ISDN numbers.

Besides telephone numbers and subscriber and equipment identifiers, several other identifiers have been defined; they are needed for the management of subscriber mobility and for addressing all the remaining network elements. The most important addresses and identifiers are presented in the following.

3.2.1 International Mobile Station Equipment Identity (IMEI)

The *International Mobile Station Equipment Identity* (IMEI) uniquely identifies mobile stations internationally. It is a kind of serial number. The IMEI is allocated by the equipment manufacturer and registered by the network operator, who stores it in the *Equipment Identity Register* (EIR). By means of the IMEI one recognizes obsolete, stolen, or nonfunctional equipment and, for example, can deny service. For this purpose, the IMEI is assigned to one or more of three categories within the EIR:

- The White List is a register of all equipment.
- The Black List contains all suspended equipment. This list is periodically exchanged among network operators.
- Optionally, an operator may maintain a Gray List, in which malfunctioning equipment or equipment with obsolete software versions is registered. Such equipment has network access, but its use is reported to the operating personnel.

The IMEI is usually requested from the network at registration, but it can be requested repeatedly. It is a hierarchical address, containing of the following parts:

- *Type Approval Code* (TAC): 6 decimal places, centrally assigned
- *Final Assembly Code* (FAC): 6 decimal places, assigned by the manufacturer
- *Serial Number* (SNR): 6 decimal places, assigned by the manufacturer
- *Spare* (SP): 1 decimal place

Thus, IMEI = TAC + FAC + SNR + SP. It uniquely characterizes a mobile station and gives clues about the manufacturer and the date of manufacturing.

3.2.2 International Mobile Subscriber Identity (IMSI)

When registering for service with a mobile network operator, each subscriber receives a unique identifier, the *International Mobile Subscriber Identity* (IMSI). This IMSI is stored in the SIM; see Section 3.3.1. A mobile station can only be operated if a SIM with a valid IMSI is inserted into equipment with a valid IMEI, since this is the only way to correctly bill the associated subscriber. The IMSI also consists of several parts:

- *Mobile Country Code* (MCC): 3 decimal places, internationally standardized
- *Mobile Network Code* (MNC): 2 decimal places, for unique identification of mobile networks within a country
- *Mobile Subscriber Identification Number* (MSIN): maximum 10 decimal places, identification number of the subscriber in his/her mobile home network

The IMSI is a GSM-specific addressing concept and is different from the ISDN numbering plan. A 3-digit MCC has been assigned to each of the GSM countries, and 2-digit MNCs have been assigned within countries (e.g. 262 as MCC for Germany; and MNC 01, 02, 03, and 07 for the networks known as D1-Telekom, D2-Privat, E-Plus, and E2-Interkom, respectively). Subscriber identification therefore uses a maximum of 15 decimal digits, and IMSI = MCC + MNC + MSIN. Whereas the MCC is defined internationally, the *National Mobile Subscriber Identity* (NMSI = MNC + MSIN) is assigned by the operator of the home PLMN.

3.2.3 Mobile Subscriber ISDN Number (MSISDN)

The ''real telephone number'' of a mobile station is the *Mobile Subscriber ISDN Number* (MSISDN). It is assigned to the subscriber (his or her SIM), such that a mobile station can have several MSISDNs depending on the SIM. With this concept, GSM is the first mobile system to distinguish between subscriber identity and number to call. The separation of call number (MSISDN) and subscriber identity (IMSI) primarily serves to protect the confidentiality of the IMSI. In contrast to the MSISDN, the IMSI need not be made public. With this separation, one cannot derive the subscriber identity from the MSISDN, unless the association of IMSI and MSISDN as stored in the HLR has been made public. It is the rule that the IMSI used for subscriber identification is not known, and thus the faking of a false identity is significantly more difficult.

In addition to this, a subscriber can hold several MSISDNs for selection of different services. Each MSISDN of a subscriber is reserved for specific service (voice, data, fax, etc.). In order to realize this service, service-specific resources have to be activated in the mobile station as well as in the network. The service desired and the resources needed for the specific call can be derived from the MSISDN. Thus, an automatic activation of service-specific resources is already possible during the setup of a connection. The MSISDN categories follow the international ISDN numbering plan and therefore have the following structure:

- *Country Code* (CC): up to 3 decimal places
- *National Destination Code* (NDC): typically 2–3 decimal places
- *Subscriber Number* (SN): maximal 10 decimal places

The CCs are internationally standardized, complying to the ITU-T E.164 series [32]. There are country codes with one, two, or three digits; e.g. the country code for the USA is 1, for the UK it is 44, and for Finland it is 358. The national operator or regulatory administration assigns the NDC as well as the subscriber number SN, which may have variable length. The NDC of the mobile networks in Germany have three digits (170, 171, 172,…). The subscriber number is the concatenation MSISDN = CC + NDC + SN and thus has a maximum of 15 decimal digits. It is stored centrally in the HLR.

3.2.4 Mobile Station Roaming Number (MSRN)

The *Mobile Station Roaming Number* (MSRN) is a temporary location-dependent ISDN number. It is assigned by the locally responsible VLR to each mobile station in its area. Calls are routed to the MS by using the MSRN. On request, the MSRN is passed from the HLR to the GMSC. The MSRN has the same structure as the MSISDN:

- *Country Code* (CC) of the visited network
- *National Destination Code* (NDC) of the visited network
- *Subscriber Number* (SN) in the current mobile network

The components CC and NDC are determined by the visited network and depend on the current location. The SN is assigned by the current VLR and is unique within the mobile network. The assignment of an MSRN is done in such a way that the currently responsible switching node MSC in the visited network (CC + NDC) can be determined from the subscriber number, which allows routing decisions to be made.

The MSRN can be assigned in two ways by the VLR: either at each registration when the MS enters a new *Location Area* (LA) or each time when the HLR requests it for setting up a connection for incoming calls to the mobile station.

In the first case, the MSRN is also passed on from the VLR to the HLR, where it is stored for routing. In the case of an incoming call, the MSRN is first requested from the HLR of this mobile station. This way the currently responsible MSC can be determined, and the call can be routed to this switching node. Additional localization information can be obtained there from the responsible VLR.

In the second case, the MSRN cannot be stored in the HLR, since it is only assigned at the time of call setup. Therefore the address of the current VLR must be stored in the tables of the HLR. Once routing information is requested from the HLR, the HLR itself goes to the current VLR and uses a unique subscriber identification (IMSI and MSISDN) to request a valid roaming number MSRN. This allows further routing of the call.

3.2.5 Location Area Identity (LAI)

Each LA of a PLMN has its own identifier. The *Location Area Identifier* (LAI) is also structured hierarchically and internationally unique (Section 3.2.2), with LAI again consisting of an internationally standardized part and an operator-dependent part:

- *Country Code* (CC): 3 decimal digits
- *Mobile Network Code* (MNC): 2 decimal places

- *Location Area Code* (LAC): maximum 5 decimal places, or maximum twice 8 bits, coded in hexadecimal (LAC $< \text{FFFF}_{\text{hex}}$)

This LAI is broadcast regularly by the base station on the *Broadcast Control Channel* (BCCH). Thus, each cell is identified uniquely on the radio channel as belonging to an LA, and each MS can determine its current location through the LAI. If the LAI that is ''heard'' by the MS changes, the MS notices this LA change and requests the updating of its location information in the VLR and HLR (location update). The significance for GSM networks is that the mobile station itself rather than the network is responsible for monitoring the local conditions of signal reception, to select the base station that can be received best, and to register with the VLR of that LA which the current base station belongs to. The LAI is requested from the VLR if the connection for an incoming call has been routed to the current MSC using the MSRN. This determines the precise location of the mobile station where the mobile can be subsequently paged. When the mobile station answers, the exact cell and therefore also the base station become known; this information can then be used to switch the call through.

3.2.6 Temporary Mobile Subscriber Identity (TMSI)

The VLR being responsible for the current location of a subscriber can assign a *Temporary Mobile Subscriber Identity* (TMSI), which has only local significance in the area handled by the VLR. It is used in place of the IMSI for the definite identification and addressing of the mobile station. This way nobody can determine the identity of the subscriber by listening to the radio channel, since this TMSI is only assigned during the mobile station's presence in the area of one VLR, and can even be changed during this period (ID hopping). The mobile station stores the TMSI on the SIM card. The TMSI is stored on the network side only in the VLR and is not passed to the HLR. A TMSI may therefore be assigned in an operator-specific way; it can consist of up to 4×8 bits, but the value FFFF FFFF_{hex} is excluded, because the SIM marks empty fields internally with logical 1.

Together with the current location area, a TMSI allows a subscriber to be identified uniquely, i.e. for the ongoing communication the IMSI is replaced by the 2-tuple (TMSI, LAI).

3.2.7 Local Mobile Subscriber Identity (LMSI)

The VLR can assign an additional searching key to each mobile station within its area to accelerate database access; this is the *Local Mobile Station Identity* (LMSI). The LMSI is assigned when the mobile station registers with the VLR and is also sent to the HLR. The LMSI is not used any further by the HLR, but each time messages are sent to the VLR concerning a mobile station, the LMSI is added, so the VLR can use the short searching key for transactions concerning this MS. This kind of additional identification is only used when the MSRN is newly assigned with each call. In this case, fast processing is very important to achieve short times for call setup. Like the TMSI, an LMSI is also assigned in an operator-specific way, and it is only unique within the administrative area of a VLR. An LMSI consists of four octets (4×8 bits).

3.2.8 Cell Identifier (CI)

Within an LA, the individual cells are uniquely identified with a *Cell Identifier* (CI), maximum 2 × 8 bits. Together with the *Global Cell Identity* (LAI + CI), cells are thus also internationally defined in a unique way.

3.2.9 Base Transceiver Station Identity Code (BSIC)

In order to distinguish neighboring base stations, these receive a unique *Base Transceiver Station Identity Code* (BSIC) which consists of two components:

- *Network Color Code* (NCC): color code within a PLMN (3 bits)
- *Base Transceiver Station Color Code* (BCC): BTS color code (3 bits)

The BSIC is broadcast periodically by the base station on a Broadcast Channel, the Synchronization Channel. Directly adjacent PLMN (and BS) must have different color codes.

3.2.10. Identification of MSCs and Location Registers

MSCs and location registers (HLR, VLR) are addressed with ISDN numbers. In addition, they may have a *Signalling Point Code* (SPC) within a PLMN, which can be used to address them uniquely within the Signaling System Number 7 network (SS#7).

The number of the VLR in whose area a mobile station is currently roaming must be stored in the HLR data for this MS, if the MSRN distribution is on a call-by-call basis (Section 3.2.4); thus the MSRN can be requested for incoming calls and the call can be switched through to the MS.

3.3 System Architecture

A GSM system has two major components: the fixed installed infrastructure (the network in the proper sense) and the mobile subscribers, which use the services of the network and communicate over the radio interface (air interface). The fixed installed GSM network can again be subdivided into three subnetworks: the radio network, the mobile switching network, and the management network [21]. These subnetworks are called subsystems in the GSM standard. The respective three subsystems are the *Base Station Subsystem* (BSS), the *Switching and Management Subsystem* (SMSS), and the *Operation and Maintenance Subsystem* (OMSS).

3.3.1 Mobile Station (MS)

Mobile stations (MS) are pieces of equipment which are used by mobile service subscribers for access to services. They consist of two major components: the *Mobile Equipment* and the *Subscriber Identity Module* (SIM). Only the SIM of a subscriber turns a piece of mobile equipment into a complete mobile station with network usage privileges, which can be used

to make calls or receive calls. The SIM can be a fixed installed chip (*plug-in SIM*) or an exchangeable SIM card. In addition to the equipment identifier IMEI, the mobile station has subscriber identification and call number (IMSI and MSISDN) as subscriber-dependent data. Thus GSM mobile stations are personalized with the SIM card (Figure 3.3).

Figure 3.3: Mobile equipment personalization with the SIM

This modern concept of the SIM used consistently for the first time in GSM achieved on one hand the separation of user mobility from equipment mobility. This enables international roaming independent of mobile equipment and network technology, provided the interface between SIM and end terminal is standardized. On the other hand, the SIM can take over substantially more tasks than the personalization of mobile stations with IMSI and MSISDN. All the cryptographic algorithms to be kept confidential are realized on the SIM, which implements important functions for the authentication and user data encryption based on the subscriber identity IMSI and secret keys. Beyond that, the SIM can store short messages and charging information, and it has a telephone book function and short list of call numbers storing names and telephone numbers for efficient and fast number selection. These functions in particular contribute to a genuine personalization of a mobile terminal, since the subscriber can use his or her normal ''environment'' plus telephone list and short message archive with any piece of mobile equipment. Besides subscriber-specific data, the SIM can also store network-specific data, e.g. lists of BCCH carrier frequencies used by the network to broadcast system information periodically, or also the current LAI. Use of the SIM and thus of the whole MS can be protected with a PIN against unauthorized access.

3.3.2 Radio Network – Base Station Subsystem (BSS)

Figure 3.4 shows the components of the GSM radio network. A GSM cell is expanded around the radio area of a *Base Transceiver Station* (BTS); *trans*mitter + re*ceiver* = *transceiver*. The BTS provides the radio channels for signaling and user data traffic in this cell. Thus, a BTS is the network part of the GSM air interface. Besides the high-frequency part (transmitter and receiver equipment) it contains only a few components for signal and protocol processing. For example, error protection coding is performed in the BTS, and the link level protocol LAPDm for signaling on the radio path is terminated here. In order to keep the base stations small, the essential control and protocol intelligence

entities reside in the *Base Station Controller* (BSC). For example, the handover protocol is executed in the BSC. BTS and BSC together form the *Base Station Subsystem* (BSS). Several BTSs can be controlled together by one BSC (Figure 3.1). Each BTS is allocated a set of frequency channels, the *Cell Allocation* (CA).

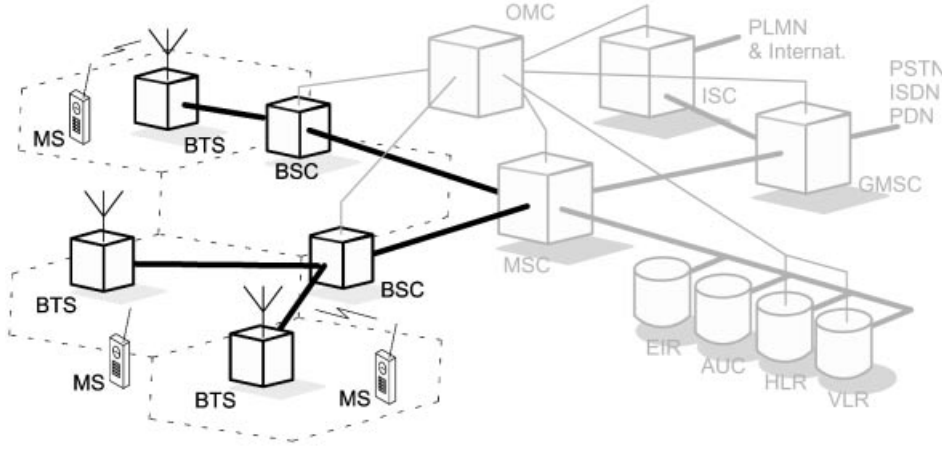

Figure 3.4: Components of the GSM radio network

Two kinds of channels are provided at the radio interface: traffic channels and signaling channels. Traffic channels are further subdivided into full-rate channels and half-rate channels. For the traffic channels, the BSS substantially comprises all the functions of OSI Layer 1.

3.3.3 Mobile Switching Network (MSS)

The *Mobile Switching and Management Subsystem* (SMSS) consists of the mobile switching centers and the databases which store the data required for routing and service provision (Figure 3.5). These components and their functions are presented briefly in the following and in more detail in later sections.

3.3.3.1 Mobile Switching Center (MSC)

The switching node of a GSM PLMN is the *Mobile Switching Center* (MSC). The MSC performs all the switching functions of a fixed-network switching node, e.g. routing path search, signal routing, and service feature processing. The main difference between an ISDN switch and an MSC is that the MSC also has to consider the allocation and administration of radio resources and the mobility of the subscribers. The MSC therefore has to provide additional functions for location registration of subscribers and for the handover of a connection in case of changing from cell to cell. A PLMN can have several MSCs with each being responsible for a part of the Service Area. The BSCs of a BSS are subordinated to a single MSC.

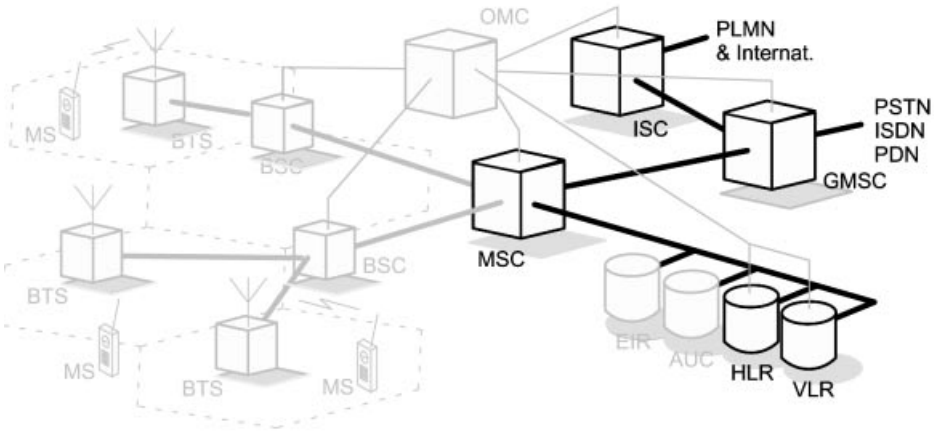

Figure 3.5: Components of the GSM mobile switching network

Dedicated *Gateway MSCs* (GMSCs) pass voice traffic between fixed networks and mobile networks. If the fixed network is unable to connect an incoming call to the local MSC (due to the inability to interrogate the HLR), it routes the connection to the next GMSC. This GMSC requests the routing information from the HLR and routes the connection to the local MSC in whose area the mobile station is currently staying. Connections to other mobile or international networks are mostly routed over the *International Switching Center* (ISC) of the respective country.

Associated with an MSC is a functional unit enabling the interworking of a PLMN and the fixed networks (PSTN, ISDN, PDN). This *Interworking Function* (IWF) performs a variety of functions depending on the service and the respective fixed network. It is needed to map the protocols of the PLMN onto those of the respective fixed network. In cases of compatible service implementation in both networks, the IWF has no functions to perform.

3.3.3.2 Home and Visitor Registers (HLR and VLR)

A GSM PLMN has several databases. Two functional units are defined for the registration of subscribers and their current location: the *Home Location Register* (HLR) and the *Visited Location Register* (VLR). In general, there is one central HLR per PLMN and one VLR for each MSC. This organization depends on the number of subscribers, the processing and storage capacity of the switches, and the structure of the network.

The HLR has entries for every subscriber and every mobile ISDN number that has his/her ''home'' in the respective network. It stores all permanent subscriber data and the relevant temporary data of all subscribers permanently registered in the HLR. Besides the fixed entries like service subscriptions and permissions, the stored data also contains a link to the current location of the mobile station (Table 3.2). The HLR is needed as the central register for routing to the subscribers, for which it has administrative responsibility. The HLR has no direct control over an MSC. All administrative activities concerning a subscriber are performed in the databases of the HLR.

The VLR as visitor register stores the data of all mobile stations which are currently staying in the administrative area of the associated MSC. A VLR can be responsible for the areas of one or more MSCs. Mobile stations are roaming freely, and therefore, depending on their current location, they may be registered in one of the VLRs of their home network or in a VLR of a ''foreign'' network (if there is a roaming agreement between both network operators). For this purpose, a mobile station has to start a registration procedure when it enters an LA. The responsible MSC passes the identity of the MS and its current LAI to the VLR, which includes these values into its database and thus registers the MS. If the mobile station has not been registered with this VLR, the HLR is informed about the current location of the MS. This process enables routing of incoming calls to this mobile station.

3.3.4 Operation and Maintenance (OMSS)

3.3.4.1 Network Monitoring and Maintenance

The ongoing network operation is controlled and maintained by the *Operation and Maintenance Subsystem* (OMSS). Network control functions are monitored and initiated from an *Operation and Maintenance Center* (OMC). Here are some of its functions:

- Administration and commercial operation (subscribers, end terminals, charging, statistics)
- Security management
- Network configuration, operation, performance management
- Maintenance tasks

Management of the network can be centralized in one or more *Network Management Centers* (NMC). The operation and maintenance functions are based on the concept of the *Telecommunication Management Network* (TMN) which is standardized in the ITU-T series M.30. The OMSS components are shown in Figure 3.6.

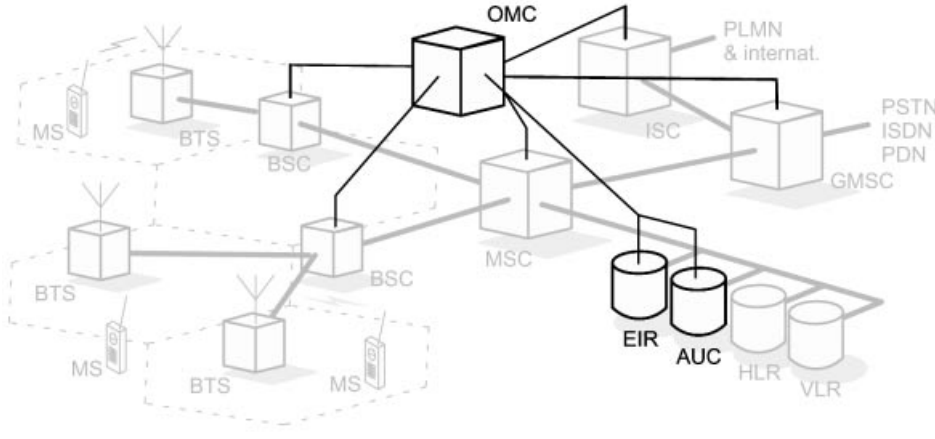

Figure 3.6: Components of the GSM OMSS

3.3.4.2 User Authentication and Equipment Registration

Two additional databases are defined in GSM besides the HLR and VLR. They are responsible for various aspects of system security. System security of GSM networks is based primarily on the verification of equipment and subscriber identity; therefore the databases serve for subscriber identification and authentication and for equipment registration. Confidential data and keys are stored or generated in the *Authentication Center* (AUC). The keys serve for user authentication and authorize the respective service access. The *Equipment Identity Register* (EIR) stores the serial numbers (supplied by the manufacturer) of the terminals (IMEI), which makes it possible to check for mobile stations with obsolete software or to block service access for mobile stations reported as stolen.

3.4 Subscriber Data in GSM

Besides data of the address type, which is the most important subscriber data of any communication network, a whole series of other service- and contract-specific data exists in GSM networks. Addresses serve to identify, authenticate, and localize subscribers, or switch connections to subscribers. Service-specific data is used to parameterize and personalize supplementary services. Finally, contracts with subscribers can define different service levels, e.g. booking of special supplementary services or subscriptions to data or teleservices. The contents of such contracts are stored in appropriate data structures in order to enable correct realization or provision of these services.

The association of the most important identifiers and their storage locations is summarized in Figure 3.7. Subscriber-related addresses are stored on the SIM and in the HLR and VLR as well. These data (IMSI, MSISDN, TMSI, MSRN) serve to address, identify, and localize a subscriber or a mobile station. Whereas IMSI and MSISDN are permanent

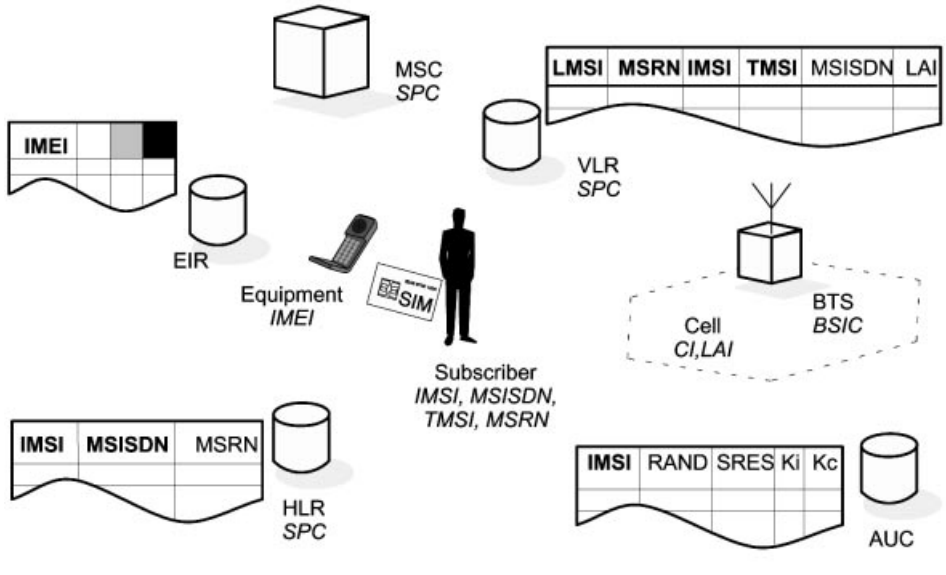

Figure 3.7: Overview of addresses and pertinent databases

data items, TMSI and MSRN are temporary values, which change according to the current location of the subscriber. Of the other data items defined for user or network equipment elements (like IMEI, LAI, or SPCs), only some are used (LAI, SPC) for localizing or routing. IMEI and BSIC/CI hold a special position by being used only for identification of network elements.

Security-relevant subscriber data is stored in the AUC, which also calculates identifiers and keys for cryptographic processing functions. Each set of data in the AUC contains the IMSI of the subscriber as a search key. For identification and authentication of a subscriber, the AUC stores the subscriber's secret key Ki from which a pair of keys RAND/SRES are precalculated and stored. Once an authentication request occurs, this pair of keys is queried by the VLR to conduct the identification/authentication process properly. The key Kc for user data encryption on the radio channel is also calculated in advance in the AUC from the secret key Ki and is requested by the VLR at connection setup.

Table 3.1: Mobile subscriber data in the HLR

Subscriber and subscription data	**Tracking and routing information**
International Mobile Subscriber Identity (IMSI)	Mobile Station Roaming Number (MSRN)
International Mobile Subscriber ISDN Number (MSISDN)	Current VLR address (if available)
Bearer and teleservice subscriptions	Current MSC address (if available)
Service restrictions, e.g. roaming restrictions	Local Mobile Subscriber Identity (LMSI) (if available)
Parameters for additional services	
Information on the subscriber's equipment (if available)	
Authentication data (subject to implementation)	

Further data about the subscriber and his or her contractual agreement with the service provider are presented in Tables 3.1 and 3.2. Above all, the HLR contains the permanent data about the subscriber's contractual relationship, e.g. information about subscribed bearer and teleservices (data, fax, etc.), service restrictions, and parameters for supplementary services. Beyond that, the registers also contain information about equipment used by the subscriber (IMEI). Depending on the implementation of the authentication center AUC and the security mechanisms, data and keys used for subscriber authentication and encryption can also be stored there.

The search keys used for retrieving subscriber information (such as IMSI, MSISDN, MSRN, TMSI and LMSI), from a register are indicated either in boldface (Figure 3.7) or in italics (Tables 3.1 and 3.2).

Table 3.2: Mobile subscriber data in the VLR

Subscriber and subscription data	Tracking and routing information
International Mobile Subscriber Identity (IMSI)	*Mobile Station Roaming Number (MSRN)*
International Mobile Subscriber ISDN Number (MSISDN)	*Temporary Mobile Station Identity (TMSI)*
Parameters for supplementary services	*Local Mobile Subscriber Identity (LMSI) (if available)*
Information on subscriber-used equipment (if available)	Local Area Identity (LAI) of LA, where MS was registered (used for paging and call setup)
Authentication data (subject to implementation)	

3.5 PLMN Configurations and Interfaces

The fixed connections for transport of signaling and user data in a GSM PLMN (Figure 3.8) are standard transmission lines. Within the SMSS, lines with a transmission rate of 2 Mbit/s (or 1.544 Mbit/s in North America) are typically used (fixed lines, mostly microwave links or leased lines). The BSS uses mostly 64 kbit/s lines. Signaling has two fundamentally different parts: GSM-specific signaling within the BSS, including the air interface, and signaling within the SMSS and with other PLMN in conformity with *Signalling System Number 7* (SS#7). User data connections are processed with an SS#7 protocol for signaling between network nodes, the *ISDN User Part* (ISUP). For the mobile

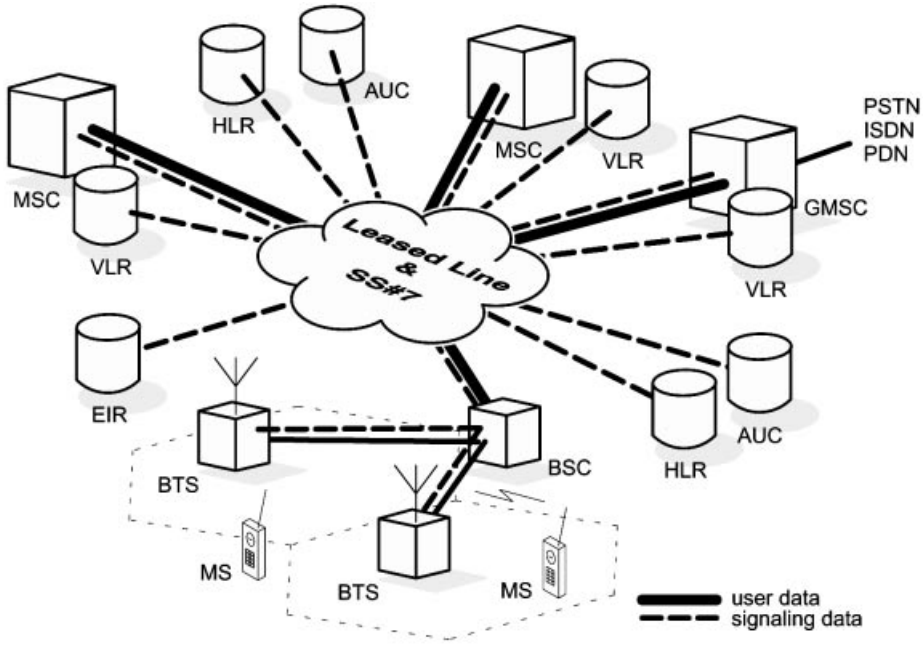

Figure 3.8: Signaling and user data transport in a GSM PLMN

network specific signaling, MSC, HLR, and VLR hold extensions of SS#7, the so-called *Mobile Application Part* (MAP). Signaling between MSC and BSS uses the *Base Station System Application Part* (BSSAP). Within the BSS and at the air interface, signaling is mobile-specific, i.e. no SS#7 protocol is used here for signaling transport.

3.5.1 Interfaces

This results in a large number of communication relationships for user data transport and signaling; for simpler structuring and standardization, these relationships have been separated by introducing a number of interfaces (Figure 3.9).

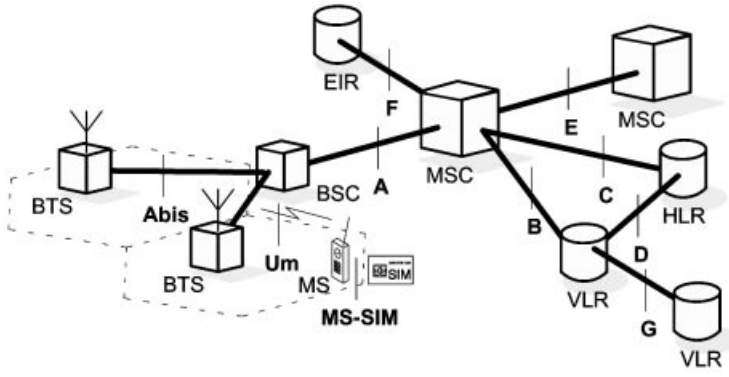

Figure 3.9: Interfaces in a GSM PLMN

The A interface between BSS and MSC is used for the transfer of data for BSS management, for connection control, and for mobility management. Within the BSS, the Abis interface between BTS and BSC and the air interface Um have been defined.

An MSC which needs to obtain data about a mobile station staying in its administrative area, requests the data from the VLR responsible for this area over the B interface. Conversely, the MSC forwards to this VLR any data generated at location updates by mobile stations. If the subscriber reconfigures special service features or activates supplementary services, the VLR is also informed first, which then updates the HLR.

This updating of the HLR occurs through the D interface. The D interface is used for the exchange of location-dependent subscriber data and for subscriber management. The VLR informs the HLR about the current location of the mobile subscriber and reports the current MSRN. The HLR transfers all the subscriber data to the VLR that is needed to give the subscriber his or her usual customized service access. The HLR is also responsible for giving a cancellation request for the subscriber data to the old VLR once the acknowledgement for the location update arrives from the new VLR. If, during location updating, the new VLR needs data from the old VLR, it is directly requested over the G interface. Furthermore, the identity of subscriber or equipment can be verified during a location update; for requesting and checking the equipment identity, the MSC has an interface F to the EIR.

An MSC has two more interfaces besides the A and B interfaces, namely the C and E interfaces. Charging information can be sent over the C interface to the HLR. Besides this, the MSC must be able to request routing information from the HLR during call setup, for calls from the mobile network as well as for calls from the fixed network. In the case of a call from the fixed network, if the fixed network's switch cannot interrogate the HLR directly, initially it routes the call to a gateway MSC (GMSC), which then interrogates the HLR. If the mobile subscriber changes during a conversation from one MSC area to another, a handover needs to be performed between these two MSCs, which occurs across the E interface.

3.5.2 Configurations

As already mentioned, the configuration of a PLMN is largely left to the network operator. Figure 3.10 shows a basic configuration of a GSM mobile communication network. This basis configuration contains a central HLR and a central VLR. All database transactions (updates, inquiries, etc.) and handover transactions between the MSC are performed with the help of the MAP over the SS#7 network. For this purpose, each MSC and register is known as a *Signalling Point* (SP) and is known by its *Signalling Point Code* (SPC) within the SS#7 network. The VLR is mainly a database which stores the location information of the mobile stations. At each change of the location area, this information must be updated. Furthermore, this database has to be interrogated: the MSC needs subscriber parameters besides location data for successful connection setup, such as service restrictions and supplementary services to be activated. Thus, there is a significant message traffic between MSC and VLR, which constitutes an ensuing load on the signaling network.

It is logical, therefore, that these two functional units are combined in one physical unit, i.e., the entire VLR is implemented in distributed form and a VLR is associated with each

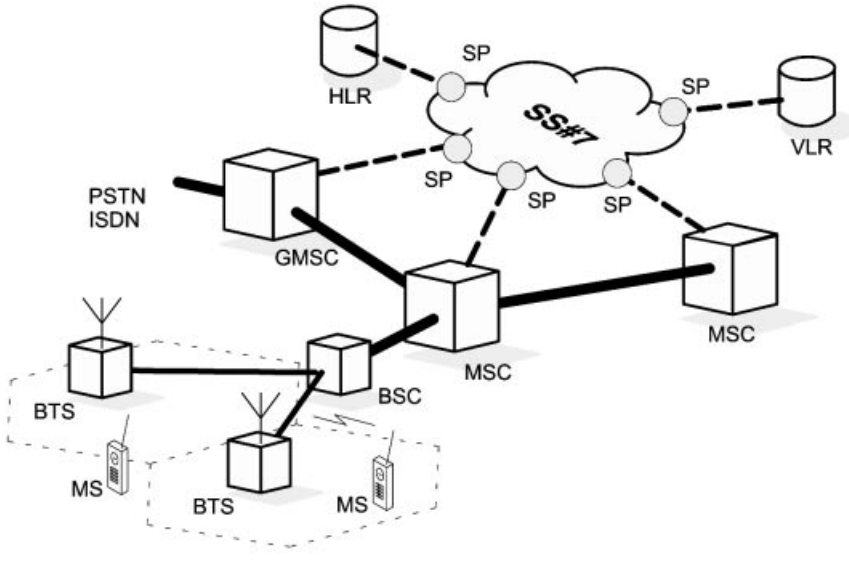

Figure 3.10: Basic configuration of a GSM PLMN

MSC (see Figure 3.11). The traffic between MSC and VLR then does not need to be transported through the SS#7 network.

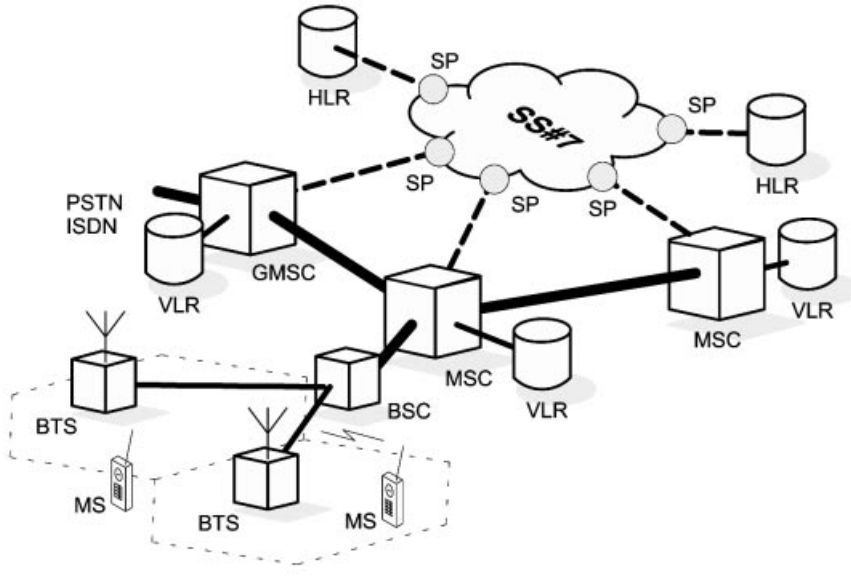

Figure 3.11: Configuration of a GSM PLMN with a VLR for each MSC

One could go one step further and also distribute the database of the HLR and thus introduce several HLRs in a mobile network. This is especially interesting for a growing pool of subscribers, since a centralized database leads to a high traffic load for this database. If there are several HLR in a PLMN, the network operator has to define an association rule between MSISDN and HLR, such that for incoming calls the routing information to an MSISDN can be derived from the associated HLR. One possible association is geographic partitioning of the whole subscriber identification space (SN field in the MSISDN, see Section 3.2.3), where, for example, the first two digits of the SN indicate the region and the associated HLR.

In extreme cases, the HLR can be realized with the VLR in a single physical unit. In this case, an HLR would also be associated with each MSC.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

4 Services

The services offered at the *User–Network Interface* (UNI) of GSM are patterned after the services offered by the *Integrated Services Digital Network* (ISDN) [7] for fixed terminals tied to telephone lines. GSM services are therefore divided just like ISDN services into three categories: *bearer services*, *teleservices*, and *supplementary services*. A bearer service offers the basic technical capability for the transmission of binary data; i.e. it offers the data transfer between end terminals at reference points R or S of the Reference Model (Figures 4.1 and 9.1). Such bearer services are made use of by the teleservices for the transfer of data with higher-level protocols.

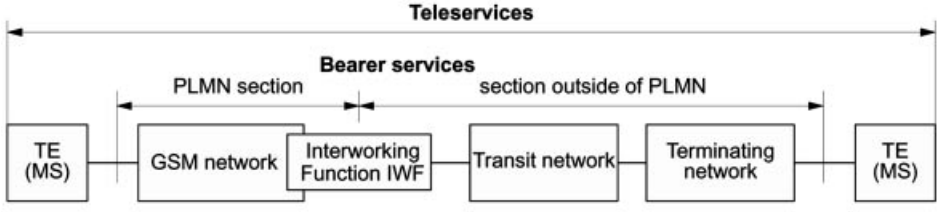

Figure 4.1: Bearer and teleservices

Notice that bearer and teleservices both require special measures not only at the air interface, but also inside the PLMN, which must offer a matching fixed-net infrastructure and special *Interworking Functions* (IWFs). Especially at the level of bearer services, the IWF must provide a mapping of GSM PLMN services within their respective service characteristics onto corresponding bearer services and characteristics of the other networks, such as PSTN and ISDN. Teleservices are end-to-end services, for which there is usually no translation in the IWF. But they do use bearer services, which again need IWF functions.

Bearer and teleservices are carried under the umbrella term *telecommunication services*. The simultaneous use of two telecommunication services is precluded, except for the case of *Short Message Services* (SMS), which can at least be received during the use of another telecommunication service. Supplementary services are complementing the control and modification of extended services and are only usable in connection with a telecommunication service.

For telecommunication services, the GSM standard includes agreements of target times for their market introduction. This fact is especially important, since GSM is an international standard which aims at worldwide compatibility of mobile stations and networks. Accordingly, only a minimum of services has been defined, which must be offered by the operators at various time phases. For this purpose, the services are divided into the categories *essential* (E) and *additional* (A). Group E must be implemented at the given date by all network operators, whereas the decision about the time of introduction of Group A services is left to the operators. Table 4.1 gives a rough overview over the implementation and introduction phases. The most important services to be implemented and introduced with their respective introduction dates are described briefly in the following.

Table 4.1: Phases of implementation and introduction

Class	Introduction	Services
E1	1991	Basic operation consisting of telephone services and some appropriate supplementary services
E2	1994	Extended operation with telephone services, first non-speech services (e.g. BS26) and an extended range of supplementary services
E3	1996	Enhanced service range with even more telecommunication and supplementary services

4.1 Bearer Services

The basic services of a GSM network are the foundation for data transmission, i.e. a basic service provides the fundamental technical facilities at the end terminal interface (reference point R) to transport user payloads. The basic services are called transport services [7] or bearer services in ISDN – and therefore in GSM, too. The GSM bearer services offer asynchronous or synchronous data transport capabilities with circuit-switched or packet-switched data rates of 300–9600 bit/s, with a 13 kbit/s bearer service for voice.

Bearer services carry only the coding- and application-independent information transport between the user–network interfaces (Figure 4.1); they represent the services of Layers 1, 2, and 3 of the OSI Reference Model. The operators of the *Terminal Equipment* (TE) can use these services and employ arbitrary higher-level protocols, but they are responsible for the compatibility of the protocols used in the terminal equipment, quite in contrast to teleservices, where the protocols in the terminal equipment are also standardized [7]. An overview of the most important bearer services is given in Table 4.2. Each bearer service has its own number, e.g. BS26 is the bearer service for circuit-switched asynchronous data transfer at 9600 bit/s.

Besides the asynchronous and synchronous circuit-switched data services (BS21–BS34), packet-switched data services are also provided. These packet services are realized either as asynchronous access to a *Packet Assembler/Disassembler* (PAD), BS41–BS46, or as direct synchronous *Packet Access*, BS51–BS53.

The bearer services for GSM data transfer are offered in two fundamentally different modes (Table 4.2): transparent (T) and nontransparent (NT). In the transparent mode,

there is a circuit-switched connection between the mobile terminal (TE) and the interworking module in the MSC, from where the connection to other networks is handled. This connection is protected by *Forward Error Correction* (FEC). The most important common characteristics for all transparent services are constant bit rate, constant transport delay, and residual bit error ratio dependent on the current channel conditions. The nontransparent mode activates a special Layer 2 protocol for the additional protection of the data transfer, the *Radio Link Protocol* (RLP) which is specially adapted to the GSM radio channel. This protocol terminates in the mobile station and in the MSC. It uses ARQ procedures to request retransmission of blocks with residual errors which could not be corrected by forward error correction.

This gives a much more significant reduction in the residual error rate. In essence an error-free information transport is achieved, hence it is approximately independent of the

Table 4.2: GSM bearer services (excerpt)[a]

Service	Structure	BS no.	Bit rates (in bit/s)	Mode	Transmission
Data	Asynch	21	300	T or NT	UDI or 3.1 kHz
		22	1200	T or NT	UDI or 3.1 kHz
		23	1200/75	T or NT	UDI or 3.1 kHz
		24	2400	T or NT	UDI or 3.1 kHz
		25	4800	T or NT	UDI or 3.1 kHz
		26	9600	T or NT	UDI or 3.1 kHz
Data	Synch	31	1200	T	UDI or 3.1 kHz
		32	2400	T or NT	UDI or 3.1 kHz
		33	4800	T or NT	UDI or 3.1 kHz
		34	9600	T or NT	UDI or 3.1 kHz
PAD	Asynch	41	300	T or NT	UDI
		42	1200	T or NT	UDI
		43	1200/75	T or NT	UDI
		44	2400	T or NT	UDI
		45	4800	T or NT	UDI
		46	9600	T or NT	UDI
Packet	Synch	51	2400	NT	UDI
		52	4800	NT	UDI
		53	9600	NT	UDI
Altern. speech/data		61	13000 or 9600		
Speech followed by data		81	13000 or 9600		

[a] T/NT, transparent/non-transparent; UDI, unrestricted digital information; PAD, packet assembler/disassembler; asynch/synch, asynchronous/synchronous

momentary channel conditions. However, with changing error behavior of the radio channel, the frequency of block repetitions also varies, and thus the average transfer delay and the net bit rate of the data service vary too. Activation of the nontransparent data service is especially interesting for rapidly moving mobile stations or for cases of bad radio conditions, where high fading rates and deep fading/shadow holes occur. In such situations, a meaningful transport of user data with transparent mode can become impossible. At the expense of a net data rate decrease, the nontransparent mode then still allows a reliable data transport.

The GSM bearer services 21–53 are further categorized into *Unrestricted Digital Information* (UDI) and 3.1 kHz (Table 4.2). The services differ mainly in the way in which they are handled outside of the PLMN, i.e. the kind of interworking function that needs to be activated. The UDI service category corresponds to the UDI of ISDN and supplies a channel for the unrestricted transfer of digital information. The data transfer is unrestricted in the sense that no bit patterns are reserved or explicitly excluded from transmission. The 3.1 kHz category is used to activate in the MSC an interworking function for 3.1 kHz audio and to select a modem. Within the GSM PLMN (from user-network access to the interworking function), the data are still transferred as UDI. The designation ''3.1 kHz'' rather refers to the fact that the transfer outside of the PLMN uses a service ''3.1 kHz Audio.'' This service is offered by conventional PSTN as well as by ISDN networks. For transfer with this service, the data has to be converted in the IWF of the MSC with a modem to an audio signal with a bandwidth of 3.1 kHz.

Further important GSM bearer services contain voice (telephone) service (BS61 to BS81), which can be (multiple times) changed during a call at the request of the user to a data service (alternate speech/data). Another alternative is that the user at first establishes a voice connection and then changes to a data connection, which cannot be changed back to voice (speech followed by data).

4.2 Teleservices

On top of the bearer services, which can be used by themselves, a number of teleservices have been defined. The most important categories are (Table 4.3) speech, SMS, access to *Message Handling Systems* (MHSs) and to videotext, teletext, and facsimile transfer.

4.2.1 Voice

Voice services had to be implemented by each operator in the start-up phase (E1) by 1991. In this category, two teleservices were distinguished: regular telephone service (TS11) and emergency service (TS12). For transmission of the digitally coded speech signals, both services use a bidirectional, symmetric, full-duplex point-to-point connection, which is set up on user demand. The sole difference between TS11 and TS12 teleservices is that regular service requires an international IWF, whereas the emergency service stays within the boundaries of a national network.

4.2.2 Fax Transmission

As teleservice for the second implementation phase (E2), implementation of transparent fax service (TS61) for Group 3 fax was planned. The fax service is called *transparent* because it uses a transparent bearer service for the transmission of fax data. The coding and transmission of the facsimile data uses the fax protocol according to the ITU-T recommendation T30. The network operator also has the option to implement TS61 on a nontransparent bearer service in order to improve the transmission quality. TS61 is transmitted over a traffic channel that is alternately used for voice or fax. Another optional alternative is designated as *fax transfer with automatic call acceptance* (TS61). This service can be offered by a network operator when *multinumbering* is used as the interworking solution. In the case of multinumbering, a subscriber is assigned several MSISDN numbers, and a separate interworking profile is stored for each of them. In this way a specific teleservice can be associated with each MSISDN, the fax service being one of them. If a mobile subscriber is called on his or her ''GSM-fax number,'' the required resources in the IWF of the MSC as well as in the MS can be activated; whereas in the case of TS61, fax calls arrive with the same number as voice calls (no multinumbering) and have to be switched over to fax reception manually.

Table 4.3: GSM teleservices (excerpt)

Category	TS no.	Service		Class
Speech	11	Telephone		E1
	12	Emergency call		E1
Fax transmission	61	Speech and fax group 3	T	E2
		alternating	NT	A
	62	Fax group 3 automatic	T	–
			NT	–
Short Message Services (SMS)	21	Short message mobile terminated, point to point		E3
	22	Short message mobile originated, point to point		A
	23	Short message cell broadcast		–
MHS access	31	Access to message handling systems		A
Videotex access	41	Videotex access profile 1		A
	42	Videotex access profile 2		A
	43	Videotex access profile 3		A
Teletext transmission	51	Teletext		A

4.2.3 Short Message Service (SMS)

Another teleservice which was assigned high priority in the service implementation strategy – and which is now very successful – is the capability to receive or send short messages at the mobile station: *Short Message Service* (SMS), TS21 and TS22. This service was supposed to be offered in the third phase (E3) at the latest from 1996 on all GSM networks. TS21 is the point-to-point version of the SMS, which allows a single station to be sent a message of up to 160 characters. Conversely, TS22 has been defined as an optional implementation of the capability to send short messages from a mobile station. The combinations of SMS with other added-value services, e.g. mailbox systems with automatic notification of newly arrived messages or the transmission by short message of incurred charges, clearly show how the services offered by GSM networks go significantly beyond the services offered in fixed networks.

For SMS, the network operator has to establish a service center which accepts short messages from the fixed network and processes them in a store-and-forward mode. The interface has not been specified and can be by DTMF signaling, special order, email, fax, etc. The delivery can be time-shifted and is of course independent of the current location of the mobile station. Conversely, a service center can accept short messages from mobile stations which can also be forwarded to subscribers in the fixed network, for example by fax or email. The transmission of short messages uses a connectionless, protected, packet-switching protocol. The reception of a message must be acknowledged by the mobile station or the service center; in case of failure, retransmission occurs.

TS21 and TS22 are the only teleservices which can be used simultaneously with other services, i.e. short messages can also be received or transmitted during an ongoing call.

A further variation of the SMS is the *Cell Broadcast Service* TS23, *Short Message Service Cell Broadcast* (SMSCB). SMSCB messages are broadcast only in a limited region of the network. They can only be received by mobile stations in idle mode, and reception is not acknowledged. A mobile station itself can not send SMSCB messages. With this service, messages contain a category designation, so that mobile stations can select categories of interest which they want to receive and store. The maximum length of SMSCB messages is 93 characters, but by using a special reassembly mechanism, the network can transmit longer messages of up to 15 subsequent SMSCB messages.

4.3 Supplementary Services

The supplementary services in GSM correspond to the supplementary services of ISDN with regard to service and performance characteristics. They can be used only in connection with a teleservice, i.e. they modify or supplement the functionality of a GSM telecommunication service (bearer or teleservice). Besides the improved network organization, the introduction of numerous ISDN-like supplementary services is the main feature of GSM Phase 2. Some GSM supplementary services are identical or similar to those offered in ISDN, but their implementation is often much more complex due to the added mobility. Beyond that, GSM offers new service characteristics which are available in ISDN networks only in restricted form or not at all.

4.3.1 Supplementary Services of Phase 1

For Phase 1 of GSM, only a small set of supplementary services concerning call forwarding and call restriction was defined (Table 4.4). If a mobile station activates call forwarding, then calls are not switched through to this MS, but forwarded to a configurable extension. Several variations can be distinguished: first, unconditional call forwarding (CFU) where all calls are diverted; then conditional call forwarding when calls are only forwarded under special conditions, such as when the MS is busy (CFB) or is not reachable (CFNRc), possibly because it is powered off or outside any covered network area.

The network operators usually offer a voice mailbox service in connection with call forwarding. This consists of an answering machine function within the network, which offers recording of voice messages for later retrieval by the subscriber for incoming calls, if the call forwarding feature has been activated. This kind of service offering clearly goes beyond what fixed ISDN networks are offering. Of course, call forwarding can also be directed at another target than the voice mailbox.

GSM Phase 1 also introduced supplementary services for barring of either outgoing or incoming calls. In this case there are also several variants. For example, all calls can be barred outgoing (BAOC) or barred incoming (BAIC), or it may be only outgoing international calls which are barred (BOIC), or perhaps incoming calls that might cause charges such as calls to an MS which is roaming outside its home network (BIC-Roam).

Table 4.4: Overview of GSM supplementary services (GSM Phase 1)

Category	Abbreviation	Service	Class
Call offering	CFU	Call forwarding unconditional	E1
	CFB	Call forwarding on mobile subscriber busy	E1
	CFNRy	Call forwarding on no reply	E1
	CFNRc	Call forwarding on mobile subscriber not reachable	E1
Call restriction	BAOC	Barring of all outgoing calls	E1
	BOIC	Barring of outgoing international calls	E1
	BAIC	Barring of all incoming calls	E1
	BOIC-exHC	Barring of outgoing international calls except calls to home PLMN	A
	BIC-roam	Barring of incoming calls when roaming outside the home PLMN	A

4.3.2 Supplementary Services of Phase 2

In the course of further evolution of the GSM standard, the menu of services known from ISDN is being made available in stages [7] and supplemented by some new GSM-specific performance characteristics. In Phase 2, which was standardized in 1996, there are some

supplementary services (Table 4.5), such as *Call Waiting* (CW) or hold (HOLD), which enable performing brokerage functions.

Two very powerful supplementary services are *Conference Calling* (CONF) allowing the interconnection of several subscribers in one call, and *Call Transfer* (CT) which allows a call to be passed to a third party. Of special interest in connection with call waiting and call transfer services are the supplementary services of the number identification category (Table 4.5). The *Calling Line Identification Presentation* (CLIP) lets the calling party's MSISDN number appear on the display of the called party, but the calling party can prevent this by activating the supplementary service *Calling Line Identification Restriction* (CLIR), in case the caller does not want to disclose his or her number. The eventually reached number may not always be the number called by the calling party, e.g. in the case of a call transfer. With the supplementary service *Connected Line Identification Presentation* (COLP) the caller can request to be shown the reached extension, but the called party can prevent this announcement by using the *Connected Line Identification Restriction* (COLR). The inquiry of current charges is also offered with a supplementary service, as well as *Reverse Charging* (REVC), which allows the called party to assume the charges for

Table 4.5: Overview of GSM supplementary services (GSM Phase 2)

Category	Abbreviation	Service	Class
Number identification	CLIP	Calling line identification presentation	A
	CLIR	Calling line identification restriction	A
	COLP	Connected line identification presentation	A
	COLR	Connected line identification restriction	A
	MCI	Malicious call identification	A
Call offering	CT	Call transfer	A
	MAH	Mobile access hunting	A
Community of interest	CUG	Closed user group	A
Charging	AoC	Advice of charge	E2
	FPH	Freephone service	A
	REVC	Reverse charging	A
Additional information transfer	UUS	User-to-user signaling	A
Call completion	CW	Call waiting	E3
	HOLD	Call hold	E2
	CCBS	Completion of call to busy subscriber	A
Multi-party	3PTY	Three-party service	E2
	CONF	Conference calling	E3

the call. These features are clearly responsible for providing a lot more calling comfort in GSM networks than ISDN networks are offering, even though digital technology enables all of them in both.

4.4 GSM Services of Phase 2+

The standardization and further development of GSM systems, however, is not completed with Phase 2 and continues to proceed. This process is generally known under the name *GSM Phase 2+*.

A broad number of topics are considered as independent standardization units. To a large extent, their implementation can be carried out independently from each other. The topics affect almost all aspects of GSM. For example, new bearer services with higher bit rates have been developed. The *General Packet Radio Service* (GPRS) has been standardized for connectionless packet switched data communication over the radio channel. GPRS is interesting for diverse applications where the mobile data communication is typically characterized by bursty traffic. These connections do not require a complete traffic channel for the entire duration. In particular, mobile Internet access with GPRS is a typical application scenario. Furthermore, new GSM speech services have been standardized in Phase 2+. Chapter 12 of this book presents some of the services of Phase 2+, and Chapter 11 deals with GPRS in detail.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

5 Air Interface – Physical Layer

The GSM physical layer, which resides on the first of the seven layers of the OSI Reference Model [55], contains very complex functions. The *physical channels* are defined here by a TDMA multiple access scheme. On top of the physical channels, a series of *logical channels* are defined, which are transmitted in the time slots of the physical channels. Logical channels perform a multiplicity of functions, such as payload transport, signaling, broadcast of general system information, synchronization, and channel assignment.

The structure of this chapter is as follows: In Section 5.1, we describe the logical channels. This serves as a foundation for understanding the signaling procedures at the air interface. The realization of the physical channels, including GSM modulation, multiple access, duplexing, and frequency hopping follows in Section 5.2. Next, Section 5.3 covers synchronization. The mapping of logical onto physical channels follows in Section 5.4, where the higher-level multiplexing of logical channels into multiframes is also covered. Section 5.5 contains a discussion of the most important control mechanisms for the air interface (channel measurement, power control, disconnection, and cell selection). The conclusion of the chapter is a power-up scenario with the sequence of events occurring, from when a mobile station is turned on to when it is in a synchronized state ready to transmit (Section 5.6).

5.1 Logical Channels

On Layer 1 of the OSI Reference Model, GSM defines a series of logical channels, which are made available either in an unassigned random access mode or in a dedicated mode assigned to a specific user. Logical channels are divided into two categories (Table 5.1): Traffic channels and signaling (control) channels.

5.1.1 Traffic Channels

The *Traffic Channels* (TCHs) are used for the transmission of user payload data (speech, fax, data). They do not carry any control information of Layer 3. Communication over a TCH can be circuit-switched or packet-switched. In the circuit-switched case, the TCH provides a transparent data connection or a connection that is specially treated according to

the carried service (e.g. telephony). For the packet-switched mode, the TCH carries user data of OSI Layers 2 and 3 according to the recommendations of the X.25 standard or similar standard packet protocols.

A TCH may either be fully used (full-rate TCH, TCH/F) or be split into two half-rate channels (half-rate TCH, TCH/H), which can be allocated to different subscribers. Following ISDN terminology, the GSM traffic channels are also designated as Bm channel (mobile B channel) or Lm channel (lower-rate mobile channel, with half the bit rate). A Bm channel is a TCH for the transmission of bit streams of either 13 kbit/s of digitally coded speech or of data streams at 14.5, 12, 6, or 3.6 kbit/s. Lm channels are TCH channels with less transmission bandwidth than Bm channels and transport speech signals of half the bit rate (TCH/H) or bit streams for data services with 6 or 3.6 kbit/s.

Table 5.1: Classification of logical channels in GSM

Group		Channel	Function	Direction
Traffic channel	Traffic channel (TCH)	TCH/F, Bm	Full rate TCH	MS ↔ BSS
		TCH/H, Lm	Half rate TCH	MS ↔ BSS
Signaling	Broadcast channel	BCCH	Broadcast control	MS ← BSS
channels (Dm)		FCCH	Frequency correction	MS ← BSS
		SCH	Synchronization	MS ← BSS
	Common control channel (CCCH)	RACH	Random access	MS → BSS
		AGCH	Access grant	MS ← BSS
		PCH	Paging	MS ← BSS
		NCH	Notification	MS ← BSS
	Dedicated control channel (DCCH)	SDCCH	Stand-alone dedicated control	MS ↔ BSS
		SACCH	Slow associated control	MS ↔ BSS
		FACCH	Fast associated control	MS ↔ BSS

5.1.2 Signalling Channels

The control and management of a cellular network demands a very high signaling effort. Even when there is no active connection, signaling information (for example location update information) is permanently transmitted over the air interface. The GSM signaling channels offer a continuous, packet-oriented signaling service to MSs in order to enable them to send and receive messages at any time over the air interface to the BTS. Following ISDN terminology, the GSM signaling channels are also called Dm channels (mobile D channel). They are further divided into: *Broadcast Channel* (BCH), *Common Control Channel* (CCCH), and *Dedicated Control Channel* (DCCH) (see Table 5.1).

The unidirectional Broadcast Channels are used by the *Base Station Subsystem* (BSS) to

broadcast the same information to all MSs in a cell. The group of Broadcast Channels consists of three channels:

- *Broadcast Control Channel* (BCCH): On this channel, a series of information elements is broadcast to the MSs which characterize the organization of the radio network, such as radio channel configurations (of the currently used cell as well as of the neighboring cells), synchronization information (frequencies as well as frame numbering), and registration identifiers (LAI, CI, BSIC). In particular, this includes information about the structural organization (formats) of the CCCH of the local BTS. The BCCH is broadcast on the first frequency assigned to the cell (the so-called *BCCH carrier*).
- *Frequency Correction Channel* (FCCH): On the FCCH, information about correction of the transmission frequency is broadcast to the MSs; see Section 5.2.2 (frequency correction burst).
- *Synchronization Channel* (SCH): The SCH broadcasts information to identify a BTS, i.e. *Base Station Identity Code* (BSIC); see Section 3.2.9. The SCH also broadcasts data for the frame synchronization of an MS, i.e. *Reduced Frame Number* (RFN) of the TDMA frame; see Section 5.3.1.

FCCH and SCH are only visible within protocol Layer 1, since they are only needed for the operation of the radio subsystem. There is no access to them from Layer 2. In spite of this fact, the SCH messages contain data which are needed by Layer 3 for the administration of radio resources. These two channels are always broadcast together with the BCCH.

The CCCH is a point-to-multipoint signaling channel to deal with access management functions. This includes the assignment of dedicated channels and paging to localize a mobile station. It comprises the following:

- *Random Access Channel* (RACH): The RACH is the uplink portion of the CCCH. It is accessed from the mobile stations in a cell without reservation in a competitive multiple-access mode using the principle of slotted Aloha [4], to ask for a dedicated signaling channel (SDCCH) for exclusive use by one MS for one signaling transaction.
- *Access Grant Channel* (AGCH): The AGCH is the downlink part of the CCCH. It is used to assign an SDCCH or a TCH to a mobile station.
- *Paging Channel* (PCH): The PCH is also part of the downlink of the CCCH. It is used for paging to find specific mobile stations.
- *Notification Channel* (NCH): The NCH is used to inform mobile stations about incoming group and broadcast calls.

The last type of signaling channel, the DCCH is a bidirectional point-to-point signaling channel. An *Associated Control Channel* (ACCH) is also a dedicated control channel, but it is assigned only in connection with a TCH or an SDCCH. The group of *Dedicated/Associated Control Channels* (D/ACCH) comprises the following:

- *Stand-alone Dedicated Control Channel* (SDCCH): The SDCCH is a dedicated point-to-point signaling channel (DCCH) which is not tied to the existence of a TCH (''stand-alone''), i.e. it is used for signaling between an MS and the BSS when there is no active connection. The SDCCH is requested from the MS via the RACH and assigned via the AGCH. After the completion of the signaling transaction, the SDCCH is released and can be reassigned to another MS. Examples of signaling transactions

which use an SDCCH are the updating of location information or parts of the connection setup until the connection is switched through (see Figure 5.1).

- *Slow Associated Control Channel* (SACCH): An SACCH is always assigned and used with a TCH or an SDCCH. The SACCH carries information for the optimal radio operation, e.g. commands for synchronization and transmitter power control and reports on channel measurements (Section 5.5). Data must be transmitted continuously over the SACCH since the arrival of SACCH packets is taken as proof of the existence of the physical radio connection (Section 5.5.3). When there is no signaling data to transmit, the MS sends a measurement report with the current results of the continuously conducted radio signal level measurements (Section 5.5.1).
- *Fast Associated Control Channel* (FACCH): By using dynamic pre-emptive multiplexing on a TCH, additional bandwidth can be made available for signaling. The signaling channel created this way is called FACCH. It is only assigned in connection with a TCH, and its short-time usage goes at the expense of the user data transport.

In addition to these channels, a *Cell Broadcast Channel* (CBCH) is defined, which is used to broadcast the messages of the *Short Message Service Cell Broadcast* (SMSCB). The CBCH shares a physical channel together with the SDCCH.

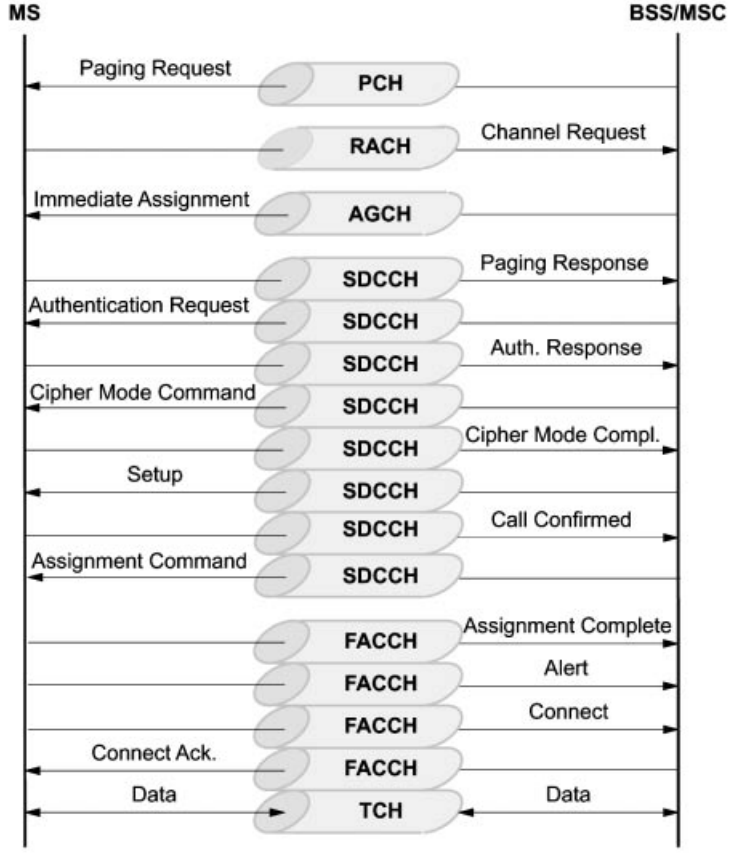

Figure 5.1: Logical channels and signaling (connection setup for an incoming call)

5.1.3 Example: Connection Setup for Incoming Call

Figure 5.1 shows an example for an incoming call connection setup at the air interface. It is illustrated how the various logical channels are used in principle. The mobile station is called via the PCH and requests a signaling channel on the RACH. It gets the SDCCH through an IMMEDIATE ASSIGNMENT message on the AGCH. Then follow authentication, start of ciphering, and start of setup over the SDCCH. An ASSIGNMENT COMMAND message gives the traffic channel to the mobile station, which acknowledges its receipt on the FACCH of this traffic channel. The FACCH is also used to continue the connection setup.

5.1.4 Bit Rates, Block Lengths, and Block Distances

Table 5.2 gives an overview of the logical channels of Layer 1, the available bit rates, block lengths used, and the intervals between transmission of blocks. The 14.4 kbit/s data service has been standardized in further GSM standardization phases. Notice that the logical channels can suffer from substantial transmission delays depending on the respective use of forward error correction (channel coding and interleaving, see Section 6.2 and Table 6.8).

Table 5.2: Logical channels of GSM Protocol Layer 1

Channel type	Net data throughput (in kbit/s)	Block length (in bit)	Block distance (in ms)
TCH (full-rate speech)	13.0	182 + 78	20
TCH (half-rate speech)	5.6	95 + 17	20
TCH (data, 14.4 kbit/s)	14.5	290	20
TCH (data, 9.6 kbit/s)	12.0	60	5
TCH (data, 4.8 kbit/s)	6.0	60	10
TCH (data, $\leq$ 2.4 kbit/s)	3.6	72	10
FACCH full rate	9.2	184	20
FACCH half rate	4.6	184	40
SDCCH	598/765	184	3060/13
SACCH (with TCH)	115/300	168 + 16	480
SACCH (with SDCCH)	299/765	168 + 16	6120/13
BCCH	598/765	184	3060/13
AGCH	n × 598/765	184	3060/13
NCH	m × 598/765	184	3060/13
PCH	p × 598/765	184	3060/13
RACH	r × 27/765	8	3060/13
CBCH	598/765	184	3060/13

5.1.5 Combinations of Logical Channels

Not all logical channels can be used simultaneously at the radio interface. They can only be deployed in certain combinations and on certain physical channels. GSM has defined several channel configurations, which are realized and offered by the base stations (Table 5.3). As already mentioned before, an SACCH is always allocated either with a TCH or with an SDCCH, which accounts for the attribute ''associated''.

Depending on its current state, a mobile station can only use a subset of the logical channels offered by the base station. It uses the channels only in the combinations indicated in Table 5.4. The combination M1 is used in the phase when no physical connection exists, i.e. immediately after the power-up of the mobile station or after a disruption due to unsatisfactory radio signal conditions. Channel combinations M2 and M3 are used by active mobile stations in standby mode. In phases requiring a dedicated signaling channel, a mobile station uses the combination M4, whereas M5 to M8 are used when there is a traffic channel up. M8 is a multislot combination (an MS transmits on several physical

Table 5.3: Channel combinations offered by the base station

	B1	B2	B3	B4	B5	B6	B7	B8	B9
TCH/F	■							■	■
TCH/H		■	■						
TCH/H			■						
BCCH				■	■	■			
FCCH				■	■				
SCH				■	■				
CCCH				■	■	■			
SDCCH					■		■		
SACCH	■	■	■		■		■	■	■
FACCH	■	■	■					■	

Table 5.4: Channel combinations used by the base station

	M1	M2	M3	M4	M5	M6	M7	M8
TCH/F					■			*n+m*
TCH/H						■	■	
TCH/H							■	
BCCH	■		■					
CCCH		■	■					
SDCCH				■				
SACCH				■	■	■	■	*n+m*
FACCH					■	■	■	■

channels), where n denotes the number of bidirectional channels, and m denotes the number of unidirectional channels ($n = 1, \ldots, 8$, $m = 0, \ldots, 7$, $n + m = 1, \ldots, 8$).

5.2 Physical Channels

After discussing the logical channels and their tasks, we now deal with the physical channels, which transport the logical channels via the air interface. We first describe the GSM modulation technique (Section 5.2.1), followed by the multiplexing structure (Section 5.2.2): GSM is a multicarrier TDMA system, i.e. it employees a combination of FDMA and TDMA for multiple access. This section also covers the explanation of the radio bursts. Finally, Section 5.2.3 briefly describes the (optional) frequency hopping technique, which has been standardized to reduce interference.

5.2.1 Modulation

The modulation technique used on the radio channel is *Gaussian Minimum Shift Keying* (GMSK). GMSK belongs to a family of continuous-phase modulation procedures, which have the special advantages of a narrow transmitter power spectrum with low adjacent channel interference on the one hand and a constant amplitude envelope on the other hand, which allows use of simple amplifiers in the transmitters without special linearity requirements (class C amplifiers). Such amplifiers are especially inexpensive to manufacture, have high degree of efficiency, and therefore allow longer operation on a battery charge [15,64].

The digital modulation procedure for the GSM air interface comprises several steps for the generation of a high-frequency signal from channel-coded and enciphered data blocks (Figure 5.2).

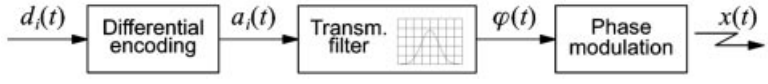

Figure 5.2: Steps of GSM digital modulation

The data d_i arrives at the modulator with a bit rate of 1625/6 kbit/s = 270.83 kbit/s (gross data rate) and are first differential-coded:

$$\hat{d}_i = (d_i + d_{i-1}) \bmod 2, \qquad d_i \in (0; 1)$$

From this differential data, the modulation data is formed, which represents a sequence of Dirac pulses:

$$a_i = 1 - 2\hat{d}_i$$

This bipolar sequence of modulation data is fed into the transmitter filter – also called a frequency filter – to generate the phase $\varphi(t)$ of the modulation signal. The impulse response $g(t)$ of this linear filter is defined by the convolution of the impulse response $h(t)$ of a

Gaussian low-pass with a rectangular step function:

$$g(t) = h(t) * \mathrm{rect}(t/T)$$

$$\mathrm{rect}(t/T) = \begin{cases} 1/T & \text{for } |t| < T/2 \\ 0 & \text{for } |t| \geq T/2 \end{cases}$$

$$h(t) = \frac{1}{\sqrt{2\pi}\sigma T} \exp\left(\frac{-t^2}{2\sigma^2 T^2} \right), \qquad \sigma = \frac{\sqrt{\ln 2}}{2\pi BT}, \qquad BT = 0.3$$

In the equations above, B is the 3 dB bandwidth of the filter $h(t)$ and T the bit duration of the incoming bit stream. The rectangular step function and the impulse response of the Gaussian lowpass are shown in Figure 5.3, and the resulting impulse response $g(t)$ of the transmitter filter is given in Figure 5.4 for some values of BT. Notice that with decreasing

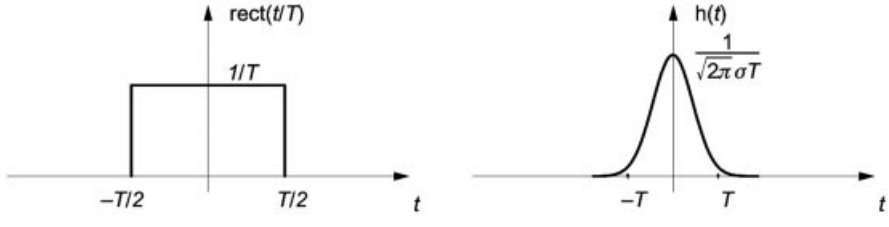

Figure 5.3: Impulse responses for the building blocks of the GMSK transmitter filter

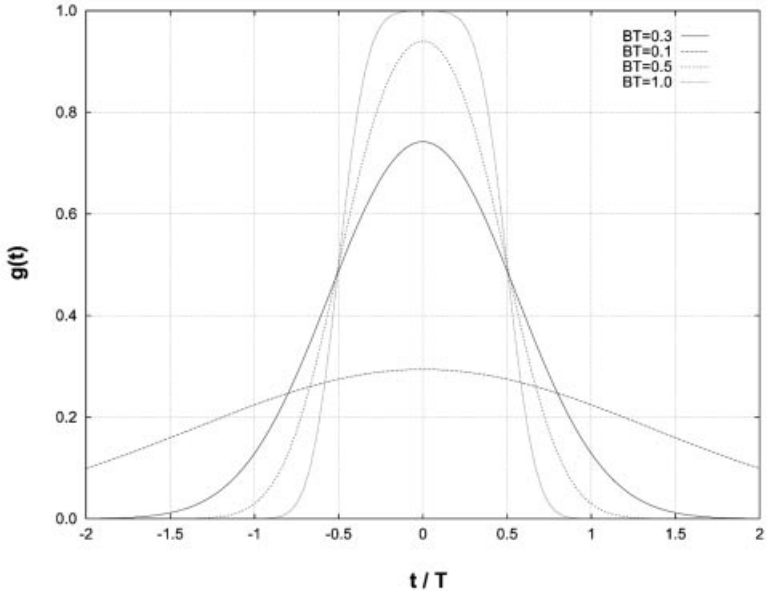

Figure 5.4: Impulse response $g(t)$ of the frequency filter (transmitter filter)

BT the impulse response becomes broader. For $BT \longrightarrow \infty$ it converges to the rect( ) function.

In essence, this modulation consists of a *Minimum Shift Keying* (MSK) procedure, where the data is filtered through an additional Gaussian lowpass before *Continuous Phase Modulation* (CPM) with the rectangular filter [15]. Accordingly it is called *Gaussian MSK* (GMSK). The Gaussian lowpass filtering has the effect of additional smoothing, but also of broadening the impulse response $g(t)$. This means that, on the one hand the power spectrum of the signal is made narrower, but on the other hand the individual impulse responses are ''smeared'' across several bit durations, which leads to increased intersymbol interference. This partial-response behavior has to be compensated for in the receiver by means of an equalizer [15].

The phase of the modulation signal is the convolution of the impulse response $g(t)$ of the frequency filter with the Dirac impulse sequence a_i of the stream of modulation data:

$$\varphi(t) = \sum_i a_i \pi\eta \int_{-\infty}^{t-iT} g(u)\mathrm{d}u$$

with the modulation index at $\eta = 1/2$, i.e. the maximal phase shift is $\pi/2$ per bit duration. Accordingly, GSM modulation is designated as 0.3-GMSK with a $\pi/2$ phase shift. The phase $\varphi(t)$ is now fed to a phase modulator. The modulated high-frequency carrier signal can then be represented by the following expression, where E_c is the energy per bit of the modulated data rate, f_0 the carrier frequency, and φ_0 is a random phase component staying constant during a burst:

$$x(t) = \sqrt{\frac{2E_c}{T}}\cos(2\pi f_0 t + \varphi(t) + \varphi_0)$$

5.2.2 Multiple Access, Duplexing, and Bursts

On the physical layer (OSI Layer 1), GSM uses a combination of FDMA and TDMA for multiple access. Two frequency bands 45 MHz apart have been reserved for GSM operation (Figure 5.5): 890–915 MHz for transmission from the mobile station, i.e. uplink, and 935–960 MHz for transmission from the base station, i.e. downlink. Each of these bands of 25 MHz width is divided into 124 single carrier channels of 200 kHz width. This variant of FDMA is also called *Multi-Carrier* (MC). In each of the uplink/downlink bands there remains a guardband of 200 kHz. Each *Radio Frequency Channel* (RFCH) is uniquely numbered, and a pair of channels with the same number form a duplex channel with a duplex distance of 45 MHz (Figure 5.5).

A subset of the frequency channels, the *Cell Allocation* (CA), is allocated to a base station, i.e. to a cell. One of the frequency channels of the CA is used for broadcasting the synchronization data (FCCH and SCH) and the BCCH. Therefore this channel is also called the *BCCH Carrier* (see Section 5.4). Another subset of the cell allocation is allocated to a mobile station, the *Mobile Allocation* (MA). The MA is used among others for the optional frequency hopping procedure (Section 5.2.3). Countries or areas which allow more than one mobile network to operate in the same area of the spectrum must have a

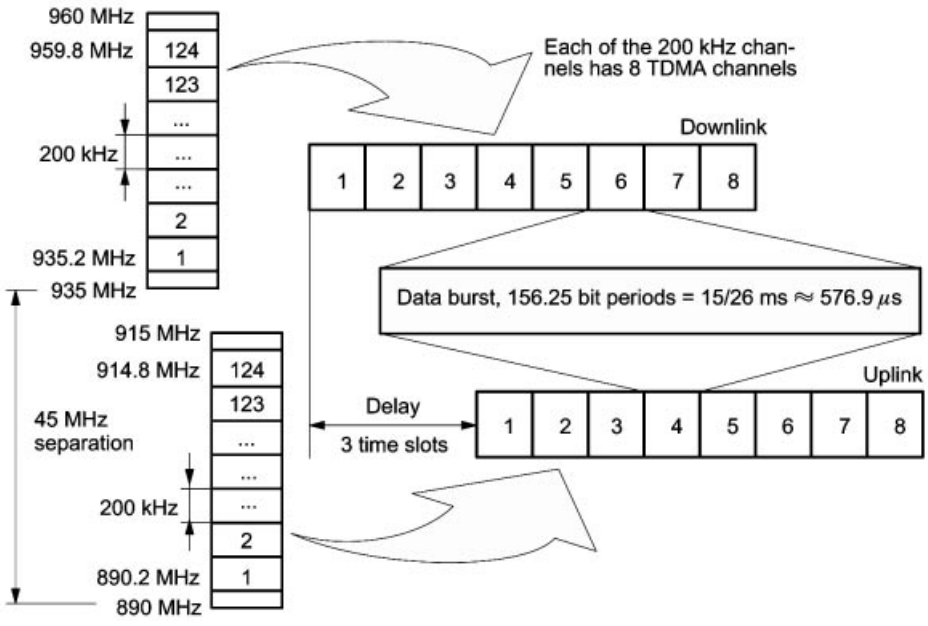

Figure 5.5: Carrier frequencies, duplexing, and TDMA frames

licensing agency which distributes the available frequency number space (e.g. the Federal Communication Commission in the USA or the ''Regulierungsbehörde für Telekommunikation und Post'' in Germany), in order to avoid collisions and to allow the network operators to perform independent network planning. Here is an example for a possible division: Operator A uses RFCH 2–13, 52–81, and 106–120, whereas operator B receives RFCH 15–50 and 83–103, in which case RFCH 1, 14, 51, 82, 104, 105, and 121–124 are left unused as additional guard bands.

Each of the 200 kHz channels is divided into eight time slots and thus carries eight TDMA channels. The eight time slots together form a TDMA frame (Figure 5.5). The TDMA frames of the uplink are transmitted with a delay of three time slots with regard to the downlink (see Figure 5.7). A mobile station uses the same time slots in the uplink as in the downlink, i.e. the time slots with the same number (TN). Because of the shift of three time slots, an MS does not have to send at the same time as it receives, and therefore does not need a duplex unit. This reduces the high-frequency requirements for the front end of the mobile and allows it to be manufactured as a less expensive and more compact unit.

So besides the separation into uplink and downlink bands – *Frequency Division Duplex* (FDD) with a distance of 45 MHz, the GSM access procedure contains a *Time Division Duplex* (TDD) component. Thus the MS does not need its own high-frequency duplexing unit, which again reduces cost as well as energy consumption.

Each time slot of a TDMA frame lasts for a duration of 156.25 bit periods and, if used, contains a data burst. The time slot lasts 15/26 ms = 576.9 μs; so a frame takes 4.615 ms. The same result is also obtained from the GMSK procedure, which realizes a gross data transmission rate of 270.83 kbit/s per carrier frequency.

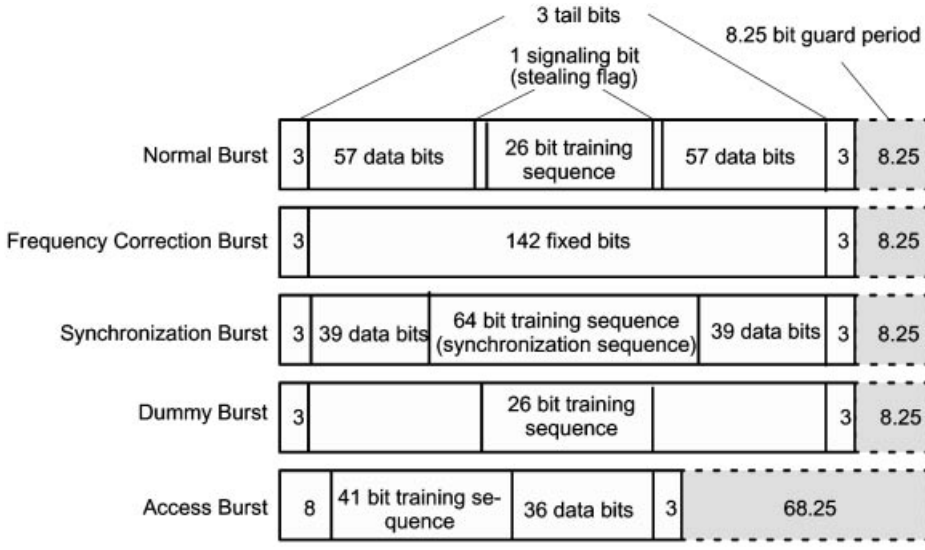

Figure 5.6: Bursts of the GSM TDMA procedure

There are five kinds of burst (Figure 5.6):

- *Normal Burst* (NB): The normal burst is used to transmit information on traffic and control (except RACH) channels. The individual bursts are separated from each other by guard periods during which no bits are transmitted. At the start and end of each burst are three tail bits which are always set to logical ''0.'' These bits fill a short time span during which transmitter power is ramped up or ramped down and during which no data transmission is possible. Furthermore, the initial zero bits are also needed for the demodulation process. The *Stealing Flags* (SF) are signaling bits which indicate whether the burst contains traffic data or signaling data. They are set to allow use of single time slots of the TCH in pre-emptive multiplexing mode, e.g. when, during a handover, fast transmission of signaling data on the FACCH is needed. This causes a loss of user data, i.e. these time slots are ''stolen'' from the traffic channel, hence the name ''stealing flag.'' A normal burst contains besides the synchronization and signaling bits (Figure 5.6) two blocks of 57 bits each of error-protected and channel-coded user data separated by a 26-bit midamble. This midamble consists of predefined, known bit patterns, the training sequences, which are used for channel estimation to optimize reception with an equalizer and for synchronization. With the help of these training sequences, the equalizer eliminates or reduces the intersymbol interferences which are caused by propagation time differences of the multipath propagation. Time differences of up to 16 μs can be compensated for. Eight different training sequences are defined for the NB which are designated by the *Training Sequence Code* (TSC). Initially, the TSC is obtained when the *Base Station Color Code* (BCC) is obtained, which is transmitted as part of the BSIC (see Section 3.2.9). Beyond that, training sequences can be individually assigned to mobile stations. In this case the TSC is contained in the Layer 3 message of the channel assignment (TCH or SDCCH). That way the base station tells a

mobile station which training sequence it should use with normal bursts of a specific traffic channel.

- *Frequency Correction Burst* (FB): This burst is used for the frequency synchronization of a mobile station. The repeated transmission of FBs is also called the *Frequency Correction Channel* (FCCH). Tail bits as well as data bits are all set to 0 in the FB. Due to the GSM modulation procedure (0.3-GMSK) this corresponds to broadcasting an unmodulated carrier with a frequency shift of 1625/24 kHz above the nominal carrier frequency. This signal is periodically transmitted by the base station on the BCCH carrier. It allows time synchronization with the TDMA frame of a mobile station as well as the exact tuning to the carrier frequency. Depending on the stability of its own reference clock, the mobile can periodically resynchronize with the base station using the FCCH.
- *Synchronization Burst* (SB): This burst is used to transmit information which allows the mobile station to synchronize time-wise with the BTS. Besides a long midamble, this burst contains the running number of the TDMA frame, the *Reduced TDMA Frame Number* (RFN) and the BSIC; the RFN is covered in Section 5.3. Repeated broadcasting of synchronization bursts is considered as the *Synchronization Channel* (SCH).
- *Dummy Burst* (DB): This burst is transmitted on one frequency of the cell allocation CA, when no other bursts are to be transmitted. The frequency channel used is the same one that carries the BCCH, i.e. it is the BCCH carrier. This ensures that the BCCH transmits a burst in each time slot which enables the mobile station to perform signal power measurements of the BCCH, a procedure also known as *quality monitoring*.
- *Access Burst* (AB): This burst is used for random access to the RACH without reservation. It has a guard period significantly longer than the other bursts. This reduces the probability of collisions, since the mobile stations competing for the RACH are not (yet) time-synchronized.

A single user gets one-eighth or 33.9 kbit/s of the gross data rate of 270.83 kbit/s. Considering a normal burst, 9.2 kbit/s are used for signaling and synchronization, i.e. tail bits, stealing flags and training sequences, including guard periods. The remaining 24.7 kbit/s are available for the transmission of (raw) user or control data on the physical layer.

5.2.3 Optional Frequency Hopping

Mobile radio channels suffer from frequency-selective interferences, e.g. frequency-selective fading due to multipath propagation phenomena. This selective frequency interference can increase with the distance from the base station, especially at the cell boundaries and under unfavorable conditions. Frequency hopping procedures change the transmission frequencies periodically and thus average the interference over the frequencies in one cell. This leads to a further improvement of the *Signal-to-Noise Ratio* (SNR) to a high enough level for good speech quality, so that conversations with acceptable quality can be conducted. GSM systems achieve a good speech quality with an SNR of about 11 dB. With frequency hopping a value of 9 dB is sufficient. GSM provides for an optional frequency hopping procedure which changes to a different frequency with each burst; this is known as *slow frequency hopping*. The resulting hopping rate is about 217 changes

per second, corresponding to the TDMA frame duration. The frequencies available for hopping, the hopping assignment, are taken from the cell allocation. The principle is illustrated in Figure 5.7, showing the time slot allocations for a full-rate TCH. The exact synchronization is determined by several parameters: the MA, a *Mobile Allocation Index Offset* (MAIO), a *Hopping Sequence Number* (HSN), and the *TDMA Frame Number* (FN); see Section 5.3. The use of frequency hopping is an option left to the network operator, which can be decided on an individual cell basis. Therefore a mobile station must be able to switch to frequency hopping if a base station notices adverse conditions and decides to activate frequency hopping.

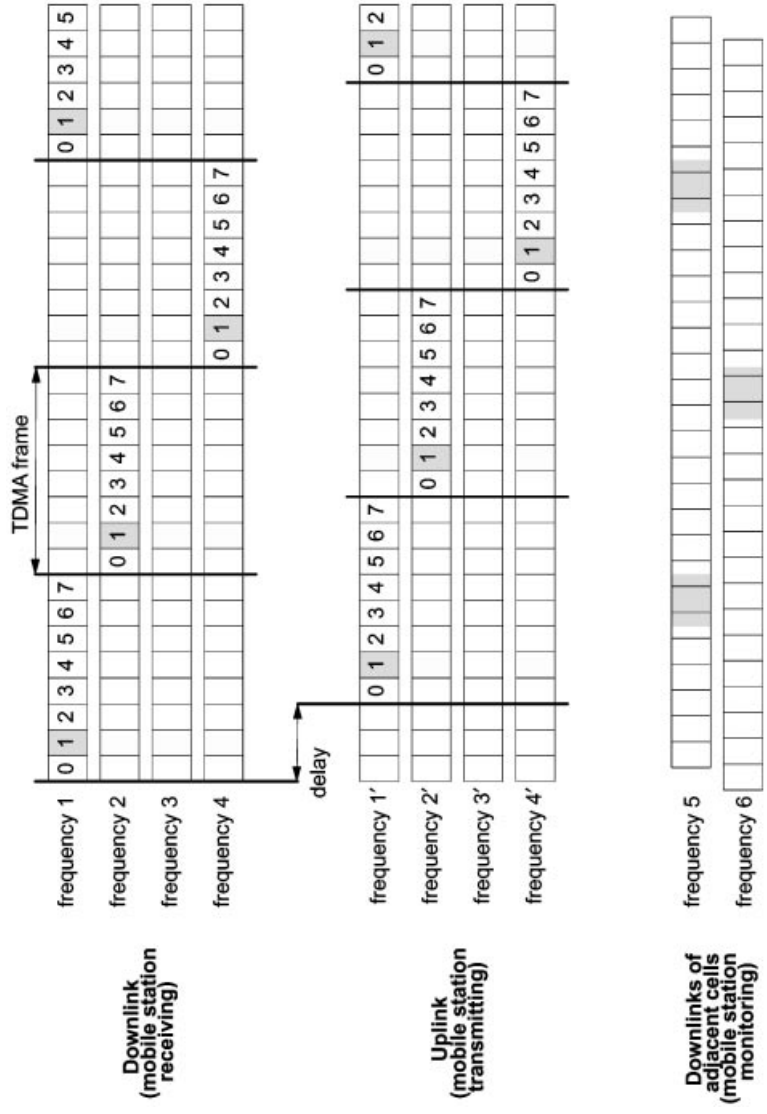

Figure 5.7: GSM full-rate traffic channel with frequency hopping

5.2.4 Summary

A physical GSM channel is defined by a sequence of frequencies and a sequence of TDMA frames. The RFCH sequence is defined by the frequency hopping parameters, and the temporal sequence of time slots of a physical channel is defined as a sequence of frame numbers and the time slot number within the frame. Frequencies for the uplink and downlink are always assigned as a pair of frequencies with a 45 MHz duplex separation.

As shown above, GSM uses a series of parameters to define a specific physical channel of a base station. Summarizing, these parameters are:

- *Mobile Allocation Index Offset* (MAIO)
- *Hopping Sequence Number* (HSN)
- *Training Sequence Code* (TSC)
- *Time Slot Number* (TN)
- *Mobile Allocation* (MA), also known as *RFCH Allocation*
- Type of logical channel carried on this physical channel
- The number of the logical subchannel (if used) – *Subchannel Number* (SCN)

Within a logical channel, there can be several subchannels (e.g. subrate multiplexing of the same channel type). The TDMA frame sequence can be derived from the type of the channel and the logical subchannel if present.

5.3 Synchronization

For the successful operation of a mobile radio system, synchronization between mobile stations and the base station is necessary. Two kinds of synchronization are distinguished: *frequency synchrony* and *time synchrony* of the bits and frames.

Frequency synchronization is necessary so that transmitter and receiver frequencies agree. The objective is to compensate for tolerances of the less expensive and therefore less stable oscillators in the mobile stations by obtaining an exact reference from the base station and to follow it.

Bit and frame synchrony are important in two regards for TDMA systems. First, the propagation time differences of signals from different mobile stations have to be adjusted, so that the transmitted bursts are received synchronously with the time slots of the base station and that bursts in adjacent time slots do not overlap and interfere with each other. Second, synchrony is needed for the frame structure since there is a higher-level frame structure superimposed on the TDMA frames for multiplexing logical signaling channels onto one physical channel. The synchronization procedures defined for GSM are explained in the following section.

5.3.1 Frequency and Clock Synchronization

A GSM base station transmits signals on the frequency carrier of the BCCH which allow a mobile station to synchronize with the base station. Synchronization means on the one hand the time-wise synchronization of mobile station and base with regard to bits and

frames, and on the other hand tuning the mobile station to the correct transmitter and receiver frequencies.

For this purpose, the BTS provides the following signals (Figure 5.6):

- *Synchronization Channel* (SCH) with extra long *Synchronization Bursts* (SB), which facilitate synchronization
- *Frequency Correction Channel* (FCCH) with *Frequency Correction Bursts* (FB)

Because of the 0.3-GMSK modulation procedure used in GSM, a data sequence of logical ''0'' generates a pure sine wave signal, i.e. broadcasting of the FB corresponds to an unmodulated carrier (frequency channel) with a frequency shift of 1625/24 kHz ($\approx$ 67.7 kHz) above the nominal carrier frequency (Figure 5.8). In this way, the mobile station can keep exactly synchronized by periodically monitoring the FCCH. On the other hand, if the frequency of the BCCH is still unknown, it can search for the channel with the highest signal level. This channel is with all likelihood a BCCH channel, because dummy bursts must be transmitted on all unused time slots in this channel, whereas not all time slots are always used on other carrier frequencies. Using the FCCH sine wave signal allows identification of a BCCH and synchronization of a mobile station's oscillator.

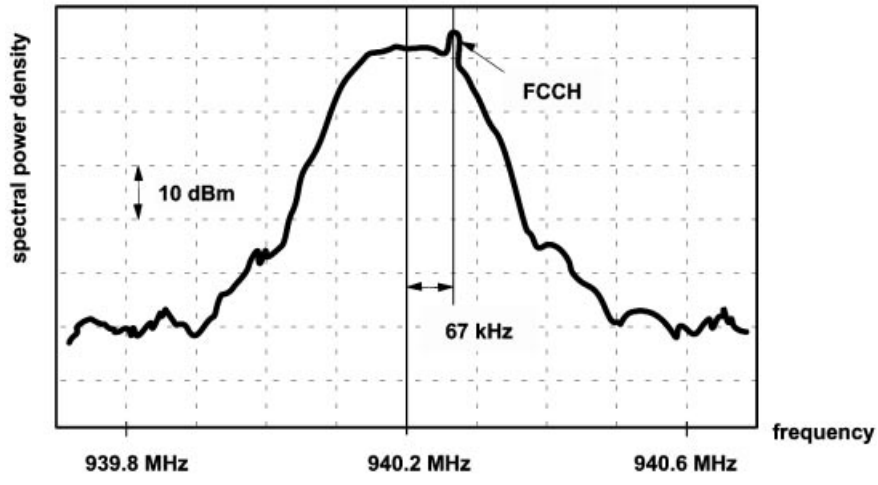

Figure 5.8: Typical power spectrum of a BCCH carrier

For the time synchronization, TDMA frames in GSM are cyclically numbered *modulo* 2 715 648 ($= 26 \times 51 \times 2^{11}$) with the FN. One cycle generates the so-called hyperframe structure which comprises 2 715 648 TDMA frames. This long numbering cycle of TDMA frames is used to synchronize the ciphering algorithm at the air interface (see Section 6.3). Each base station BTS periodically transmits the *Reduced TDMA Frame Number* (RFN) on the SCH. With each SB the mobiles thus receive information about the number of the current TDMA frame. This enables each mobile station to be time-synchronized with the base station.

The reduced TDMA frame number (RFN) has a length of 19 bits. It consists of three fields:

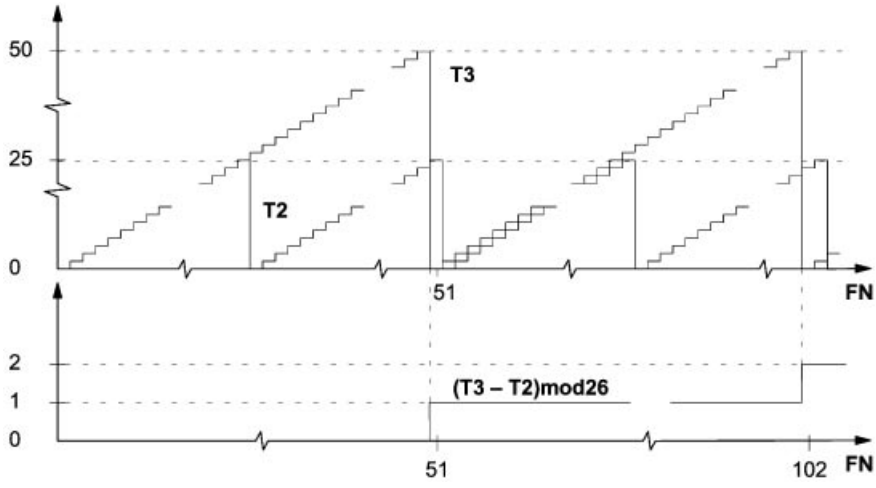

Figure 5.9: Values T2 and T3 for the calculation of RFN

T1 (11 bits), T2 (5 bits), and T3′ (3 bits). These three fields are defined by (with div designating integer division):

$$T1 = FN \text{ div } (26 \times 51) \; [0 - 2047]$$

$$T2 = FN \bmod 26 \; [0 - 25]$$

$$T3' = (T3 - 1) \text{ div } 10 \; [0 - 4]$$

$$\text{with } T3 = FN \bmod 51 \; [0 - 50]$$

The sequences of running values of T2 and T3 are illustrated in Figure 5.9. The value crucial for the reconstruction of the frame number FN is the difference (T3 − T2) between the two fields. The time synchronization of a mobile station and its time slots, TDMA frames, and control channels is based on a set of counters which run continuously, independent of mobile or base station transmission. Once these counters have been started and correctly initialized, the mobile station is in a synchronized state with the base station. The following four counters are kept for this purpose:

- Quarter Bit Counter counting the *Quarter Bit Number* (QN)
- Bit Counter counting the *Bit Number* (BN)
- Time Slot Counter counting the *Time Slot Number* (TN)
- Frame Counter counting the FN

Because of the bit and frame counting, these counters are of course interrelated, namely in such a way that the subsequent counter counts the overflows of the preceding counter. The following principle is used (Figure 5.10): QN is incremented every 12/13 μs; BN is obtained from it by integer division (BN = QN div 4). With each transition from 624 to 0 the time slot number TN is incremented, and each overflow of TN increments the frame counter FN by 1.

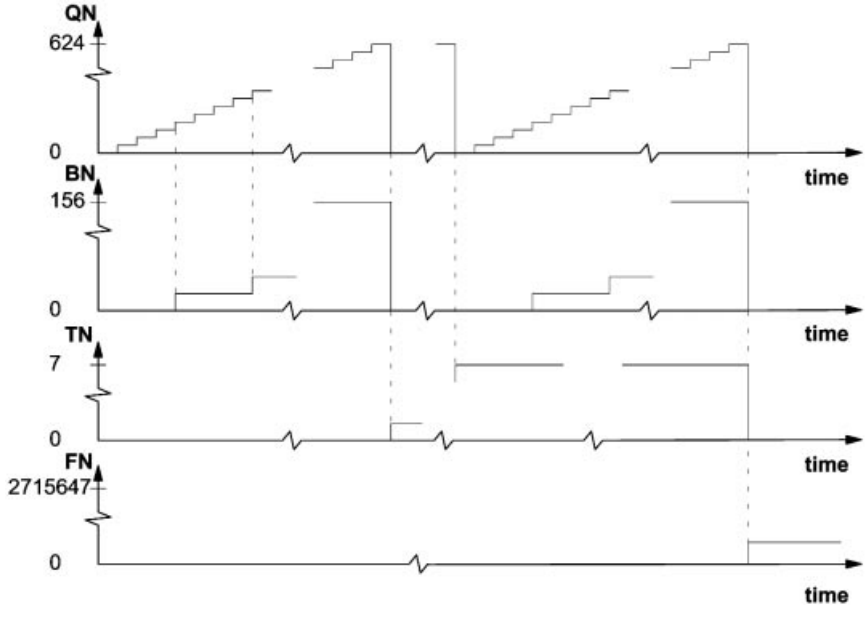

Figure 5.10: Synchronization timers, simplified: the TDMA frame duration is 156.25 bit times

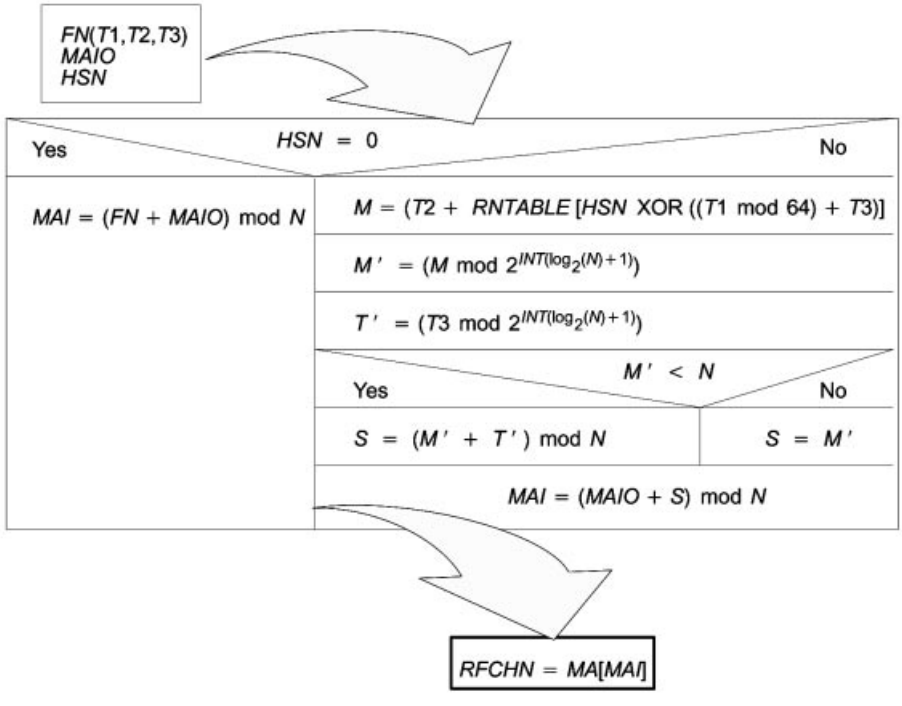

Figure 5.11: Generation of the GSM frequency hopping sequence

The timers can be reset and restarted when receiving an SB. The Quarter Bit Counter is set by using the timing of the training sequence of the burst, whereas the TN is reset to 0 with the end of the burst. The FN can then be calculated from the RFN transmitted on the SCH:

$$\mathrm{FN} = 51 \times ((\mathrm{T3} - \mathrm{T2}) \bmod 26) + \mathrm{T3} + 51 \times 26 \times \mathrm{T1}$$

$$\text{with } \mathrm{T3} = 10 \times \mathrm{T3}' + 1$$

It is important to recalculate T3 from $T3'$, although, because of the binary representation, only the integer part of the division by 10 is taken into account.

If the optional frequency hopping procedure is used (see Section 5.2.3), an additional mapping of the TDMA frame number onto the frequency to be used is required besides the evaluation of the synchronization signals from the FCCH and SCH. One has to obtain the index number of the frequency channel on which the current burst has to be transmitted from the MA table. This process uses a predefined RFNTABLE, the FN, and a HSN; see Figure 5.11. The MA holds *N* frequencies, with a maximum value of 64 for *N*. With this procedure, every burst is sent on a different frequency in a cyclic way.

5.3.2 Adaptive Frame Synchronization

The mobile station can be anywhere within a cell, which means the distance between mobile and base station may vary. Thus the signal propagation times between mobile and base station vary. Due to the mobility of the subscribers, the bursts received at the base would be offset. The TDMA procedure cannot tolerate such time shifts, since it is based on the exact synchronization of transmitted and received data bursts. Bursts transmitted by different mobile stations in adjacent time slots must not overlap when received at the base by more than the guard period (Figure 5.6), even if the propagation times within the cell are very different. To avoid such collisions, the start of transmission time from the mobile station is advanced in proportion to the distance from the base station. The process of adapting the transmissions from the mobile stations to the TDMA frame is called *adaptive frame alignment*.

For this purpose, the parameter *Timing Advance* (TA) in each SACCH Layer 1 protocol block is used (Figure 5.18). The mobile station receives from the base station on the SACCH downlink the TA value it must use; it reports the actually used value on the SACCH uplink. There are 64 steps for the timing advance which are coded as 0 to 63. One step corresponds to one bit period. Step 0 means no timing advance, i.e. the frames are transmitted with a time shift of 3 slots or 468.75 bit durations with regard to the downlink. At step 63, the timing of the uplink is shifted by 63 bit durations, such that the TDMA frames are transmitted on the uplink only with a delay of 405.75 bit durations. So the required adjustment always corresponds to twice the propagation time or is equal to the round-trip delay (Figure 5.12). In this way, the available range of values allows a compensation over a maximum propagation time of 31.5 bit periods ($\approx$ 113.3 μs). This corresponds to a maximum distance between mobile and base station of 35 km. A GSM cell may therefore have a maximum diameter of 70 km. The distance from the base station or the currently valid TA value for a mobile station is therefore an important handover criterion in GSM networks (see Section 8.4.3).

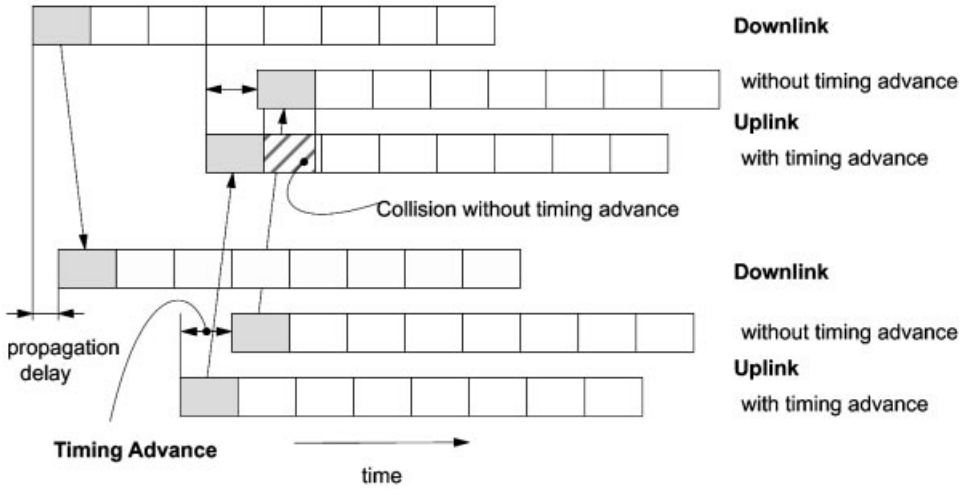

Figure 5.12: Operation of timing advance

The adaptive frame alignment technique is based on continuous measurement of propagation delays by the base station and corresponding timing advance activity by the mobile station. In the case of an (unreserved) random access to the RACH, a channel must first be established. The base station has in this case not yet had the opportunity to measure the distance of the mobile station and to transmit a corresponding timing advance command. If a mobile station transmits an access burst in the current time slot, it uses a timing advance value of 0 or a default value. To minimize collisions with subsequent time slots at the base station, the access burst AB has to be correspondingly shorter than the time slot duration (Figure 5.13). This explains the long duration AB of the guard period of 68.25 bit periods, which can compensate for the propagation delay if a mobile station sends an access burst from the boundary of a cell of 70 km diameter.

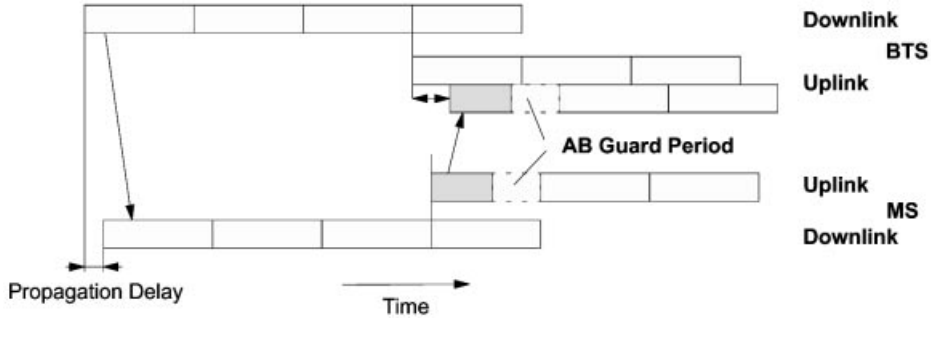

Figure 5.13: Timing for RACH random multiple access

5.4 Mapping of Logical Channels onto Physical Channels

The mapping of logical channels onto physical channels has two components: mapping in frequency and mapping in time. The mapping of a logical channel onto a physical channel

in the frequency domain is based on the *TDMA frame number* (FN), the frequencies allocated to base and mobile stations – CA and MA – and the rules for the optional frequency hopping (see Section 5.2.3).

In the time domain, logical channels are transported in the corresponding time slots of the physical channel. They are mapped onto physical channels in certain time-multiplexed combinations, where they can occupy a complete physical channel or just a part of a physical channel. Whereas user payload data is allocated a dedicated full-rate or half-rate channel, logical signaling (control) channels have to share a physical channel.

The logical channels are organized by the definition of complex superstructures on top of the TDMA frames, forming so-called multiframes, superframes and hyperframes (Figure 5.14). For the mapping of logical onto physical channels, we are interested in the multiframe domain. These multiframes allow us to map (logical) subchannels onto physical channels. Two kinds of multiframes are defined (Figure 5.15): a multiframe consisting of 26 TDMA frames (predominantly payload – speech and data – frames) and a multiframe of 51 TDMA frames (predominantly signaling frames).

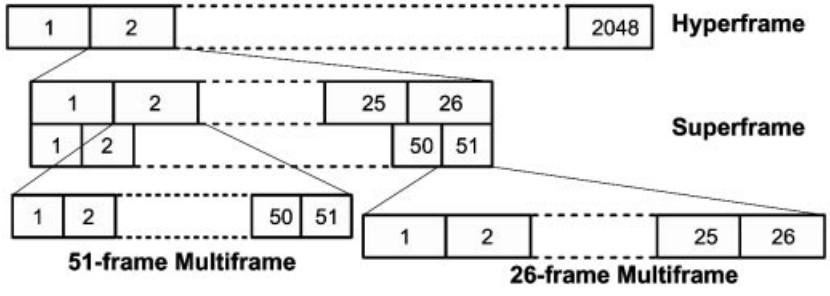

Figure 5.14: GSM frame structures

Each hyperframe is divided into 2048 superframes. With its long cycle period of 3 h 28 min 53.760 s, it is used for the synchronization of user data encryption. A superframe consists of 1326 consecutive TDMA frames which therefore lasts for 6.12 s, like 51 multiframes of 26 TDMA frames or 26 multiframes of 51 TDMA frames. These multiframes are again used to multiplex the different logical channels onto a physical channel as shown below.

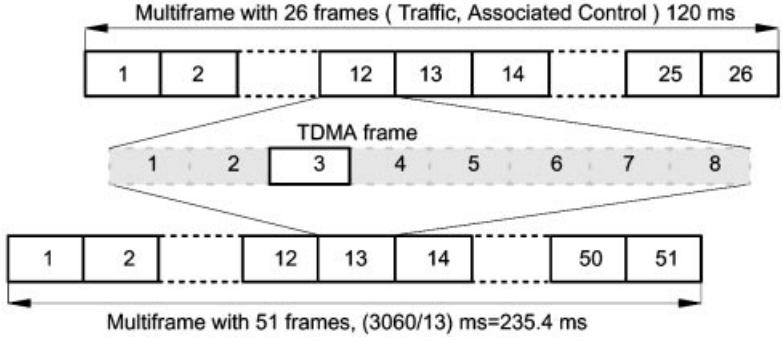

Figure 5.15: GSM multiframes

5.4.1 26-Frame Multiframe

Each 26 subsequent TDMA frames form a multiframe which multiplexes two logical channels, a TCH and the SACCH, onto the physical channel (Figure 5.16). This process uses only one time slot per TDMA frame for the corresponding multiframe (e.g. time slot 3 in Figure 5.15), since a physical channel consists of just one time slot per TDMA frame. Besides the 24 TCH frames for user data, this multiframe also contains an AC frame for signaling data (SACCH data). One frame (the 26th) remains unused in the case of a full-rate TCH (IDLE/AC); it is reserved for the introduction of two half-rate TCHs; then the 26th frame will be used to carry the SACCH channels of the other half.

The data of the *Fast Associated Control Channel* (FACCH) is transmitted by occupying one half of the bits in eight consecutive bursts, by ''stealing'' these bits from the TCH. For this purpose, the Stealing Flags of the normal bursts are set (Figure 5.6).

A subscriber has available a gross data rate of 271 kbit/s ÷ 8 = 33.9 kbit/s (Section 5.2). Of this budget, 9.2 kbit/s are for signaling, synchronization, and guard periods of the burst. Of the remaining 24.7 kbit/s, in the case of the 26-frame multiframe, 22.8 kbit/s are left for the coded and enciphered user data of a full-rate channel, and 1.9 kbit/s remain for the SACCH and IDLE.

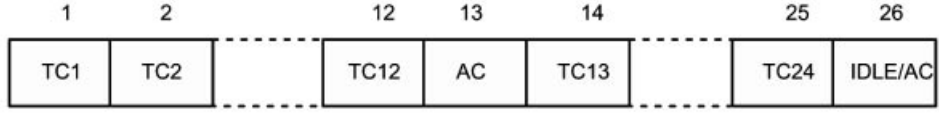

Figure 5.16: Channel organization in a 26-frame multiframe

5.4.2 51-Frame Multiframe

For the transmission of the control channels which are not associated with a TCH (all except FACCH and SACCH), a multiframe is formed consisting of 51 consecutive TDMA frames (Figure 5.6). According to channel configuration (Section 5.1), the multiframe is used differently. In each case, multiframes of 51 TDMA frames serve the purpose of mapping several logical channels onto a physical channel.

Furthermore, some of these control channels are unidirectional, which results in different structures for uplink and downlink. For some configurations, two adjacent multiframes are required to map all the logical channels. Some examples are illustrated in Figure 5.17. They correspond to the combinations B2, B3, and B4 in Table 5.3 whereas for channels SDCCH and SACCH some 4 or 8 logical subchannels have been defined (D0, D1, …, A0, A1, …). One of the frequency channels of the CA of a base station is used to broadcast synchronization data (FCCH and SCH) and the BCCH. Since the base station has to transmit in each time slot of the BCCH carrier to enable a continuous measurement of the BCCH carrier by the mobile station, a *Dummy Burst* (DB) is transmitted in all time slots with no traffic.

On time slot 0 of the BCCH carrier, only two combinations of logical channels may be transmitted, the combinations B2 or B3 from Table 5.3: (BCCH + CCCH + FCCH +

SCH + SDCCH + SACCH or BCCH + CCCH + FCCH + SCH). No other time slot of the CA must carry this combination of logical channels.

As one can see in Figure 5.17, in the time slot 0 of the BCCH carrier of a base station (downlink) the frames 1, 11, 21,… are FCCH frames, and the subsequent frames 2, 12, 22,… form SCH frames. Frames 3, 4, 5, 6 of the 51-frame BCCH multiframe transport the

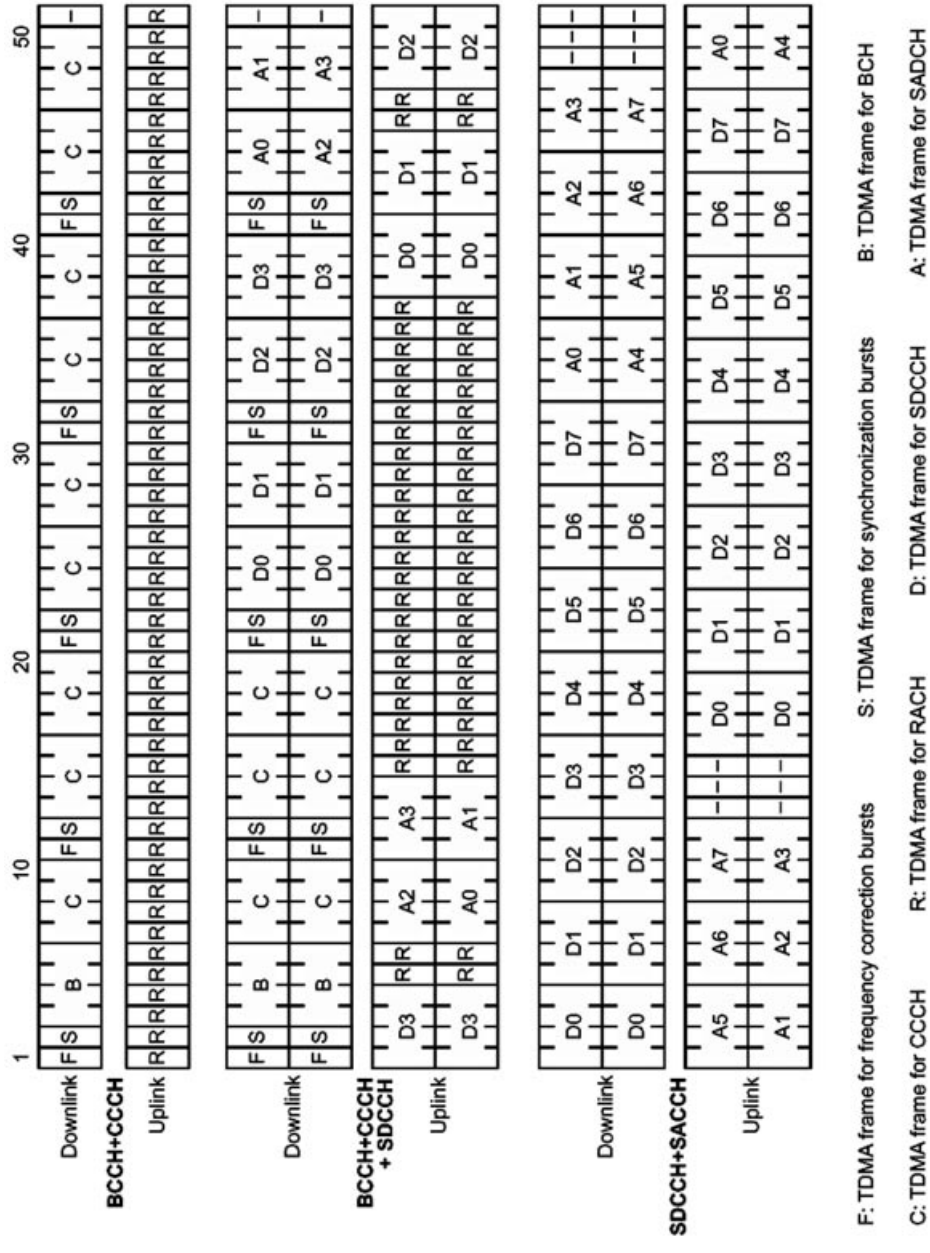

Figure 5.17: Channel organization in a 51-frame multiframe

appropriate BCCH information, whereas the remaining frames may contain different combinations of logical channels. Once the mobile station has synchronized by using the information from FCCH and SCH, it can determine from the information in the FCCH and SCCH how the remainder of the BCCH is constructed. For this purpose, the base station *Radio Resource Management* periodically transmits a set of messages to all mobile stations in this cell.

These *System Information Messages* comprise six types, of which only Types 1–4 are of interest here. Using the TDMA frame number (FN), one can determine which type is to be sent in the current time slot by calculating a *Type Code* (TC):

$$TC = (FN \text{ div } 51) \bmod 8$$

Table 5.5 shows how the TC determines the type of the system information message to be sent within the current multiframe.

Of the parameters contained in such a message, the following are of special interest: BS_CC_CHANS determines the number of physical channels which support a CCCH. The first CCCH is transmitted in time slot 0, the second one in time slot 2, the third one in time slot 4, and the fourth one in time slot 6 of the BCCH carrier. Another parameter, BS_CCCH_SDCCH_COMB, determines whether the DCCHs SDCCH(0–3) and SACCH(0–3) are transmitted together with the CCCH on the same physical channel. In this case, each of these dedicated control channels consists of four subchannels.

Table 5.5: Mapping of frame number onto BCCH message

TC	System information message
0	Type 1
1	Type 2
2, 6	Type 3
3, 7	Type 4
4, 5	Any (optional)

Each of the CCCHs of a base station is assigned a group CCCH_GROUP of mobile stations. Mobile stations are allowed random access (RACH) or receive paging information (PCH) only on the CCCH assigned to this group. Furthermore, a mobile station needs only to listen for paging information on every *N*th block of the *Paging Channel* (PCH). The number *N* is determined by multiplying the number of paging blocks per 51-frame multiframe of a CCCH with the parameter BS_PA_MFRMS designating the number of multiframes between paging frames of the same *Paging Group* (PAGING_GROUP).

Especially in cells with high traffic, the CCCH and paging groups serve to subdivide traffic and to reduce the load on the individual CCCHs. For this purpose, there is a simple algorithm which allows each mobile station to calculate its respective CCCH_GROUP

and PAGING_GROUP from its IMSI and parameters BS_CC_CHANS, BS_PA_MFRMS and *N*.

5.5 Radio Subsystem Link Control

The radio interface is characterized by another set of functions of which only the most important ones are discussed in the following. One of these functions is the control of the radio link: *Radio Subsystem Link Control*, with the main activities of received-signal quality measurement (quality monitoring) for cell selection and handover preparation, and of transmitter power control.

If there is no active connection, i.e. if the mobile station is at rest, the BSS has no tasks to perform. The MS, however, is still committed to continuously observing the BCCH carrier of the current and neighboring cells, so that it would be able to select the cell in which it can communicate with the highest probability. If a new cell needs to be selected, a *Location Update* may become necessary.

During a connection (TCH or SDCCH), the functions of channel measurement and power control serve to maintain and optimize the radio channel; this also includes adaptive frame alignment (Section 5.3.1) and frequency hopping (Section 5.2.3). Both need to be done until the current base can hand over the current connection to the next base station.

These link control functions are performed over the SACCH channel. Two fields are defined in an SACCH block (Figure 5.18) for this purpose, the power level and the TA. On the downlink, these fields contain values as assigned by the BSS. On the uplink, the MS inserts its currently used values. The quality monitoring measurement values are transmitted in the data part of the SACCH block.

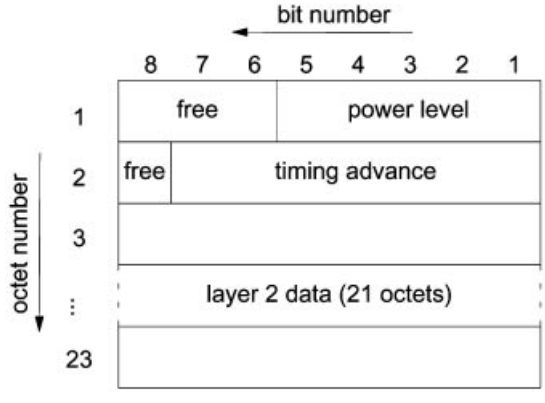

Figure 5.18: SACCH block format

The following illustrates the basic operation of the *Radio Subsystem Link Control* at the BSS side for an existing connection; the detailed explanation of the respective functions is given later. In principle, the radio link control can be subdivided into three tasks: measurement collection and processing, transmitter power control, and handover control.

In the example of Figure 5.19, the process BSS_Link_Control starts at initialization the processes BSS_Power_Control and BSS_HO_Control and then enters a measurement loop, which is only left when the connection is terminated. In this loop, measurement data is periodically received (every 480 ms) and current mean values are calculated. At first, these measurement data are supplied to the transmitter power control to adapt the power of MS and BSS to a new situation if necessary. Thereafter, the measurement data and the result of the power control activity are supplied to the handover process, which can then decide whether a handover is necessary or not.

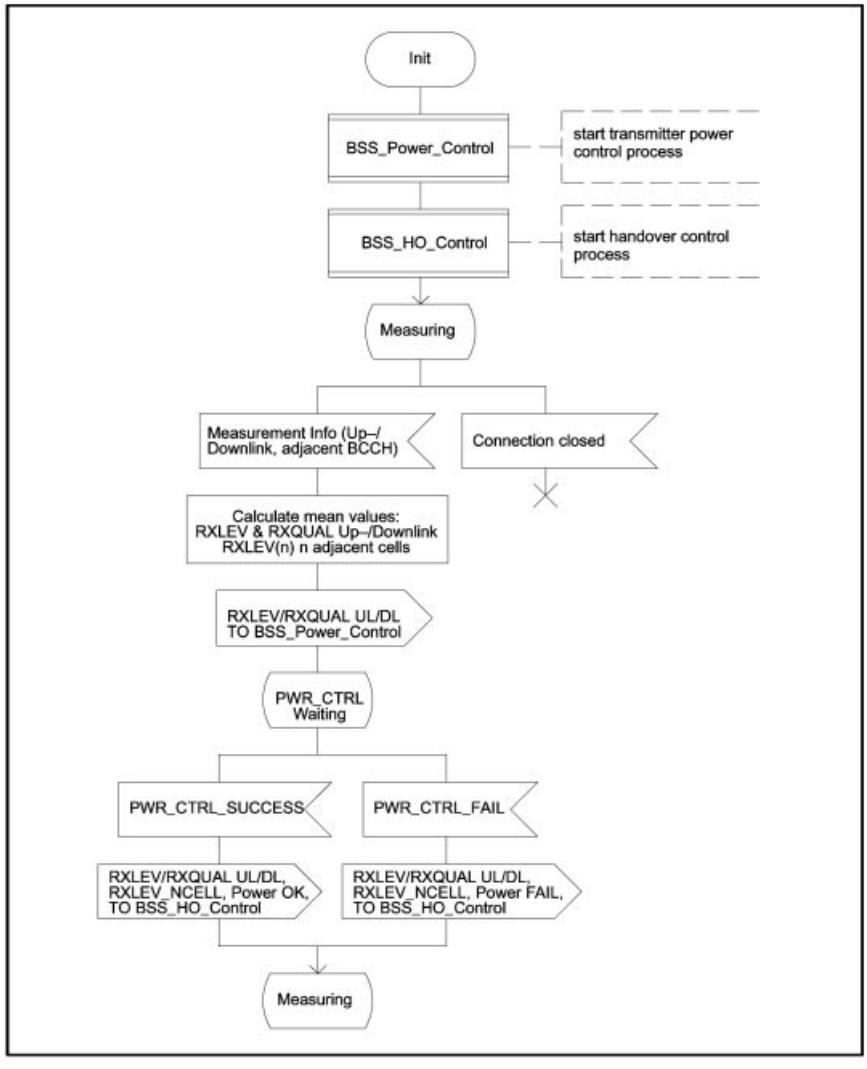

Figure 5.19: Principal operation of the radio subsystem link control

5.5.1 Channel Measurement

The task of *Radio Subsystem Link Control* in the mobile station includes identification of the reachable base stations and measurement of their respective received signal level and channel quality (quality monitoring task). In idle mode, these measurements serve to select the current base station, whose PCH is then periodically examined and on whose RACH desired connections can be requested.

During a connection, i.e. on a TCH or SDCCH with respective SACCH/FACCH, this measurement data is transmitted on the SACCH to the base station as a *measurement report/measurement info*. These reports serve as inputs for the handover and power control algorithms.

The measurement objects are on the one hand the uplink and downlink of the current channel (TCH or SDCCH), and on the other hand the BCCH carriers which are continuously broadcast with constant power by all BTSs in all time slots. It is especially important to keep the transmitter power of the BCCH carriers constant to allow comparisons between neighboring base stations. A list of neighboring base station's BCCH carrier frequencies, called the *BCCH Allocation* (BA) is supplied to each mobile by its current BTS, to enable measurement of all cells which are candidates for a handover. The cell identity is broadcast as the BSIC on the BCCH. Furthermore, up to 36 BCCH carrier frequencies and their BSICs can be stored on the SIM card. In principle, GSM uses two parameters to describe the quality of a channel: the *Received Signal Level* (RXLEV), measured in dBm, and the *Received Signal Quality* (RXQUAL), measured as bit error ratio in percent before error correction (Tables 5.6 and 5.7). The received signal power is measured continuously by mobile and base stations in each received burst within a range of -110 dBm to -48 dBm. The respective RXLEV values are obtained by averaging.

The bit error ratio before error correction can be determined in a variety of ways. For example, it can be estimated from information obtained from channel estimation for equalization from the training sequences, or the number of erroneous (corrected) bits can be determined through repeated coding of the decoded, error-corrected data blocks and comparison with the received data. Since the data before error correction is presented as blocks of 456 bits (see Section 6.2 and Figure 6.10), the bit error ratio can only be given with a quantizing resolution of 2×10^{-3}. Again, the value of RXQUAL is determined from this information by averaging.

Table 5.6: Measurement range of the received signal level

Level	Received signal level (dBm)	
	From	To
RXLEV_0	–	−110
RXLEV_1	−110	−109
⋮	⋮	⋮
RXLEV_62	−49	−48
RXLEV_63	−48	–

Table 5.7: Measurement range of bit error ratio

Level	Bit error ratio (%)	
	From	To
RXQUAL_0	–	0.2
RXQUAL_1	0.2	0.4
RXQUAL_2	0.4	0.8
RXQUAL_3	0.8	1.6
RXQUAL_4	1.6	3.2
RXQUAL_5	3.2	6.4
RXQUAL_6	6.4	12.8
RXQUAL_7	12.8	–

5.5.1.1 Channel Measurement during Idle Mode

In idle mode (see also Figure 7.17) the mobile station must always stay aware of its environment. The main purpose is to be able to assign a mobile station to a cell, whose BCCH carrier it can decode reliably. If this is the case, the mobile station is able to read system and paging information. If there is a desire to set up a connection, the mobile station can most likely communicate with the network.

There are two possible starting situations:

- The MS has no a priori knowledge about the network at hand, especially which BCCH carrier frequencies are in use.
- The MS has a stored list of BCCH carriers.

In the first case, the more unfavorable of the two, the mobile has to search through all the 124 GSM frequencies, measure their signal power level, and calculate an average from at least five measurements. The measurements of the individual carriers should be evenly distributed over an interval of 3–5 s. After at most 5 s, a minimum of 629 measurement values are available that allow the 124 RXLEV values to be determined. The carriers with the highest RXLEV values are very likely BCCH carriers, since continuous transmission is required on them. Final identification occurs with the frequency correction burst of the FCCH. Once the received BCCH carriers have been found, the mobile station starts to synchronize with each of them and reads the system information, beginning with the BCCH with the highest RXLEV value.

This orientation concerning the current location can be accelerated considerably, if a list of BCCH carriers has been stored on the SIM card. Then the mobile station tries first to synchronize with some known carrier. Only if it cannot find any of the stored BCCH carrier frequencies, it does start with the normal BCCH search. A mobile station can store several lists for the recently visited networks.

5.5.1.2 Channel Measurement during a Connection

During a traffic (TCH) or signaling (SDCCH) connection, the channel measurement of the mobile station occurs over an SACCH interval, which comprises 104 TDMA frames in the case of a TCH channel (480 ms) or 102 TDMA frames (470.8 ms) in the case of an SDCCH channel.

For the channel at hand, two parameters are determined: the received signal level RXLEV and the signal quality RXQUAL. These two values are averaged over a SACCH interval (480 or 470.8 ms) and transmitted to the base station on the SACCH as a measurement report/measurement info. This way the downlink quality of the channel assigned to the mobile station can be judged. In addition to these measurements of the downlink by the mobile station, the base station also measures the RXLEV and RXQUAL values of the respective uplink.

In order to make a handover decision, information about possible handover targets must be available. For this purpose, the mobile station has to observe continuously the BCCH carriers of up to six neighboring base stations. The RXLEV measurements of the neighboring BCCH carriers are performed during the mobile station's unused time slots (see Figure 5.7). The BCCH measurement results of the six strongest signals are included in the measurement report transmitted to the BSS.

However, the received signal power level and the frequency of a BCCH carrier alone are not a sufficient criterion for a successful handover. Because of the frequency reuse in cellular networks, and especially in the case of small clusters, it is possible that a cell can receive the same BCCH carrier from more than one neighboring cell, i.e. there exist several neighboring cells which use the same BCCH carrier. It is therefore necessary, to also know the identity (BSIC) of each neighboring cell. Simultaneously with the signal level measurement, the mobile station has to synchronize with each of the six neighboring BCCHs and read at least the SCH information.

For this purpose, one must first search for the FCCH burst of the BCCH carrier; then the SCH can be found in the next TDMA frame. Since the FCCH/SCH/BCCH is always transmitted in time slot 0 of the BCCH carrier, the search during a conversation for FCCHs can only be conducted in unused frames, i.e. in case of a full-rate TCH in the IDLE frame of the multiframe (frame number 26 in Figures 5.16 and 5.20). These free frames are therefore also known as *search frames*.

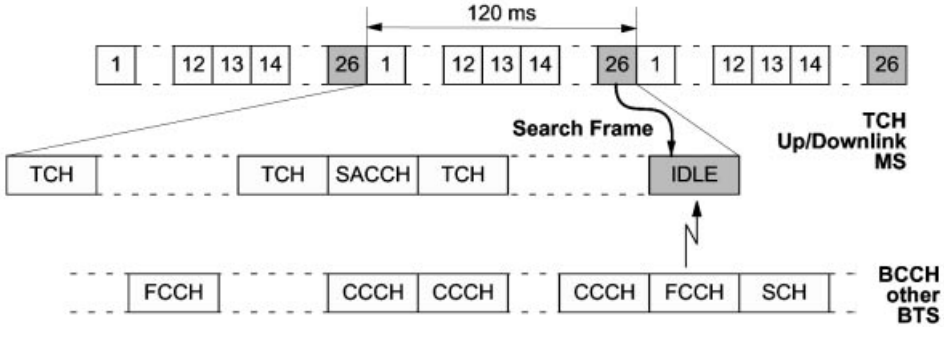

Figure 5.20: Synchronization with adjacent cells during a call

Therefore there are exactly four search frames within an SACCH block of 480 ms (four 26-frame multiframes of 120 ms). The mobile station has to examine the surrounding BCCH carriers for FCCH bursts, in order to synchronize with them and to decode the SCH. But how can one search for synchronization points exactly within these frames during synchronized operation?

This is possible because the actual traffic channel and the respective BCCH carriers use different multiframe formats. Whereas the traffic channel uses the 26-frame multiframe format, time slot 0 of the BCCH carrier with the FCCH/SCH/BCCH is carried on a 51-frame multiframe format. This ratio of the different multiframe formats has the effect that the relative position of the search frames (frame 26 in a TCH multiframe) is shifting with regard to the BCCH multiframe by exactly one frame each 240 ms (Figure 5.21). Figura-

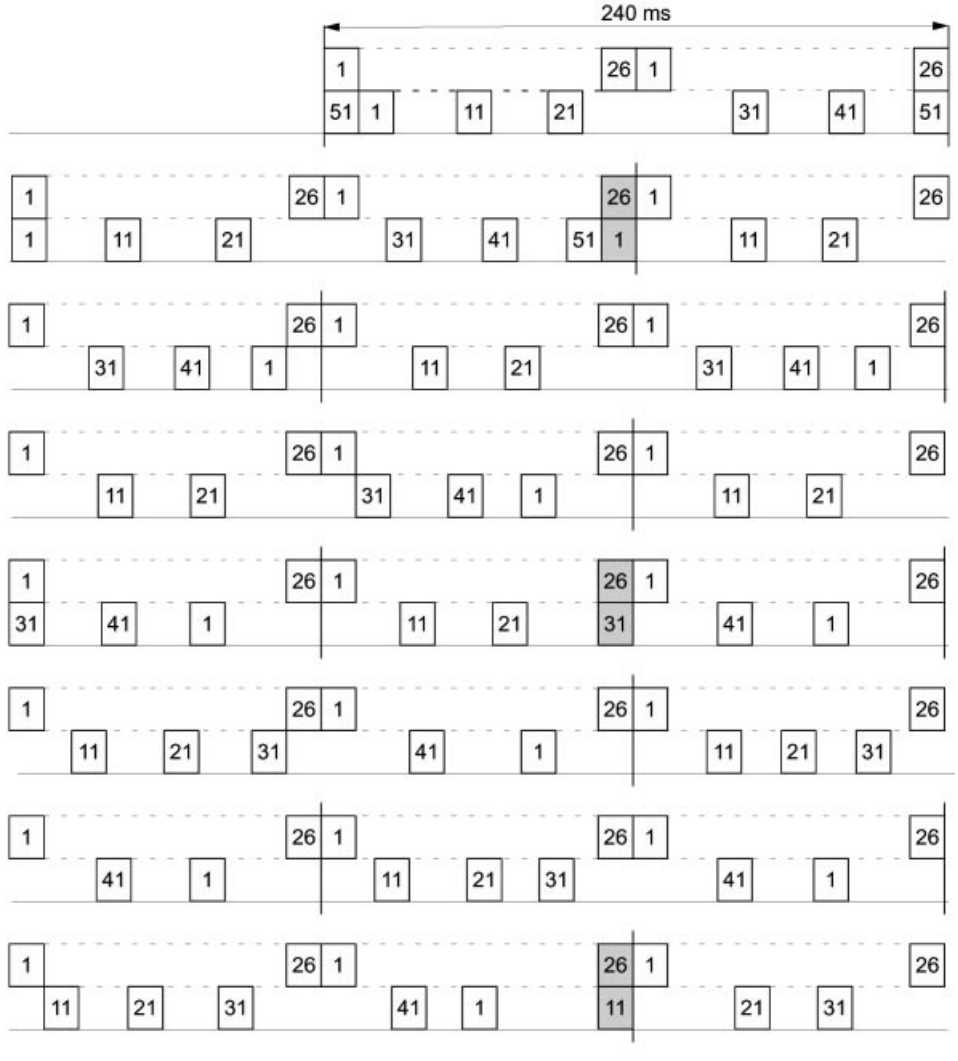

Figure 5.21: Principle of FCCH search during the search frame

tively speaking, the search frame is travelling along the BCCH multiframe in such a way that at most after 11 TCH multiframes (=1320 ms) a frequency correction burst of a neighboring cell becomes visible in a search frame.

In this way, the mobile station is able to determine the BSIC for the respective RXLEV measurement value. Only BCCH carrier measurements whose identity can be established without doubt are included in the measurement report to the base station.

The base station can now make a handover decision based on these values, on the distance of the mobile station, and on the momentary interference of unused time slots.

The algorithm for handover decisions has not been included in the GSM standard. The network operators may use algorithms which are optimized for their network or the local situation. GSM only gives a basic proposal which satisfies the minimum requirements for a handover decision algorithm. This algorithm defines threshold values, which must be violated in one or the other direction to arrive at a safe handover decision and to avoid so-called ping-pong handovers, which oscillate between two cells. Although the decision algorithm is part of *Radio Subsystem Link Control*, its discussion is postponed and it is treated together with handover signaling (see Section 8.4.3).

5.5.2 Transmission Power Control

Power classes (Table 5.8) are used for classification of base and mobile stations. The transmission power can also be controlled adaptively. As part of the *Radio Subsystem Link Control*, the mobile station's transmitter power is controlled in steps of 2 dBm.

The GSM transmitter power control has the purpose of limiting the mobile station's transmitter power to the minimum necessary level, in such a way that the base station receives signals from different mobile stations at approximately the same power level. Sixteen power control steps are defined for this purpose: Step 0 (43 dBm = 20 W) to Step 15 (13 dBm). Starting with the lowest, Step 15, the base station can increment the transmitter power of the mobile station in steps of 2 dBm up to the maximum power level of the

Table 5.8: GSM power classes

Power class	Max. peak transmission power (W)	
	Mobile station (dBm)	Base station
1	20 (43)	320
2	> 8 (39)	160
3	> 5 (37)	80
4	> 2 (33)	40
5	> 0.8 (29)	20
6	–	10
7	–	5
8	–	2.5

Table 5.9: Thresholds for transmitter power control

Threshold parameter	Typical value (dBm)	Meaning
L_RXLEV_UL_P	−103 to −73	
L_RXLEV_DL_P	−103 to −73	Threshold for raising of transmission power
L_RXQUAL_UL_P	–	in uplink or downlink
L_RXQUAL_DL_P	–	
U_RXLEV_UL_P	–	
U_RXLEV_DL_P	–	Threshold for reducing of transmission
U_RXQUAL_UL_P	–	power in uplink or downlink
U_RXQUAL_DL_P	–	

respective power class of the mobile station. Similarly, the transmitter power of the base station can be controlled in steps of 2 dBm, with the exception of the BCCH carrier of the base station, which must remain constant to allow comparative measurements of neighboring BCCH carriers by the mobile stations.

Transmission power control is based on the measurement values RXLEV and RXQUAL, for which one has defined upper and lower thresholds for uplink and downlink (Table 5.9). Network management defines the adjustable parameters *P* and *N*. If the values of *P* for the last *N* calculated mean values of the respective criterion (RXLEV or RXQUAL) are above or below the respective threshold value, the BSS can adjust the transmitter power (Figure 5.22).

If the thresholds U_xx_UL_P of the uplink are exceeded, the transmission power of the mobile station is reduced; in the other case, if the signal level is below the threshold L_xx_UL_P, the mobile station is ordered to increase its transmitter power. In an analogous way, the transmitter power of the base station can be adjusted, when the criteria for the downlink are exceeded in either direction.

Even if the mobile or base station signal levels stay within the thresholds, the current RXLEV/RXQUAL values can cause a change to another channel of the same or another cell based on the handover thresholds (Table 8.1). For this reason, checking for transmitter thresholds is immediately followed by a check of the handover thresholds as the second part of the *Radio Subsystem Link Control* (Figures 5.19 and 8.17). If one of the threshold values is exceeded in either direction and the transmitter power cannot be adjusted accordingly, i.e. the respective transmitter power has reached its maximum or minimum value, this is an overriding cause for handover (PWR_CTRL_FAIL, see Table 8.2) which the BSS must communicate immediately to the MSC (see Section 8.4).

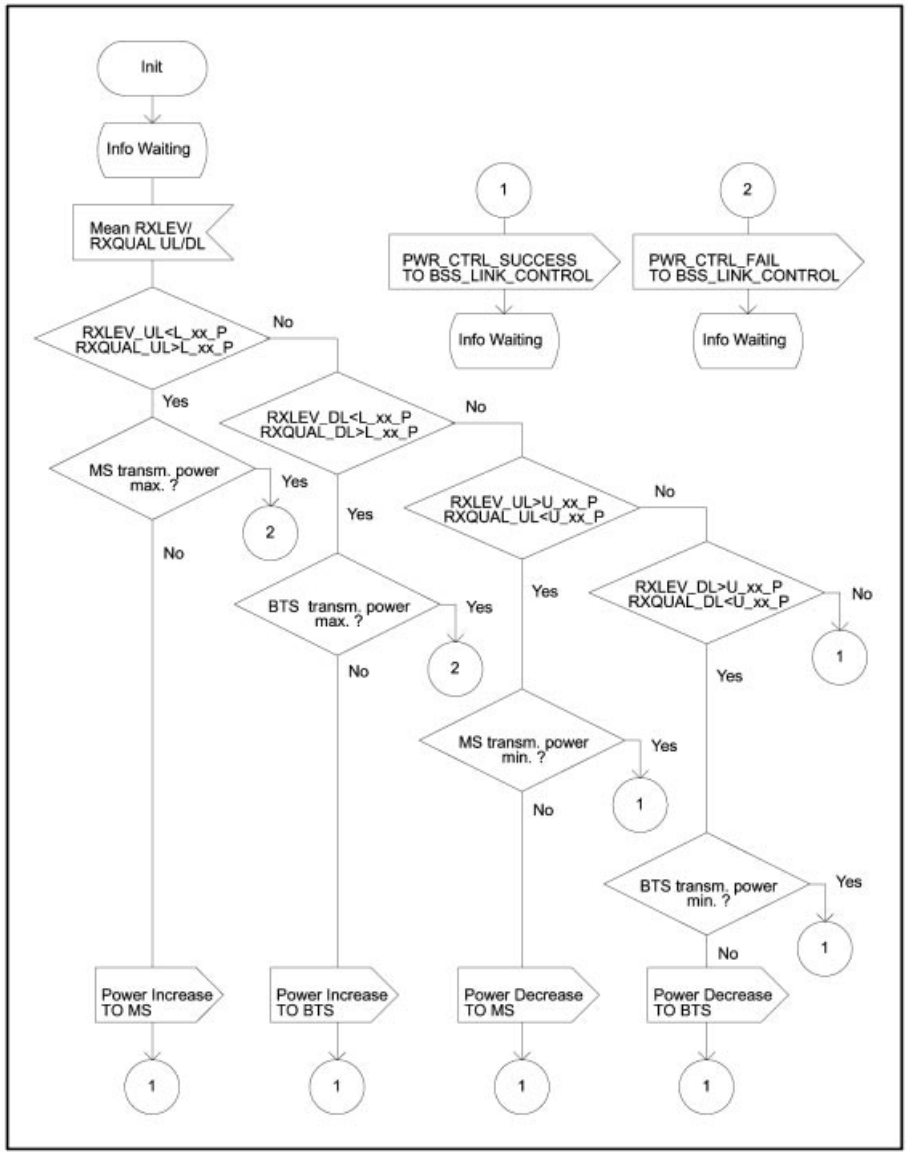

Figure 5.22: Schematic operation of transmitter power control

5.5.3 Disconnection due to Radio Channel Failure

The quality of a radio channel can vary considerably during an existing connection, or it can even fail in the case of shadowing. This should not lead to immediate disconnection, since such failures are often of short duration. For this reason GSM has a special algorithm within the *Radio Subsystem Link Control* which continuously checks for connectivity. It consists of recognizing a radio link failure by the inability to decode signaling information on the SACCH. This connectivity check is done both in the mobile as well as in the base

station. The connection is not immediately terminated, but is delayed so that only repeated consecutive failures (erroneous messages) represent a valid disconnect criterion. On the downlink, the mobile station must check the frequency of erroneous, nondecodable messages on the SACCH. The error protection on the SACCH has very powerful error correction capabilities and thus guarantees a very low probability of 10^{-10} for nonrecognized, wrongly corrected bits in SACCH messages.

In this way, erroneous SACCH messages supply a measure for the quality of the downlink, which is already quite low when errors on the SACCH cannot be corrected any more. If a consecutive number of SACCH messages is erroneous, the link is considered bad, and the connection is terminated. For this purpose, a counter S has been defined which is incremented by 2 with each arrival of an error-free message, and decremented by 1 for each erroneous SACCH message (Figure 5.23). When the counter reaches the value $S = 0$, the downlink is considered as failing, and the connection is terminated. This failure is signaled to the upper layers, *Mobility Management* (MM), which can start a *call reestablishment procedure*. The maximum value RADIO_LINK_TIMEOUT for the counter S therefore determines the interval length during which a channel has to fail before a connection is terminated. After assignment of a dedicated channel (TCH or SDCCH), the mobile station starts the checking process by initializing the counter S with this value (Figure 5.23), which can be set individually per cell and is broadcast on the BCCH.

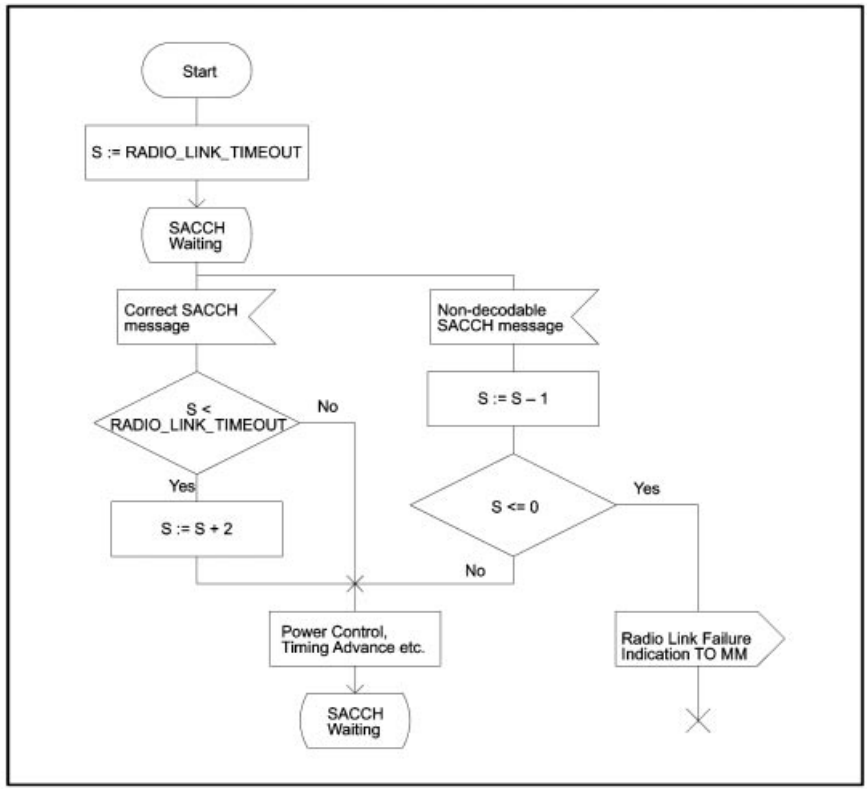

Figure 5.23: MS disconnect procedure

The corresponding checks are also conducted on the uplink. In both cases, however, this requires continuous transmission of data on the SACCH, i.e. when no signaling data has to be sent, filling data is transmitted. On the uplink, current measurement reports are transmitted, whereas the downlink carries system information of Type 5 and Type 6 (see also Section 7.4.3).

5.5.4 Cell Selection and Operation in Power Conservation Mode

5.5.4.1 Cell Selection and Cell Reselection

A mobile station in idle mode must periodically measure the receivable BCCH carriers of the base stations in the area and calculate mean values RXLEV(n) from this data (see Section 5.5.1.1). Based on these measurements, the mobile station selects a cell, namely the one with the best reception, i.e. the mobile station is committed to this cell. This is called ''camping'' on this cell. In this state, accessing a service becomes possible, and the mobile station listens periodically to the PCH. Two criteria are defined for the automatic selection of cells: the path loss criterion C1 and the reselection criterion C2. The path loss criterion serves to identify cell candidates for camping. For such cells, C1 has to be greater than zero. At least every 5 s, a mobile station has to recalculate C1 and C2 for the current and neighboring cells. If the path loss criterion of the current cell falls below zero, the path loss to the current base station has become too large. A new cell has to be selected, which requires use of the criterion C2. If one of the neighboring cells has a value of C2 greater than zero, it becomes the new current cell.

The cell selection algorithm uses two further threshold values, which are broadcast on the BCCH:

- the minimum received power level RXLEV_ACCESS_MIN (typically −98 to −106 dBm) required for registration into the network of the current cell
- the maximum allowed transmitter power MS_TXPWR_CCH (typically 31–39 dBm) allowed for transmission on a control channel (RACH) before having received the first power control command

In consideration of the maximal transmitter power P of a mobile station, the *Path Loss Criterion* C1 is now defined using the minimal threshold RXLEV_ACCESS_MIN for network access and the maximal allowed transmitter power MS_TXPWR_MAX_CCH:

$$C1(n) = (\text{RXLEV}\ (n) - \text{RXLEV\_ACCESS\_MIN} - \text{maximum}(0, (\text{MS\_TXPWR\_MAX\_CCH} - P)))$$

The values of the path loss criterion C1 are determined for each cell for which a value RXLEV(n) of a BCCH carrier can be obtained. The cell with the lowest path loss can thus be determined using this criterion. It is the cell for which $C1 > 0$ has the largest value. During cell selection, the mobile station is not allowed to enter power conservation mode (DTX, see Section 5.5.4.2).

A prerequisite for cell selection is that the cell considered belongs to the home PLMN of the mobile station or that access to the PLMN of this cell is allowed. Beyond that, a *Limited Service Mode* has been defined with restricted service access, which still allows emergency

calls if nothing else. In limited service mode, a mobile station can be camping on any cell but can only make emergency calls. Limited service mode exists when there is no SIM card in the mobile station, when the IMSI is unknown in the network or the IMEI is barred from service, but also if the cell with the best value of C1 does not belong to an allowed PLMN.

Once a mobile station is camping on a cell and is in idle mode, it should keep observing all the BCCH carriers whose frequencies, the BA, are broadcast on the current BCCH. Having left idle mode, e.g. if a TCH has been assigned, the mobile station monitors only the six strongest neighboring BCCH carriers. A list of these six strongest neighboring BCCH carriers has already been prepared and kept up to date in idle mode. The BCCH of the camped-on cell must be decoded at least every 30 s. At least once every 5 min, the complete set of data from the six strongest neighboring BCCH carriers has to be decoded, and the BSIC of each of these carriers has to be checked every 30 s. This allows the mobile station to stay aware of changes in its environment and to react appropriately. In the worst case, conditions have changed so much that a new cell to camp on needs to be selected (cell reselection).

For this cell reselection, a further criterion C2, the Reselection Criterion, has been defined:

$$C2(n) = C1(n) + \text{CELL\_RESELECT\_OFFSET} - (\text{TEMPORARY\_OFFSET} \times H(\text{PENALTY\_TIME} - T))$$

$$\text{with } H(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$$

The interval T in this criterion is the time passed since the mobile station observed the cell n for the first time with a value of $C1 > 0$. It is set back to 0 when the path loss criterion C1 falls to $C1 < 0$. The parameters CELL_RESELECT_OFFSET, TEMPORARY_OFFSET, and PENALTY_TIME are announced on the BCCH. But as a default, they are set to 0. Otherwise, the criterion C2 introduces a time hysteresis for cell reselection. It tries to ensure that the mobile station is camping on the cell with the highest probability of successful communication.

One exception for cell reselection is the case when a new cell belongs to another location area. In this case C2 must not only be larger than zero, but $C2 >$ CELL_RESELECT_HYSTERESIS to avoid too frequent location updates.

5.5.4.2 Discontinuous Reception

To limit power consumption in idle mode and thus increase battery life in standby mode, the mobile station can activate the *Discontinuous Reception* (DRX) mode. In this mode, the receiver is turned on only for the phases of receiving paging messages and is otherwise in the power conservation mode which still maintains synchronization with BCCH signals through internal timers. In this DRX mode, measurement of BCCH carriers is performed only during unused time slots of the paging blocks.

5.6 Power-up Scenario

At this point, all the functions, protocols and mechanisms of the GSM radio interface have been presented which are needed to illustrate a basic power-up scenario. The following describes the basic events that occur during a power up of the mobile station. The scenario can be divided into several steps:

- Provided a SIM card is present, immediately after turning on power, a mobile station starts the search for BCCH carriers. Normally, the station has a stored list of up to 32 carriers (Figure 5.24) of the current network. Signal level measurements are done on each of these frequencies (RXLEV). Alternatively, if no list is available, all GSM frequencies have to be measured to find potential BCCH carriers. Using the path loss criterion C1 and the threshold values stored with the list of carriers (RXLEV_ACCESS_MIN, MS_TXPWR_MAX_CCH), a first ordering can be done.

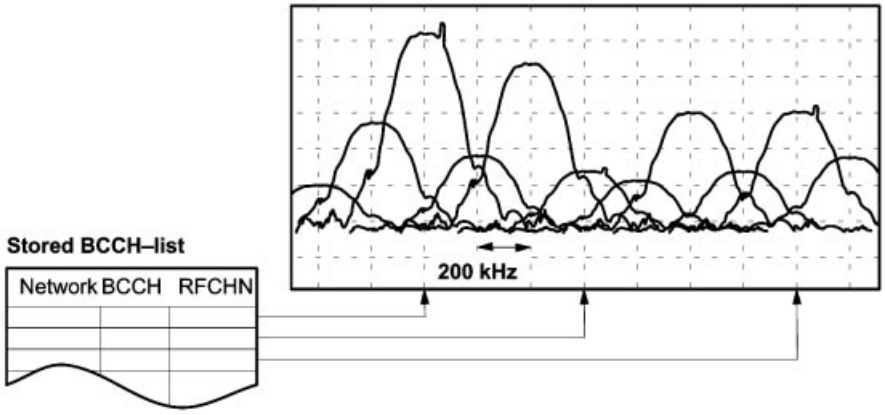

Figure 5.24: BCCH search in the power density spectrum (schematic)

- After having found potential candidates based on the received signal level RXLREV, each carrier is investigated for the presence of an FCCH signal, beginning with the strongest signal. Its presence identifies the carrier as a BCCH carrier for synchronization. Using the sine wave signal allows coarse time synchronization as well as fine tuning of the oscillator.
- The synchronization burst of the SCH in the TDMA frame immediately following the FCCH burst (Figure 5.17) has a long training sequence of 64 bits (Figure 5.6) which is used for fine tuning of the frequency correction and time synchronization. This way the mobile station is able to read and decode synchronization data from the SCH, the BSIC and the RFN. This process starts with the strongest of all BCCH carriers. If a cell is identified using BSIC and path loss criterion C1, the cell is selected for camping on it.
- The exact channel configuration of the selected cell is obtained from the BCCH data as well as the frequencies of the neighboring cells. The mobile station can now monitor the PCH of the current cell and measure the signal levels of the neighboring cells.

- The mobile station must now prepare synchronization with the six cells with the strongest signal level (RXLEV) and read out their BCCH/SCH information, i.e. steps 1–4 above are to be performed continuously for the six neighboring cells with the best RXLEV values.
- If significant changes are noticed using the path loss criterion C1 and the reselection criterion C2, the mobile station can start reselection of a new cell. Both criteria are determined periodically for the current BCCH and the six strongest neighbors.

To limit power consumption and to extend standby time of the battery, the mobile station can activate the DRX mode.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

6 Coding, Authentication, and Ciphering

The previous chapter explained the basic functions of the physical layer at the air interface, e.g. the definition of logical and physical channels, modulation, multiple access techniques, duplexing, and the definition of bursts. In this chapter, we discuss several additional functions that are performed to transmit the data in an efficient, reliable, and secure way over the radio channel: source coding and speech processing (Section 6.1), channel coding and burst mapping (Section 6.2), and security related functions, such as encryption and authentication (Section 6.3).

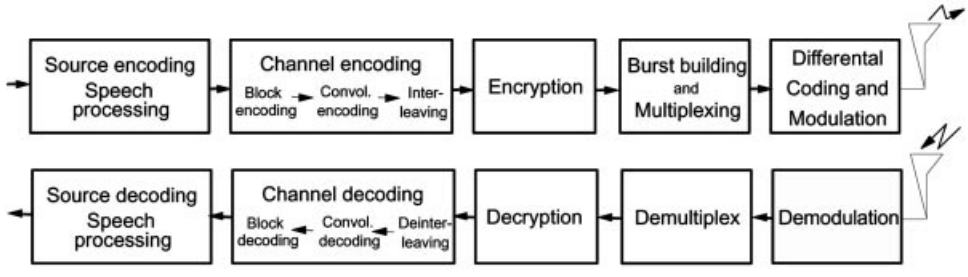

Figure 6.1: Basic elements of GSM transmission chain on the physical layer at the air interface

Figure 6.1 gives a schematic overview of the basic elements of the GSM transmission chain. The stream of sampled speech data is fed into a source encoder, which compresses the data by removing unnecessary redundancy (Section 6.1). The resulting information bit sequence is passed to the channel encoder (Section 6.2). Its purpose is to add, in a controlled manner, some redundancy to the information sequence. This redundancy serves to protect the data against the negative effects of noise and interference encountered in the transmission through the radio channel. On the receiver side, the introduced redundancy allows the channel decoder to detect and correct transmission errors. GSM uses a combination of block and convolutional coding. Moreover, an interleaving scheme is used to deal with burst errors that occur over multipath and fading channels. Next, the encoded and interleaved data is encrypted to guarantee secure and confident data transmission. The encryption technique as well as the methods for subscriber authentication and secrecy of the subscriber identity is explained in Section 6.3. The encrypted data is subsequently

mapped to bursts (Section 6.2.4), which are then multiplexed as explained in the previous chapter. Finally the stream of bits is differential coded and modulated.

After transmission, the demodulator processes the signal, which was corrupted by the noisy channel. It attempts to recover the actual signal from the received signal. The next steps are demultiplexing and decryption. The channel decoder attempts to reconstruct the original information sequence, and, as a final step, the source decoder tries to reconstruct the original source signal.

6.1 Source Coding and Speech Processing

Source coding reduces redundancy in the speech signal and thus results in signal compression, which means that a significantly lower bit rate is achieved than needed by the original speech signal. The speech coder/decoder is the central part of the GSM speech processing function, both at the transmitter (Figure 6.2) as well as at the receiver (Figure 6.3). The functions of the GSM speech coder and decoder are usually combined in one building block called the codec (COder/DECoder).

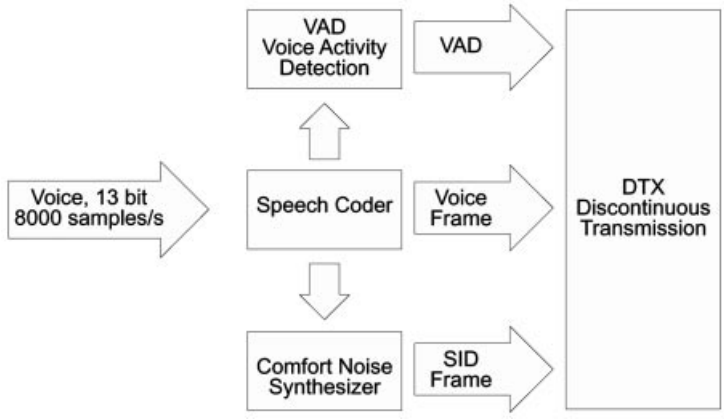

Figure 6.2: Schematic representation of speech functions at the transmitter

The analog speech signal at the transmitter is sampled at a rate of 8000 samples/s, and the samples are quantized with a resolution of 13 bits. This corresponds to a bit rate of 104 kbit/s for the speech signal. At the input to the speech codec, a speech frame containing 160 samples of 13 bits arrives every 20 ms. The speech codec compresses this speech signal into a source-coded speech signal of 260-bit blocks at a bit rate of 13 kbit/s. Thus the GSM speech coder achieves a compression ratio of 1 to 8. The source coding procedure is briefly explained in the following; detailed discussions of speech coding procedures are given in [54].

A further ingredient of speech processing at the transmitter is the recognition of speech pauses, called *Voice Activity Detection* (VAD). The voice activity detector decides, based on a set of parameters delivered by the speech coder, whether the current speech frame (20 ms) contains speech or a speech pause. This decision is used to turn off the transmitter

amplifier during speech pauses, under control of the *Discontinuous Transmission* (DTX) block.

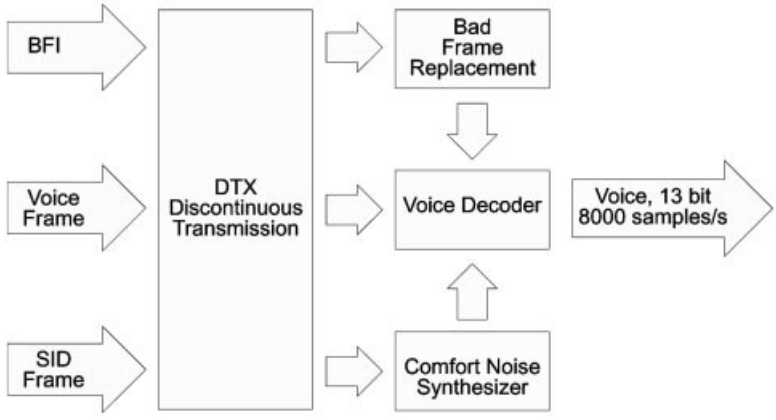

Figure 6.3: Schematic representation of speech functions at the receiver

The discontinuous transmission mode takes advantage of the fact, that during a normal telephone conversation, both parties rarely speak at the same time, and thus each directional transmission path has to transport speech data only half the time. In DTX mode, the transmitter is only activated when the current frame indeed carries speech information. This decision is based on the VAD signal of speech pause recognition. The DTX mode can reduce the power consumption and hence prolong the battery life. In addition, the reduction of transmitted energy also reduces the level of interference and thus improves the spectral efficiency of the GSM system. The missing speech frames are replaced at the receiver by a synthetic background noise signal called *Comfort Noise* (Figure 6.3). The parameters for the *Comfort Noise Synthesizer* are transmitted in a special *Silence Descriptor* (SID) frame.

This silence descriptor is generated at the transmitter from continuous measurements of the (acoustic) background noise level. It represents a speech frame which is transmitted at the end of a speech burst, i.e. at the beginning of a speech pause. In this way, the receiver recognizes the end of a speech burst and can activate the comfort noise synthesizer with the parameters received in the SID frame. The generation of this artificial background noise prevents that in DTX mode the audible background noise transmitted with normal speech bursts suddenly drops to a minimal level at a speech pause. This modulation of the background noise would have a very disturbing effect on the human listener and would significantly deteriorate the subjective speech quality. Insertion of comfort noise is a very effective countermeasure to compensate for this so-called noise-contrast effect.

Another loss of speech frames can occur, when bit errors caused by a noisy transmission channel cannot be corrected by the channel coding protection mechanism, and the block is received at the codec as a speech frame in error, which must be discarded. Such bad speech frames are flagged by the channel decoder with the *Bad Frame Indication* (BFI). In this case, the respective speech frame is discarded and the lost frame is replaced by a speech

frame which is predictively calculated from the preceding frame. This technique is called *Error Concealment*. Simple insertion of comfort noise is not allowed. If 16 consecutive speech frames are lost, the receiver is muted to acoustically signal the temporary failure of the channel.

The speech compression takes place in the speech coder. The GSM speech coder uses a procedure known as *Regular Pulse Excitation– Long-Term Prediction– Linear Predictive Coder* (RPE-LTP). This procedure belongs to the family of hybrid speech coders. This hybrid procedure transmits part of the speech signal as the amplitude of a signal envelope, a pure wave form encoding, whereas the remaining part is encoded into a set of parameters. The receiver reconstructs these signal parts through speech synthesis (vocoder technique). Examples of envelope encoding are *Pulse Code Modulation* (PCM) or *Adaptive Delta Pulse Code Modulation* (ADPCM). A pure vocoder procedure is *Linear Predictive Coding* (LPC). The GSM procedure RPE-LTP as well as *Code Excited Linear Predictive Coding* (CELP) represent mixed (hybrid) approaches [15,46,54].

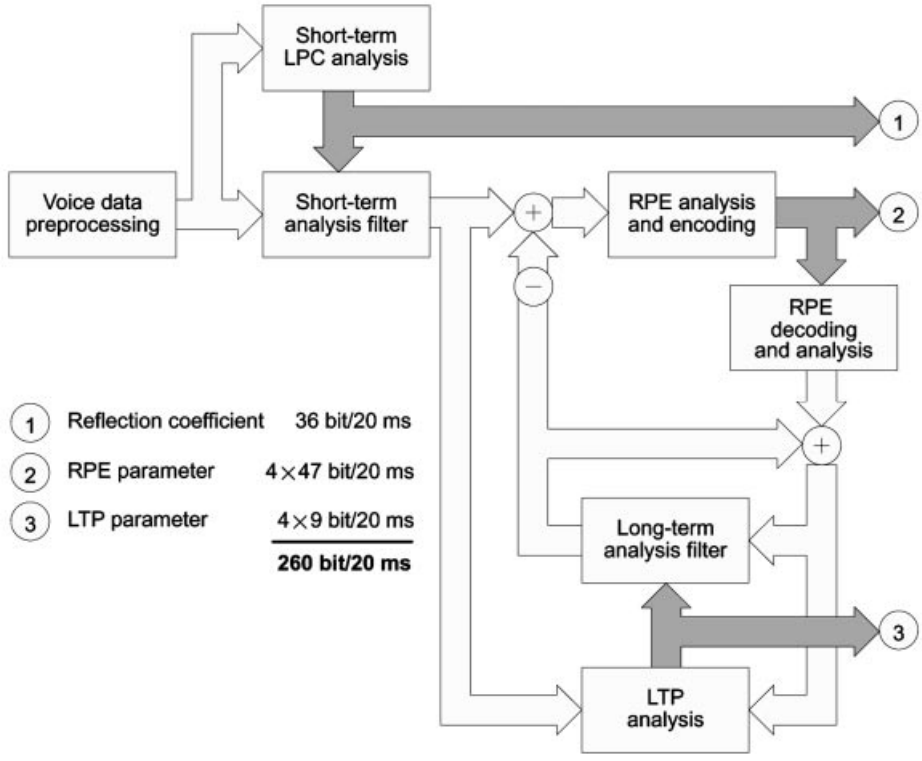

Figure 6.4: Simplified block diagram of the GSM speech coder

A simplified block diagram of the RPE-LTP coder is shown in Figure 6.4. Speech data generated with a sampling rate of 8000 samples/s and 13 bit resolution arrive in blocks of 160 samples at the input of the coder. The speech signal is then decomposed into three components: a set of parameters for the adjustment of the short-term analysis filter (LPC)

also called *reflection coefficients*; an excitation signal for the RPE part with irrelevant portions removed and highly compressed; and finally a set of parameters for the control of the LTP long-term analysis filter. The LPC and LTP analyses supply 36 filter parameters for each sample block, and the RPE coding compresses the sample block to 188 bits of RPE parameters. This results in the generation of a frame of 260 bits every 20 ms, equivalent to a 13 kbit/s GSM speech signal rate.

The speech data preprocessing of the coder (Figure 6.4) removes the DC portion of the signal if present and uses a preemphasis filter to emphasize the higher frequencies of the speech spectrum. The preprocessed speech data is run through a nonrecursive lattice filter (LPC filter, Figure 6.4) to reduce the dynamic range of the signal. Since this filter has a ''memory'' of about 1 ms, it is also called short-term prediction filter. The coefficients of this filter, called reflection coefficients, are calculated during LPC analysis and transmitted in a logarithmic representation as part of the speech frame, *Log Area Ratios* (LARs).

Further processing of the speech data is preceded by a recalculation of the coefficients of the long-term prediction filter (LTP analysis in Figure 6.4). The new prediction is based on the previous and current blocks of speech data. The resulting estimated block is finally subtracted from the block to be processed, and the resulting difference signal is passed on to the RPE coder.

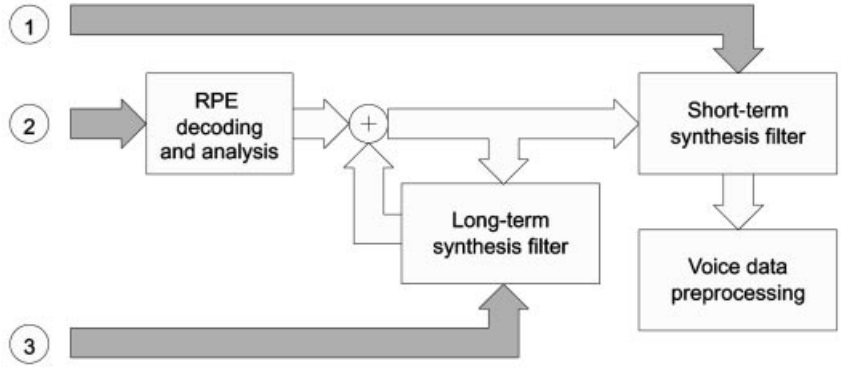

Figure 6.5: Simplified block diagram of the GSM speech decoder

After LPC and LTP filtering, the speech signal has been redundancy reduced, i.e. it already needs a lower bit rate than the sampled signal; however, the original signal can still be reconstructed from the calculated parameters. The irrelevance contained in the speech signal is reduced by the RPE coder. This irrelevance represents speech information that is not needed for the understandability of the speech signal, since it is hardly noticeable to human hearing and thus can be removed without loss of quality. On one hand, this results in a significant compression (factor $160 \times 13/188 \approx 11$); on the other hand, it has the effect that the original signal cannot be reconstructed uniquely. Figure 6.5 summarizes the reconstruction of the speech signal from RPE data, as well as the long-term and short-term synthesis from LTP and LPC filter parameters. In principle, at the receiver site, the functions performed are the inverse of the functions of the encoding process.

The irrelevance reduction only minimally affects the subjectively perceived speech qual-

ity, since the main objective of the GSM codec is not just the highest possible compression but also good subjective speech quality. To measure the speech quality in an objective manner, a series of tests were performed on a large number of candidate systems and competing codecs.

The base for comparison used is the *Mean Opinion Score* (MOS), ranging from MOS = 1, meaning quality is very bad or unacceptable, to MOS = 5, quality very good, fully acceptable. A series of coding procedures were discussed for the GSM system; they were examined in extensive hearing tests for their respective subjective speech quality [46]. Table 6.1 gives an overview of these test results; it includes as reference also ADPCM and frequency-modulated analog transmission. The GSM codec with the RPE-LTP procedure generates a speech quality with an MOS value of about 4 for a wide range of different inputs.

Table 6.1: MOS results of codec hearing tests [46]

CODEC	Process	Bit rate (in kbit/s)	MOS
FM	Frequency Modulation	–	1.95
SBC-ADPCM	Subband-CODEC – Adaptive Delta-PCM	15	2.92
SBC-APCM	Subband-CODEC – Adaptive PCM	16	3.14
MPE-LTP	Multi-Pulse Excited LPC-CODEC – Long Term Prediction	16	3.27
RPE-LPC	Regular-Pulse Excited LPC-CODEC	13	3.54
RPE-LTP	Regular Pulse Excited LPC-CODEC – Long Term Prediction	13	≈ 4
ADPCM	Adaptive Delta Modulation	32	$\geqslant 4$

6.2 Channel Coding

The heavily varying properties of the mobile radio channel (see Section 2.1) result in an often very high bit error ratio, on the order of 10^{-3} to 10^{-1}. The highly compressed, redundancy-reduced source coding makes speech communication with acceptable quality almost impossible; moreover, it makes reasonable data communication impossible. Suitable error correction procedures are therefore necessary to reduce the bit error probability into an acceptable range of about 10^{-5} to 10^{-6}. Channel coding, in contrast to source coding, adds redundancy to the data stream to enable detection and correction of transmission errors. It is the modern high-performance coding and error correction techniques which essentially enable the implementation of a digital mobile communication system.

The GSM system uses a combination of several procedures: besides a *block code*, which generates parity bits for error detection, a *convolutional code* generates the redundancy needed for error correction. Furthermore, sophisticated *interleaving* of data over several

blocks reduces the damage done by burst errors. The individual steps of channel coding are shown in Figure 6.6:

- Calculation of parity bits (block code) and addition of fill bits
- Error protection coding through convolutional coding
- Interleaving

Finally, the coded and interleaved blocks are enciphered, distributed across bursts, modulated and transmitted on the respective carrier frequencies.

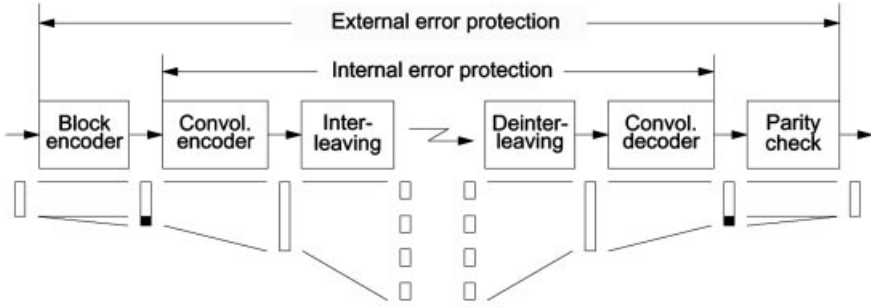

Figure 6.6: Stages of channel coding

The sequence of data blocks that arrives at the input of the channel encoder is combined into blocks, partially supplemented by parity bits (depending on the logical channel), and then complemented to a block size suitable for the convolutional encoder. This involves appending zero bits at the end of each data block, which allow a defined resetting procedure of the convolutional encoder (zero-termination) and thus a correct decoding decision. Finally, these blocks are run through the convolutional encoder. The ratio of uncoded to coded block length is called the *rate* of the convolutional code. Some of the redundancy bits generated by the convolutional encoder are deleted again for some of the logical channels. This procedure is known as *puncturing*, and the resulting code is a punctured convolutional code [3,28,38]. Puncturing increases the rate of the convolutional code, so it reduces the redundancy per block to be transmitted, and lowers the bandwidth requirements, such that the convolution-encoded signal fits into the available channel bit rate. The convolution-encoded bits are passed to the interleaver, which shuffles various bit streams. At the receiving site, the respective inverse functions are performed: deinterleaving, convolutional decoding, parity checking. Depending on the position within the transmission chain (Figure 6.6), one distinguishes between external error protection (block code) and internal protection (convolutional code).

In the following, the GSM channel coding is presented according to these stages. Section 6.2.1 explains the block coding, Section 6.2.2 deals with convolutional coding, and, finally, Section 6.2.3 presents the interleaving procedures used in GSM. The error protection measures have different parameters depending on channel and type of transported data. Table 6.2 gives an overview. (Note that the tail bits indicated in the second column are the fill bits needed by the decoding process; they should not be confused with the tail bits of the bursts (see Section 5.2).)

Table 6.2: Error protection coding and interleaving of logical channels

Channel type	Abbr.	Block distance (ms)	Bits per block			Convol. code rate	Encoded bits per block	Interleaver depth
			Data	Parity	Tail			
TCH, full rate, speech	TCH/FS	20	260				456	8
Class I			182	3	4	1/2	378	
Class II			78	0	0	–	78	
TCH, half rate, speech	TCH/HS	20	112				228	4
Class I			95	3	6	104/211	211	
Class II			17	0	0	–	17	
TCH, full rate, 14.4 kbit/s	TCH/F14.4	20	290	0	4	294/456	456	19
TCH, full rate, 9.6 kbit/s	TCH/F9.6	5	4 × 60	0	4	244/456	456	19
TCH, full rate, 4.8 kbit/s	TCH/F4.8	10	60	0	16	1/3	228	19
TCH, half rate, 4.8 kbit/s	TCH/H4.8	10	4 × 60	0	4	244/456	456	19
TCH, full rate, 2.4 kbit/s	TCH/F2.4	10	2 × 36	0	4	1/6	456	8
TCH, half rate, 2.4 kbit/s	TCH/H2.4	10	2 × 36	0	4	1/3	228	19
FACCH, full rate	FACCH/F	20	184	40	4	1/2	456	8
FACCH, half rate	FACCH/H	40	184	40	4	1/2	456	6
SDCCH, SACCH			184	40	4	1/2	456	4
BCCH, NCH, AGCH, PCH		235	184	40	4	1/2	456	4
RACH		235	8	6	4	1/2	36	1
SCH			25	10	4	1/2	78	1
CBCH		235	184	40	4	1/2	456	4

The basic unit for all coding procedures is the data block. For example, the speech coder delivers to the channel encoder a sequence of data blocks. Depending on the logical channel, the length of the data block is different; after convolutional coding at the latest, data from all channels are transformed into units of 456 bits. Such a block of 456 bits transports a complete speech frame or a protocol message in most of the signaling channels, except for the RACH and SCH channels. The starting points are the blocks delivered to the input of the channel encoder from the protocol processing in higher layers (Figure 6.7).

Speech traffic channels – One block of the full-rate speech codec consists of 260 bits of speech data, i.e. each block contains 260 information bits, which must be encoded. They are graded into two classes (Class I, 182 bits; Class II, 78 bits) which have different sensitivity against bit errors. Class I includes speech bits that have more impact on speech quality and hence must be better protected. Speech bits of Class II, however, are less

important. They are therefore transmitted without convolutional coding, but are included in the interleaving process. The individual sections of a speech frame are therefore protected to differing degrees against transmission errors (*Unequal Error Protection* (UEP)). In the case of a half-rate speech codec, data blocks of 112 information bits are input to the channel encoder. Of these, 95 bits belong to Class I and 17 bits belong to Class II. Again, one data block corresponds to one speech frame.

Data traffic channels – Blocks of traffic channels for data services have a length of N0 bits, the value of N0 being a function of the data service bit rate. We take for example the 9.6 kbit/s data service on a full-rate traffic channel (TCH/F9.6). Here, a bit stream organized in blocks of 60 information bits arrives every 5 ms at the input of the encoder. Four subsequent blocks are combined for the encoding process.

Signalling channels – The data streams of most of the signaling channels are constructed of blocks of 184 bits each; with the exception of the RACH and SCH which supply blocks of length P0 to the channel coder. The block length of 184 bits results from the fixed length of the protocol message frames of 23 octets on the signaling channels. The channel coding process maps pairs of subblocks of 57 bits onto the bursts such that it can fill a normal data burst NB (Figure 5.6).

6.2.1 External Error Protection: Block Coding

The block coding stage in GSM has the purpose of generating parity bits for a block of data, which allow the detection of errors in this block. In addition, these blocks are supplemented by fill bits (tail bits) to a block length suitable for further processing. Since block coding is the first or external stage of channel coding, the block code is also known as *external protection*. Figure 6.7 gives a brief overview showing which codes are used for which channels. In principle, only two kinds of codes are used: a *Cyclic Redundancy Check* (CRC) and a *Fire code*.

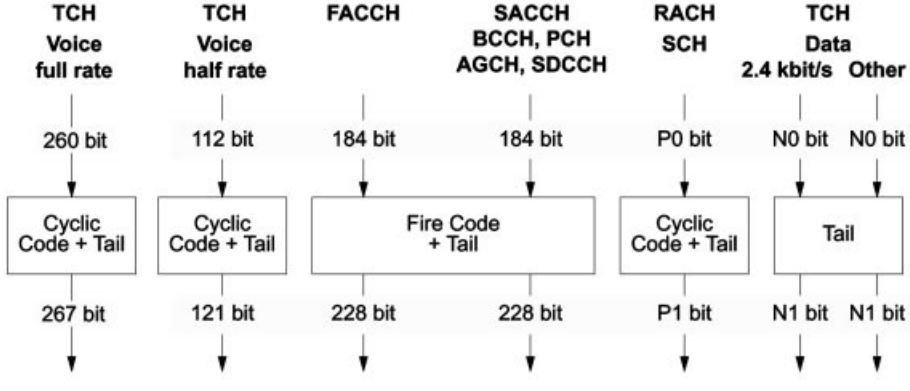

Figure 6.7: Overview of block coding for logical channels (also see Table 6.2)

6.2.1.1 Block Coding for Speech Traffic Channels

As mentioned above, speech data occurs on the TCH in speech frames (blocks) of 260 bits for TCH/F and 112 bits for TCH/H, respectively. The bits belonging to Class I are error-protected, whereas the bits of Class II and are not protected. A 3-bit Cyclic Redundancy Check (CRC) code is calculated for the first 50 bits of Class I (in the case of TCH/F). The generator polynomial for this CRC is

$$G_{CRC}(x) = x^3 + x + 1$$

In the case of a TCH/H speech channel, the most significant 22 bits of Class I are protected by 3 parity bits, using the same generator polynomial.

We now explain the block coding process in more detail with focus on the TCH/F speech codec. Since cyclic codes are easily generated with a feedback shift register, they are often defined directly with this register representation. Figure 6.8 shows such a shift register with storage locations (delay elements) and modulo-2 adders. For initialization, the register is primed with the first three bits of the data block. The other data are shifted bitwise into the feedback shift register; after the last data bit has been shifted out of the register, the register contains the check sum bits, which are then appended to the block.

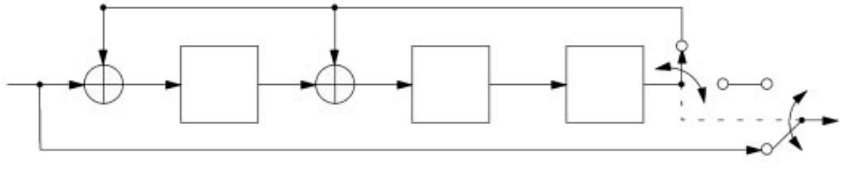

Figure 6.8: Feedback shift register for CRC

The operation of this shift register can be easily explained, if the bit sequences are also represented as polynomials like the generating function. The first 50 bits of a speech frame $D_0,D_1,\ldots,D_{49}$ are denoted as

$$D(x) = D_{49}x^{49} + D_{48}x^{48} + \cdots + D_1x + D_0$$

If this data sequence is shifted through the register of Figure 6.8, after the register was primed with D_{47}, D_{48}, D_{49} followed by 50 shift operations, then the check sum bits $R(x)$ correspond to the remainder, which is left by dividing the data sequence $x^3D(x)$ (supplemented by three zero bits) by the generator polynomial:

$$R(x) = \text{Remainder}\left[\frac{x^3D(x)}{G_{CRC}(x)}\right]$$

In the case of error-free transmission, the codeword $C'(x) = \mathrm{x}^3D(x) + R(x)$ is therefore divisible by $G_{CRC}C(x)$ without remainder. But since the check sum bits $R(x)$ are transmitted in inverted form, the division yields a remainder:

$$S(x) = \text{Remainder}\left[\frac{C(x)}{G_{CRC}(x)}\right] = \text{Remainder}\left[\frac{x^3D(x) + \bar{R}(x)}{G_{CRC}(x)}\right] = x^2 + x + 1$$

This is equivalent to shifting the whole codeword $C(x)$ through an identical shift register on the decoder side, after priming it with C_{50}, C_{51}, C_{52}. After shifting in the last check sum bit (50 shift operations), this register should contain a 1. If this is not the case, the block contains erroneous bits. Inversion of the parity bits avoids the generation of null code-words, i.e. bursts which contain only zeros cannot occur on the traffic channel.

The speech data $d(k)$ ($k = 1,\ldots,182$) of Class I of a block are combined with the parity bits $p(k)$ ($k = 1,2,3$) and fill bits to form a new block $u(k)$ ($k = 1,\ldots,189$):

$$u(k) = \begin{cases} d(2k) & k = 1,\ldots,90 \\ d(2 \times (184 - k) + 1) & k = 94,\ldots,184 \\ p((k - 91) + 1) & k = 91,92,93 \\ 0 & k = 185,\ldots,189 \end{cases}$$

The bits in even or odd positions are shifted to the upper or lower half of the block, respectively, and separated by the three check sum bits; additionally, the order of the odd bits is reversed. Finally the block is filled to 189 bits. Combination with the speech bits of Class II yields a block of 267 bits, which serves as input to the convolutional coder.

This enormous effort is taken because of the high compression rate and sensitivity against bit errors of the speech data. A speech frame in which the bits of Class I have been recognized as erroneous can therefore be reported as erroneous to the speech codec using the *Bad Frame Indication* (BFI); see Section 6.1. In order to maintain a constantly good speech quality, speech frames recognized as faulty are discarded, and the last correctly received frame is repeated, or an extrapolation of received speech data is performed.

6.2.1.2 Block Coding for Data Traffic Channels

Block coding of traffic channels is somewhat simpler for data services. In this case, no parity bits are determined. Blocks of length N0 arriving at the input of the encoder are supplemented by fill bits to a size of N1 suitable for further coding. Table 6.3 gives an overview of the different block lengths, which depend on the data rate and channel type, i.e. whether the channel is a full-rate (TCH/Fxx) or half-rate (TCH/Hxx) channel.

Table 6.3: Block formation for data traffic channels

Data channel	N0		Tail bits		N1
TCH/F14.4	290	+	4	=	294
TCH/F9.6	4 × 60	+	4	=	244
TCH/F4.8	(2 ×)60	+	(2 ×)16	=	(2 ×)76
TCH/H4.8	4 × 60	+	4	=	244
TCH/F2.4	2 × 36	+	4	=	76
TCH/H2.4	(2 ×)2 × 36	+	(2 ×)4	=	(2 ×)76

The 9.6 kbit/s data service is only offered on a full-rate traffic channel. The data comes in blocks of 60 bits to the channel encoder (every 5 ms). Four blocks each are combined and supplemented by four appended tail bits (zero bits). In the case of nontransparent data service, these four blocks make up exactly one protocol frame of the RLP protocol (240 bits). The procedures for other data services are similar. As shown in Table 6.2, for the 4.8 kbit/s and 2.4 kbit/s services, blocks of 60 or 36 bit length arrive every 10 ms. Subsequent blocks are combined and are then supplemented with tail bits (zero bits) to form blocks of 76 or 244 bits, respectively. The bit stream for the 14.4 kbit/s data service (TCH/F14.4) is offered to the encoder in blocks of 290 information bits every 20 ms. Here, four tail bits are added, resulting in 294 bits (see Table 6.3).

6.2.1.3 Block Coding for Signaling Channels

The majority of the signaling channels (SACCH, FACCH, SDCCH, BCCH, PCH, AGCH) use an extremely powerful block code for error detection. This is a so-called Fire code, i.e. a shortened binary cyclic code which appends 40 redundancy bits to the 184-bit data block. Its pure error detection capability is sufficient to let undetected errors go through only with a probability of 2^{-40}. (A Fire code can also be used for error correction, but here it is used only for error detection.) Error detection with the Fire code in the SACCH channel is used to verify connectivity (Figure 5.23), and is used, if indicated, to decide about breaking a connection. The Fire code can be defined like the CRC by way of a generator polynomial:

$$TG_F(x) = (x^{23} + 1)(x^{17} + x^3 + 1)$$

The check sum bits $R_F(x)$ of this code are calculated in such a way that a 40-bit remainder $S_F(x)$ is left after dividing the codeword $C_F(x)$ by the generator polynomial $G_F(x)$. In the case of no errors, the remainder contains only ''1'' bits:

$$S_F(x) = \text{Remainder}\left[\frac{C_F(x)}{G_F(x)}\right] = \text{Remainder}\left[\frac{x^{40}D_F(x) + R_F(x)}{G_F(x)}\right]$$

$$= x^{39} + x^{38} + \cdots + x^2 + x + 1$$

The codeword generated with the redundancy bits of the Fire code is supplemented with ''0'' bits to a total length of 228 bits, which are then delivered to the convolutional coder.

Another approach has been used for error detection in the RACH channel. The very short random access burst in the RACH allows only a data block length of P0 = 8 bits, which is supplemented in a cyclic code by six redundancy bits. The corresponding generator polynomial is

$$G_{RACH}(x) = x^6 + x^5 + x^3 + x^2 + x + 1$$

In the *Access Burst* (AB), the mobile station also has to indicate a target base station. The BSIC of the respective base station is used for this purpose. The six bits of the BSIC are added to the six redundancy bits modulo 2, and the resulting sequence is inserted as the redundancy of the data block. The total codeword to be convolution-coded for the RACH

thus has a length of 18 bits; i.e. four fill bits (''0'') are also added in the RACH to this block. In exactly the same way, block coding is performed for the *Handover Access* burst, which is in principle also a random access burst.

The SCH channel, as an important synchronization channel, uses a somewhat more elaborate error protection than the RACH channel. The SCH data blocks have a length of 25 bits and receive, besides the fill bits, another 10 bits of redundancy for error detection through a cyclic code with somewhat better error detection capability than on the RACH:

$$G_{SCH}(x) = x^{10} + x^8 + x^6 + x^5 + x^4 + x^2 + 1$$

Thus the length of the codewords delivered to the channel coder in the SCH channel is 39 bits. Table 6.4 summarizes the block parameters of the RACH and SCH channels. Table 6.5 presents an overview of the cyclic codes used in GSM.

Table 6.4: Block lengths for the RACH and SCH channels

Data channel	P0		Parity bits		Tail bits		P1
RACH	8	+	6	+	4	=	18
SCH	25	+	10	+	4	=	39

Table 6.5: Cyclic codes used for block coding in GSM

Channel	Polynomial
TCH/FS	$x^3 + x + 1$
DCCH and CCCH (part.)	$(x^{23} + 1)(x^{17} + x^3 + 1)$
RACH	$x^6 + x^5 + x^3 + x^2 + x + 1$
SCH	$x^{10} + x^8 + x^6 + x^5 + x^4 + x^2 + x + 1$

6.2.2 Internal Error Protection: Convolutional Coding

After block coding has supplemented the data with redundancy bits for error detection (parity bits), added fill bits and thus generated sorted blocks, the next stage is calculation of additional redundancy for error correction to correct the transmission errors caused by the radio channel. The internal error correction of GSM is based exclusively on convolutional codes.

Convolutional codes [35] can also be defined using shift registers and generator polynomials. Figure 6.9 illustrates a possible convolutional encoder realization. It basically consists of a shift register with modulo-2 adders and K storage locations (here $K = 4$). One data/information symbol d_i is read into the shift register per tact interval. A symbol consists of k (here $k = 1$) data/information bits, each of which is moved into the shift register. A data symbol could also consist of more than one bit ($k > 1$), but this is not

implemented in GSM. The symbol read is combined with up to K of its predecessor symbols $d_{i-1},\ldots,d_{i-K}$ in several modulo 2 additions. The results of these operations are given to the interleaver as coded user payload symbols c_j. The value K determines the number of predecessor symbols to be combined with a data symbol and is therefore also called the *memory* of the convolutional encoder. The number v of combinatorial rules (here $v = 2$) determines the number of coded bits in a code symbol c_j generated for each input symbol d_i. In Figure 6.9, the combinatorial results are scanned from top to bottom to generate the code symbol c_j. The combinatorial rules are defined by the generator polynomial $G_i(d)$. It is important to note that a specific convolutional code can be generated by various encoders. Thus, it must be carefully distinguished between code properties and encoder properties.

As mentioned in Section 6.2.1, block coding appends at least four zero bits to each block. These bits not only serve as fill bits at the end of a block, but they are also important for the channel coding procedure. Shifted at the end of each block into the encoder, these bits serve to reset the encoder into the defined starting position (zero-termination of the encoder), such that in principle adjacent data blocks can be coded independently of each other.

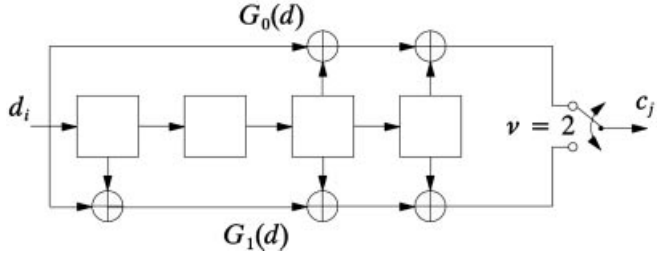

Figure 6.9: Principle of a convolutional encoder

The rate r of a convolutional code indicates how many data (information) bits are processed for each coded bit. Consequently, $1/r$ is the number of coded bits per information bit. This rate is the essential measure of the redundancy produced by the code and hence its error correction capability:

$$r = k/v, \qquad \text{here}: \; r = 1/v = \frac{1}{2}$$

The code rate is therefore determined by the number of bits k per input data symbol and the number of combinatorial rules v which are used for the calculation of a code symbol. In combination with the memory K, the code rate r determines the error correction capability of the code. In a simplified way: with decreasing r and increasing K, the number of corrigible errors per codeword increases, and, thus, the error correction capabilities of the code are improved. The encoding procedure is expressed in the combinatorial operations (modulo 2 additions). These coding rules can be described with polynomials. In the case of the convolutional encoder of Figure 6.9, the two generator polynomials are

$$G_0(d) = d^4 + d^3 + 1$$

$$G_1(d) = d^4 + d^3 + d + 1$$

They give a compact representation of the encoding procedure. The maximal exponent of a generator polynomial is known as its *constraint length*. The maximal of all constraint lengths (i.e. the maximal exponent of all polynomials) defines the memory K of the convolutional encoder. The number of polynomials determines the rate r. The exponents represent how an input symbol d_i processed in the encoder. For example, in the upper path of the encoder (represented by $G_0(d)$), an input symbol d_i is immediately forwarded to the output (exponent ''0''), and it is processed again in the third and fourth tact interval.

GSM defines different convolutional codes for the different logical channels (see Figure 6.10). Table 6.6 lists the seven generator polynomials (G_0,…,G_6) used in different combinations. The convolutional encoder used for half-rate speech channels (TCH/HS) has memory 6. All other encoders have memory 4, but they differ in the code rate and the polynomials used.

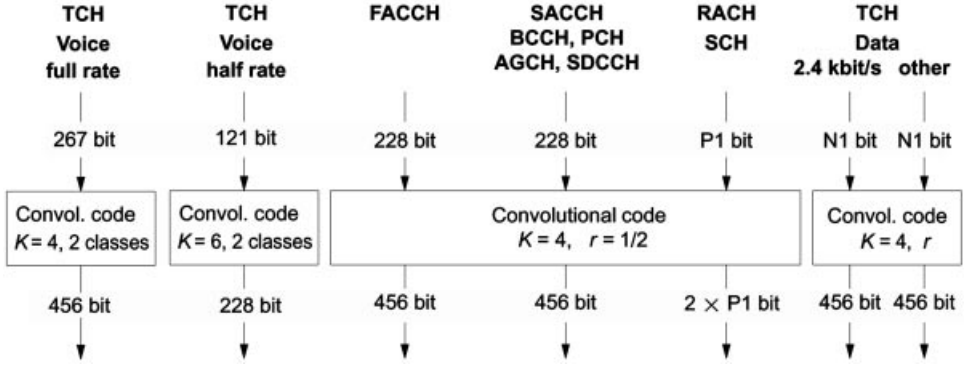

Figure 6.10: Overview of convolutional coding of logical channels (continued from Figure 6.7; also see Table 6.2)

Table 6.6: Generator polynomials for convolutional codes

Type	Polynomial
G0	$1 + d^3 + d^4$
G1	$1 + d + d^3 + d^4$
G2	$1 + d^2 + d^4$
G3	$1 + d + d^2 + d^3 + d^4$
G4	$1 + d^2 + d^3 + d^5 + d^6$
G5	$1 + d + d^4 + d^6$
G6	$1 + d + d^2 + d^3 + d^4 + d^6$

Table 6.7 gives an overview of the uses and combinations of generator polynomials. Most logical channels use a convolutional code of rate 1/2 based on polynomials G0 and G1.

Speech traffic channels – Convolutional coding of Class I speech bits on the full-rate speech channel generates $1/r \times (182 + 3 + 4)$ bits $= 378$ bits. The 78 bits of Class II are

not encoded at all, which results in a total number of 456 bits. This is the uniform block size needed for mapping these data blocks onto the bursts with 114-bit payload. In a similar way, for most of the remaining channels, two coded blocks of 228 bits are combined.

In the case of a half-rate speech traffic channel (TCH/HS), the Class I bits are encoded using a punctured version of a rate-1/3 convolutional encoder defined by G4, G5, and G6. Including puncturing, the net rate of the encoder is $r' = 104/211 \approx 1/2$. The 95 information bits, 3 parity bits, and 6 tail bits are thus mapped to 211 bits. The 17 Class II bits are not convolutional encoded, which results in a total number of 228 bits.

Data traffic channels – The 4.8 kbit/s data service on a full-rate channel (TCH/F4.8) and the 2.4 kbit/s data service on a half-rate channel (TCH/H2.4) use a code of rate 1/3 based on polynomials G1, G2, and G3. The 2.4 kbit/s data service on a full-rate channel (TCH/F2.4) uses these three polynomials twice in a row to generate a convolutional code of rate 1/6.

The data on a TCH/F9.6 and TCH/H4.8 is encoded using the rate-1/2 code defined by the polynomials G0 and G1. At the input of the convolution encoder, blocks of 244 bits arrive which the encoder maps to blocks of 488 bits. These blocks are reduced to 456 bits by removing (puncturing) every 15th bit beginning with the 11th bit, i.e. a total of 32 bits are punctured. On the one hand, puncturing cuts the block size to a length suitable for further processing; on the other hand, puncturing removes redundancy. The resulting net code rate of $r' = 244/456$ is therefore somewhat higher than the rate of 1/2 for the

Table 6.7: Usage of generator polynomials

Channel type	generator polynomial						
	G0	G1	G2	G3	G4	G5	G6
TCH, full rate, speech							
Class I	■	■					
Class II							
TCH, half rate, speech							
Class I					■	■	■
Class II							
TCH, full rate, 14.4 kbit/s	■	■					
TCH, full rate, 9.6 kbit/s	■	■					
TCH, full rate, 4.8 kbit/s		■	■	■			
TCH, half rate, 4.8 kbit/s	■	■					
TCH, full rate, 2.4 kbit/s		■	■	■			
TCH, half rate, 2.4 kbit/s		■	■	■			
FACCHs	■	■					
SDCCHs, SACCHs	■	■					
BCCH, AGCH, PCH	■	■					
RACH	■	■					
SCH	■	■					

convolutional encoder. Thus, the code has slightly lower error correction capability. Puncturing cuts down the convolutional coded blocks of the TCH/F9.6 and TCH/H4.8 channels to the standard format of 456 bits. Thus, blocks of these channels can also be processed in a standardized way (interleaving, etc.), and the amount of redundancy contained in a block is also matched to the bit rate available for transmission.

For the encoding of the TCH/F14.4, again a punctured version of the (G0, G1) convolutional encoder is employed. The 294 bits at the input of the encoder are mapped to 588 bits, followed by a puncturing of 132 bits.

Convolutional decoding – In most cases, the decoding of convolutional code employs the Viterbi algorithm. It uses a suitable metric to determine the data sequence that most likely equals the transmitted data (*maximum likelihood decoding*) [9]. Using the knowledge of the generator polynomials, the decoder can determine the original data sequence.

6.2.3 Interleaving

The decoding result of the convolutional code strongly depends on the frequency and grouping of bit errors that occur during transmission. Especially burst errors during long and deep fading periods, i.e. a series of erroneous sequential bits, have negative impact on error correction. In such cases, the channel is not a binary channel without memory, rather the single-bit errors have statistical dependence, which diminishes the result of the error correction procedure of the convolutional code. To achieve good error correction results, the channel should have no memory, i.e. the bit errors should be statistically independent. Therefore, burst errors occurring frequently on the radio channel should be distributed uniformly across the transmitted codewords. This can be accomplished through the interleaving technique described in the following.

The interleaving approach is to distribute codewords from the convolutional encoder by spreading in time and merging them across several bursts for transmission. This principle is shown in Figure 6.11. By time spreading, each of the codewords is distributed across a

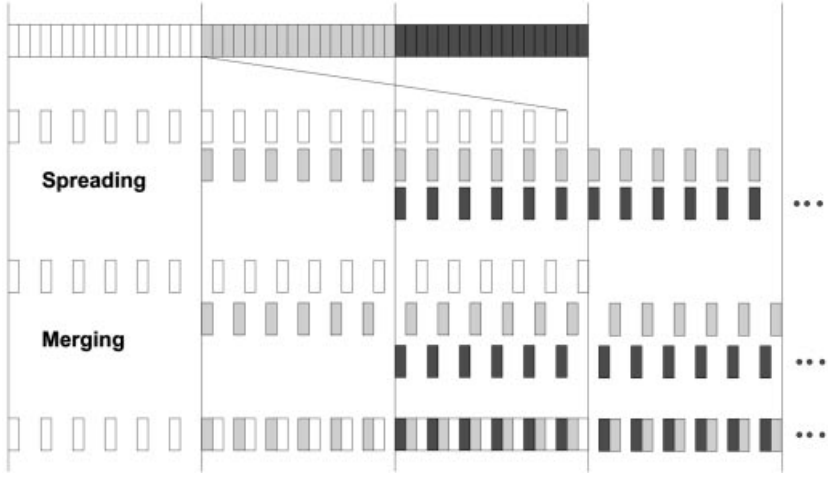

Figure 6.11: Interleaving: spreading and merging

threefold length. Merging the bit sequences generated in this way has the effect that the individual bits from each of the three codewords are sorted into alternate bursts; this way each codeword is transmitted as distributed over a total of three bursts, and two bits of a data block are never transmitted adjacent to each other.

This kind of interleaving is also known as *diagonal interleaving*. The number of bursts over which a codeword is spread is called the *interleaving depth*; a spreading factor can be defined analogously. A burst error is therefore distributed uniformly over several subsequently transmitted codewords because of the distribution of the data over several bursts. This generates bit error sequences which are less dependently distributed in the data stream, hence it improves the success of the error correction process.

Figure 6.12 shows an example. During the third burst of transmission, severe fading of the signal leads to a massive burst error. This burst is now heavily affected by a total of six single-bit errors. In the process of deinterleaving (inversion of merging, despreading) these bit errors are distributed across three data blocks, corresponding to the bit positions which were sorted into the respective bursts during interleaving. The number of errors per data block is now only two, which can be much more easily corrected.

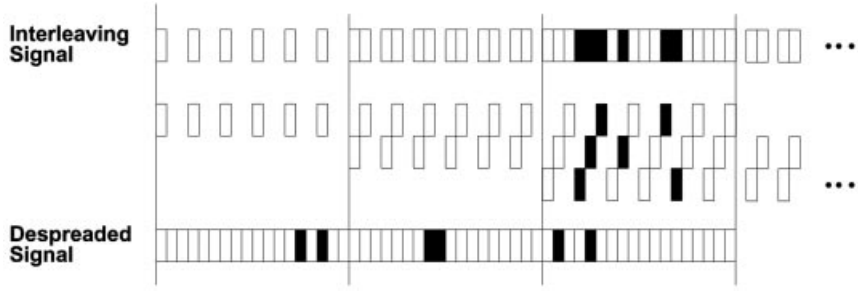

Figure 6.12: Distributing bit errors through deinterleaving

Another kind of interleaving is *block interleaving*. In this principle, codewords are written line by line into a matrix (Figure 6.13), which is subsequently read out column by column. The number of lines of the interleaving matrix determines the interleaving depth. As long as the length of a burst error is shorter than the interleaving depth, the burst error generates only single-bit errors per codeword if block interleaving is used [9,54].

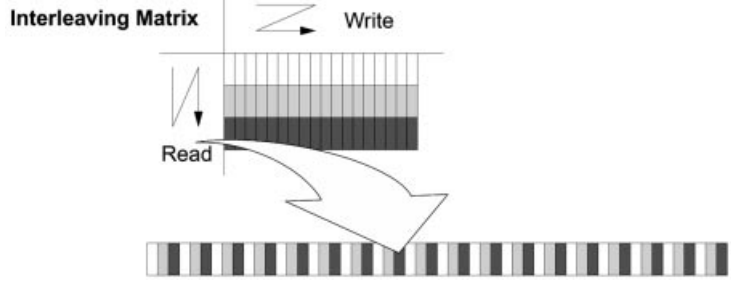

Figure 6.13: Principle of block interleaving

However, the great advantage of interleaving, to alleviate the effect of burst errors for optimal error correction with a convolutional code, is traded for a not insignificant disadvantage for speech and data communication. As evident from Figures 6.11 and 6.13, the bits of a codeword are spread across several bursts (here: three). For a complete reconstruction of a codeword, one has to wait for the complete transmission of three bursts. This forces a transmission delay, which is a function of the interleaving depth.

In GSM, both methods of interleaving are used (Figure 6.14), blockwise as well as bitwise. With a maximal interleaving depth of 19, this can lead to delays of up to 360 ms (Table 6.8).

Full-rate speech channel, TCH/F2.4, and FACCH – The speech channel TCH/FS in GSM uses block-diagonal interleaving. The 456 bits of a codeword are distributed across eight interleaving blocks, where one interleaving block has 114 bit positions. The exact interleaving rule for mapping the coded bits $c(n, k = 0, \ldots, 455)$ of the nth codeword, onto bit position $i(b,j = 0,\ldots,114)$ of the bth interleaving block, is

$$i(b, j) = c(n, k)$$

$$\text{with} \begin{cases} n = 0, 1, 2, \ldots, N, N + 1, \ldots \\ k = 0, 1, 2, \ldots, 455 \\ b = b_0 + 4n + (k \bmod 8) \\ j = 2((49k) \bmod 57)) + ((k \bmod 8) \operatorname{div} 4) \end{cases}$$

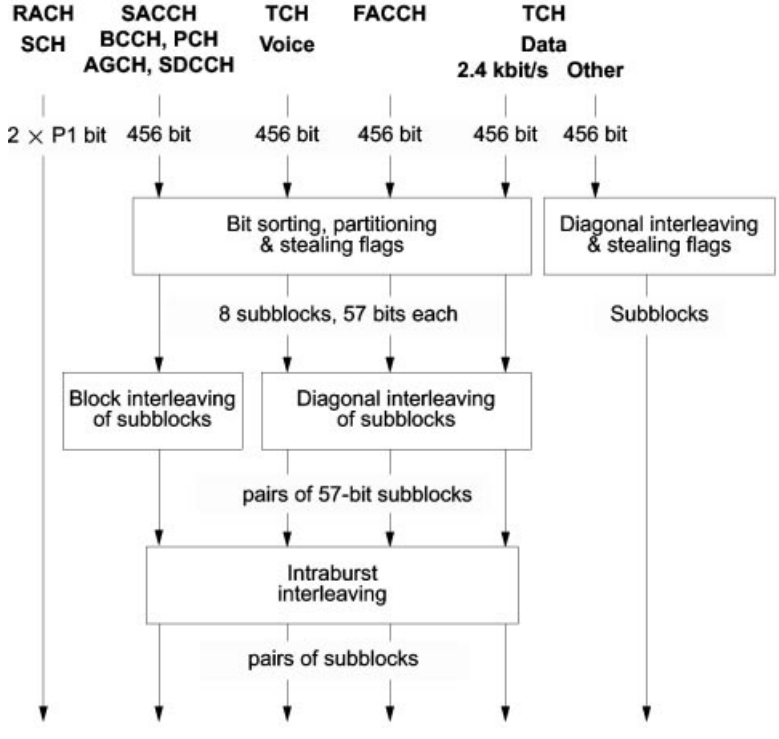

Figure 6.14: Overview: interleaving of (full-rate) logical channels

The bits of the *n*th codeword (data block *n* in Figure 6.15) are distributed across eight interleaving blocks, beginning with block $B = b_0 + 4n$. To do so, the coded bits are mapped to the even bits of the first four interleaving blocks $(B + 0, \ldots, B + 3)$ and to the odd bits of the other four interleaving blocks $(B + 4, \ldots, B + 7)$. The even bits of the last four interleaving blocks $(B + 4, \ldots, B + 7)$ are occupied by data from codeword $n + 1$. Each interleaving block thus contains 57 bits of the current codeword *n* and 57 bits of the following codeword $n + 1$ or the preceding codeword $n - 1$, respectively. In this way, a new codeword is started after each fourth merged interleaving block.

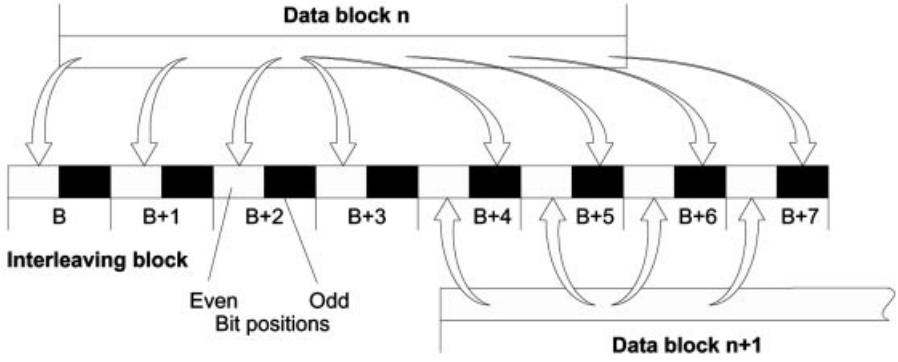

Figure 6.15: Interleaving TCH/FS: block mapping

The individual bits of codeword *n* are alternatively distributed across the interleaving blocks, e.g. every eighth bit is in the same interleaving block according to the term (*k* mod 8), whereas bit position *j* within an interleaving block $b = B + 0, B + 1, \ldots, B + 7$ is determined by two terms: the term (*k* mod 8) div 4 is used to determine the even/odd bit positions; and the term 2 ((49*k*) mod 57) determines the offset within the interleaving block. The first interleaving block *B* derived from codeword n thus contains bit numbers 0,8,16,…,448,456 of this codeword.

The placement of these bits for the first block *B* in the interleaving block is illustrated in Figure 6.16. This placement is chosen in such a way that no two directly adjacent bits of the interleaving block belong to the same codeword. In addition, the mapped bits are

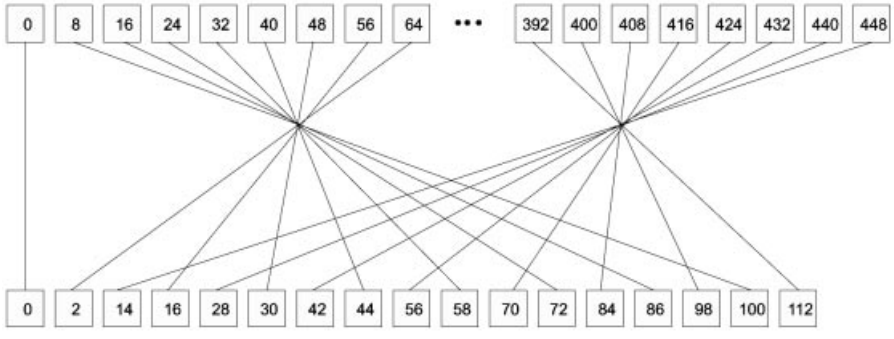

Figure 6.16: Mapping of codeword *n* onto interleaving block *B* for a TCH/FS

combined into groups of eight bits each, which are distributed as uniformly as possible across the entire interleaving block. This achieves additional spreading of error bursts within a data block. Therefore, the interleaving for the TCH/FS is block-diagonal interleaving with additional merging of data bits within the interleaving block. This is also called *intraburst interleaving* (Figure 6.14). The data channel TCH/F2.4 and the FACCH in GSM use the same interleaving methods as the TCH/FS.

Table 6.8: Transmission delay caused by interleaving

Channel type	Interleaving depth	Transmission delay (ms)
TCH, full-rate, voice	8	38
TCH, half-rate, voice	4	
TCH, full-rate, 14.4 kbit/s	19	93
TCH, full-rate, 9.6 kbit/s	19	93
TCH, full-rate, 4.8 kbit/s	19	93
TCH, half-rate, 4.8 kbit/s	19	185
TCH, full-rate, 2.4 kbit/s	8	38
TCH, half-rate, 2.4 kbit/s	19	185
FACCH, full-rate	8	38
FACCH, half-rate	8	74
SDCCH	4	14
SACCH/TCH	4	360
SACCH/SDCCH	4	14
BCCH, AGCH, PCH	4	14

Other data traffic channels – For the other data services in the traffic channel (TCH/F14.4, TCH/F9.6, TCH/F4.8, TCH/H4.8, and TCH/H2.4) the interleaving is somewhat simpler. A pure bitwise diagonal interleaving with an interleaving depth of 19 is used. In this case, the interleaving rule is

$$i(b,j) = c(n,k)$$

$$\text{with} \begin{cases} n = 0,1,2,\ldots,N,N+1,\ldots \\ k = 0,1,2,\ldots,455 \\ b = b_0 + 4n + (k \bmod 19) + k \\ j = k \bmod 19 + 19(k \bmod 6) \operatorname{div} 114 \end{cases}$$

The bits of a data block (n,k) are distributed in groups of 114 bits across 19 interleaving blocks, whereby groups of six bits are distributed uniformly over one interleaving block.

With this diagonal interleaving, each interleaving block also starts a new 114-bit block of data. A closer look at this interleaving rule reveals that the input to the interleaver consists of blocks of 456 coded data bits as codewords. The whole codeword is therefore really spread across 22 interleaving blocks; the nominal interleaving depth of 19 results historically from 114-bit block interleaving.

Half-rate speech channel – The interleaving rule for the half-rate speech channel (TCH/HS) is given by

$$i(b,j) = c(n,k)$$

$$\text{with} \begin{cases} n = 0,1,2,\ldots,N,N+1,\ldots \\ k = 0,1,2,\ldots,227 \\ b = b_0 + 2n + (k \bmod 4) \end{cases}$$

and j according to a table in the GSM standard. The 228 bits of a codeword n are distributed over 4 blocks. Beginning with interleaving block $B = b_0 + 2n$, it occupies the even numbered bits of the first two interleaving blocks ($B + 0$, $B + 1$) and the odd numbered bits of the other two blocks ($B + 2$, $B + 3$). Consequently, the following codeword $n + 1$ uses the even numbered bits of the blocks $B + 2$ ($= b_0 + 2(n + 1) + 0$) and $B + 3 (= b_0 + 2(n + 1) + 1)$ as well as the odd numbered bits of the interleaving blocks $b_0 + 2(n + 1) + 2$ and $b_0 + 2(n + 1) + 3$. As with the TCH/FS, one interleaving block contains 57 bits from codeword n and 57 bits from codeword $n + 1$ or $n - 1$. In summary, a new codeword starts every second interleaving block.

Signalling channels – Most signaling channels use an interleaving depth of 4, such as SACCH, BCCH, PCH, AGCH, and SDCCH. The interleaving scheme is almost identical to the one used for the TCH/FS, however, the codewords $c(n,k)$ are spread across four rather than eight interleaving blocks:

$$i(b,j) = c(n,k)$$

$$\text{with} \begin{cases} n = 0,1,2,\ldots,N,N+1,\ldots \\ k = 0,1,2,\ldots,455 \\ b = b_0 + 4n + (k \bmod 4) \\ j = 2((49k) \bmod 57)) + ((k \bmod 8) \operatorname{div} 4) \end{cases}$$

With this kind of interleaving, there are also eight blocks generated, just like in the case of the TCH/FS, however, at the same time a block of 57 even bits is combined with a block of 57 odd bits to form a complete interleaving block. This has the consequence that consecutive coded signaling messages are not block-diagonally interleaved, but that each four consecutive interleaving blocks are fully occupied with the data of just one, and only one, codeword. Also, a new codeword starts after every four interleaving blocks. Therefore, this interleaving of GSM signaling messages is in essence also a block interleaving procedure. This is especially important for signaling channels to ensure the transmission of individual protocol messages independent of preceding or succeeding messages. This also enables

some kind of asynchronous communication of signaling information. The signaling data of the RACH and SCH must each be transmitted in single data bursts; no interleaving occurs.

6.2.4 Mapping onto the Burst Plane

After block encoding, convolutional encoding, and interleaving, the data are available in form of 114-bit interleaving blocks. This corresponds exactly to the amount of data which can be carried by a normal burst (Figure 5.6). Each interleaving block is mapped directly onto one burst (Figure 6.17). After setting the stealing flags, the bursts can be composed and passed to the modulator. The stealing flags indicate whether high-priority signaling messages are present (FACCH messages), which must be transmitted as fast as possible, instead of the originally planned data of the traffic channel.

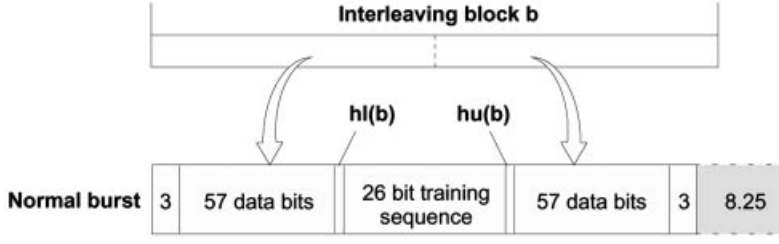

Figure 6.17: Mapping onto a burst

An essential component of GSM channel coding is the correct treatment of FACCH signaling messages which are multiplexed in a pre-emptive way into the traffic channel. At the burst level, each FACCH codeword displaces a codeword of the current TCH/FS traffic channel, i.e. the codewords must be tied into the interleaving structure instead of regular data blocks of the traffic channel. The interleaving rule for the FACCH is the same as for the TCH/FS: from an FACCH codeword the even positions are occupied in one set of four interleaving blocks, and the odd bit positions are occupied in another set of interleaving blocks; in addition, the bit positions within the interleaving blocks are shuffled (intra-burst interleaving).

When an FACCH message needs to be transmitted (e.g. a handover command), the current data block n is replaced by the convolutional coded FACCH message, and it is interleaved in a block-diagonal way with the data blocks $(n - 1)$ and $(n + 1)$ of the traffic channel (Figure 6.18). In the eight blocks involved in this procedure, $B, B + 1, \ldots, B + 7$, the respective stealing flags hl(b) and hu(b) have to be set (Figure 6.17). If neither flag is set, the burst contains data of the traffic channel. If the even bits of the burst are occupied by FACCH data, hu(b) is set; in the case of the odd bits being used for FACCH, hl(b) is set (Figure 6.18).

If the current burst is not available for traffic channel data, the data block n has to be discarded. Bits ‘‘stolen’’ in this way have varying effects: For example:

- A complete speech frame of the TCH/FS is lost (20 ms speech).

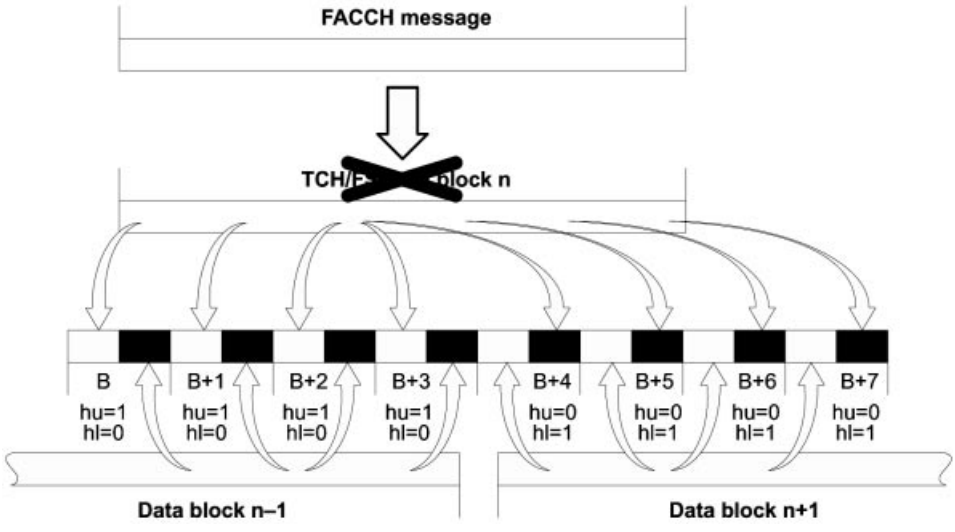

Figure 6.18: Insertion of an FACCH message into the TCH/FS data stream

- In the case of TCH/F9.6 and TCH/H4.8 channels, three bits are stolen from each of the eight interleaving blocks, which belong to the same data block, such that a maximum of 24 coded data bits are interfered with.
- In the case of TCH/F4.8 and TCH/H2.4 channels, six bits are stolen from each of the eight interleaving blocks, which belong to the same data block, such that a maximum of 48 coded data bits are interfered with.
- In the case of TCH/F2.4 channels, the same interleaving rules as in the TCH/FS are used, such that a complete data block is displaced by the FACCH.

In summary, the FACCH signaling needed for fast reactions causes data losses or bit errors in the accompanying traffic channel, and they have to be totally or partially corrected by the convolutional code.

6.3 Security-Related Network Functions and Encryption

Methods of encryption for user data and for the authentication of subscribers, like all techniques for data security and data protection, are gaining enormous importance in modern digital systems [17]. GSM therefore introduced powerful algorithms and encryption techniques. The various services and functions concerned with security in a GSM PLMN are categorized in the following way:

- Subscriber identity confidentiality
- Subscriber identity authentication
- Signalling information element confidentiality
- Data confidentiality for physical connections

In the following, the security functions concerning the subscriber are presented.

6.3.1 Protection of Subscriber Identity

The intent of this function is to prevent disclosing which subscriber is using which resources in the network, by listening to the signaling traffic on the radio channel. On one hand this should ensure the confidentiality of user data and signaling traffic, on the other hand it should also prevent localizing and tracking of a mobile station. This means above all that the *International Mobile Subscriber Identity* (IMSI) should not be transmitted as clear text, i.e. unencrypted.

Instead of the IMSI, one uses a *Temporary Mobile Subscriber Identity* (TMSI) on the radio channel for identification of subscribers. The TMSI is temporary and has only local validity, which means that a subscriber can only be uniquely identified by TMSI and the *Location Area ID* (LAI). The association between IMSI and TMSI is stored in the VLR.

The TMSI is issued by the VLR, at the latest, when the mobile station changes from one *Location Area* (LA) into another (location updating). When a new location area is entered, this is noticed by the mobile station (Section 3.2.5) which reports to the new VLR with the old LAI and TMSI (LAI*old* and TMSI*old*, Figure 6.19). The VLR then issues a new TMSI for the MS. This TMSI is transmitted in encrypted form.

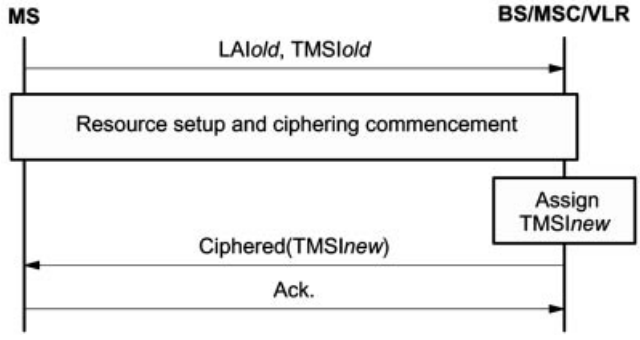

Figure 6.19: Encrypted transmission of the temporary subscriber identity

The subscriber identity is thus protected against eavesdropping in two ways: first, the temporary TMSI is used on the radio channel instead of the IMSI; second, each new TMSI is transmitted in encrypted form.

In the case of database failures, if the VLR database is partially lost or no correct subscriber data is available (loss of TMSI, TMSI unknown at VLR, etc.), the GSM standard provides for a positive acknowledgement of the subscriber identity. For this subscriber identification, the IMSI must be transmitted as clear text (Figure 6.20) before encryption is turned on. Once the IMSI is known, encryption can be restarted and a new TMSI can be assigned.

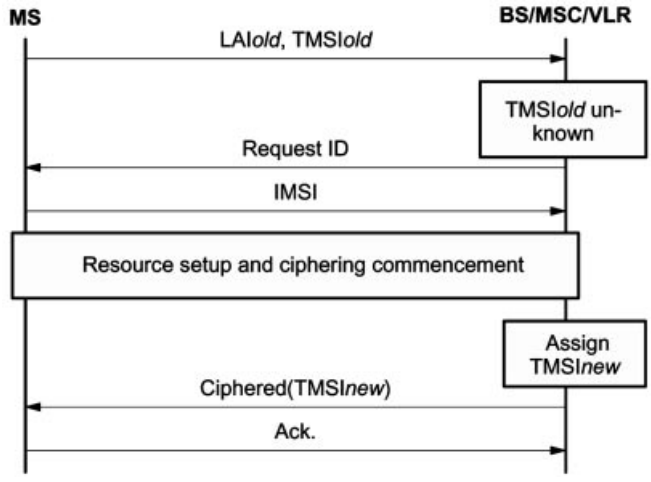

Figure 6.20: Clear text transmission of the IMSI when the TMSI is unknown

6.3.2 Verification of Subscriber Identity

When a subscriber is added to a home network for the first time, a *Subscriber Authentication Key* (Ki) is assigned in addition to the IMSI to enable the verification of the subscriber identity (also known as authentication). All security functions are based on the secrecy of this key. At the network side, the key Ki is stored in the *Authentication Center* (AUC) of the home PLMN. At the subscriber side, it is stored on the SIM card of the subscriber.

The process of authenticating a subscriber is essentially based on the A3 algorithm, which is performed at the network side as well as at the subscriber side (Figure 6.21). This algorithm calculates independently on both sides (MS and network) the *Signature Response* (SRES) from the authentication key Ki and a *Random Number* (RAND) offered by the network. The MS transmits its SRES value to the network which compares it with its calculated value. If both values agree, the authentication was successful. Each execution of the algorithm A3 is performed with a new value of the random number RAND which

Figure 6.21: Principle of subscriber authentication

cannot be predetermined; in this way recording the channel transmission and playing it back cannot be used to fake an identity.

6.3.3 Generating Security Data

At the network side, the 2-tuple (RAND, SRES) need not be calculated each time when authentication has to be done. Rather the AUC can calculate a set of (RAND, SRES) 2-tuples in advance, store them in the HLR, and send them on demand to the requesting VLR. The VLR stores this set (RAND[n], SRES[n]) and uses a new 2-tuple from this set for each authentication procedure. Each 2-tuple is used only once; so new 2-tuples continue to be requested from the HLR/AUC.

This procedure, to let security data (Kc, RAND, SRES) be calculated in advance by the AUC has the advantage that the secret authentication key Ki of a subscriber can be kept exclusively within the AUC, which ensures a higher level of confidentiality. A somewhat less secure variant is to supply the currently needed key Ki to the local VLR which then generates the security data locally.

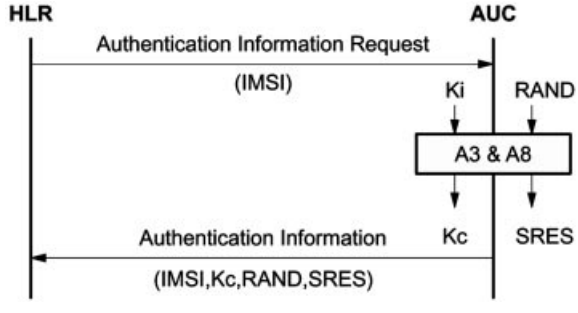

Figure 6.22: Generation of a set of security data for the HLR

If the key Ki is kept exclusively in the AUC, the AUC has to generate a set of security data for a specific IMSI on demand from the HLR (Figure 6.22): the random number RAND is generated and the pertinent signature SRES is calculated with the A3 algorithm, whereas the A8 algorithm generates the encryption key Kc.

The set of security data, a 3-tuple consisting of Kc, RAND, and SRES, is sent to the HLR and stored there. In most cases, the HLR keeps a supply of security data (e.g. 5), which can then be transmitted to the local VLR, so that one does not have to wait for the AUC to generate and transmit a new key. When there is a change of LA into one belonging to a new VLR, the sets of security data can be passed on to the new VLR. This ensures that the subscriber identity IMSI is transmitted only once through the air, namely when no TMSI has yet been assigned (see registration) or when this data has been lost. Afterwards the (encrypted) TMSI can be used for communicating with the MS.

If the IMSI is stored on the network side only in the AUC, all authentication procedures can be performed with the 2-tuples (RAND, SRES) which were precalculated by the AUC. Besides relieving the load on the VLR (no execution of the A3 algorithm), this kind of

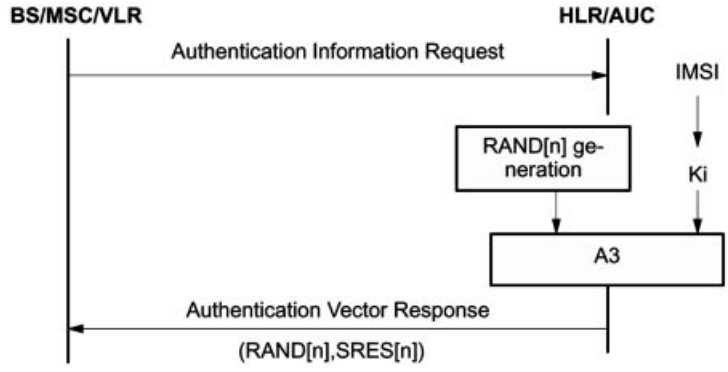

Figure 6.23: Highly secure authentication (no transmission of Ki)

subscriber identification (Figure 6.23) has the other advantage of being particularly secure, because confidential data, especially Ki, need not be transmitted over the air. It should be used especially when the subscriber is roaming in a network of a foreign operator, since it avoids passing of security-critical data over the network boundary.

The less secure variant (Figure 6.24) should only be used within a PLMN. In this case, the secret (security-critical) key Ki is transmitted each time from the HLR/AUC to the current VLR, which executes the algorithm A3 for each authentication.

Figure 6.24: Weakly secure authentication (transmission of Ki to VLR)

6.3.4 Encryption of Signaling and Payload Data

The encryption of transmitted data is a special characteristic of GSM networks that distinguishes the offered service from analog cellular and fixed ISDN networks. This encryption is performed at the transmitting side after channel coding and interleaving and immediately preceding modulation (Figure 6.25). On the receiving side, decryption directly follows the demodulation of the data stream.

A *Cipher Key* (Kc) for the encryption of user data is generated at each side using the generator algorithm A8 and the random number RAND of the authentication process

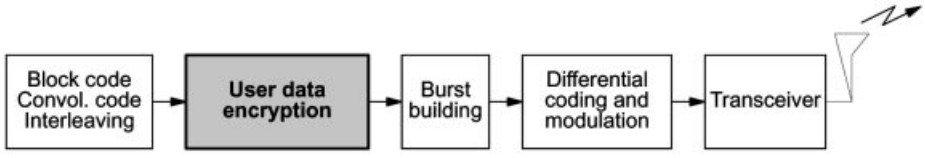

Figure 6.25: Encryption of payload data in the GSM transport chain

(Figure 6.26). This key Kc is then used in the encryption algorithm A5 for the symmetric encryption of user data. At the network side, the values of Kc are calculated in the AUC/HLR simultaneously with the values for SRES. The keys Kc are combined with the 2-tuples (RAND, SRES) to produce 3-tuples, which are stored at the HLR/AUC and supplied on demand, in case the subscriber identification key Ki is only known to the HLR (Section 6.3.2). In the case of the VLR having access to the key Ki, the VLR can calculate Kc directly.

Figure 6.26: Generation of the cipher key Kc

The encryption of signaling and user data is performed at the mobile station as well as at the base station (Figure 6.27). This is a case of symmetric encryption, i.e. ciphering and deciphering are performed with the same key Kc and the A5 algorithm.

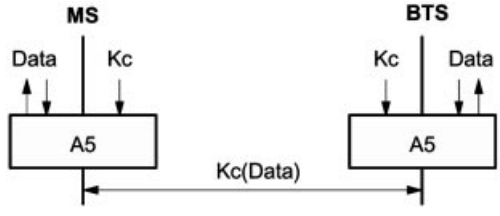

Figure 6.27: Principle of symmetric encryption of user data

Based on the secret key Ki stored in the network, the cipher key Kc for a connection or signaling transaction can be generated at both sides, and the BTS and MS can decipher each other's data. Signaling and user data are encrypted together (TCH/SACCH/FACCH);

for dedicated signaling channels (SDCCH) the same method is used as for traffic channels. This process is also called a *stream cipher*, i.e. ciphering uses a bit stream which is added bitwise to the data to be enciphered (Figure 6.28).

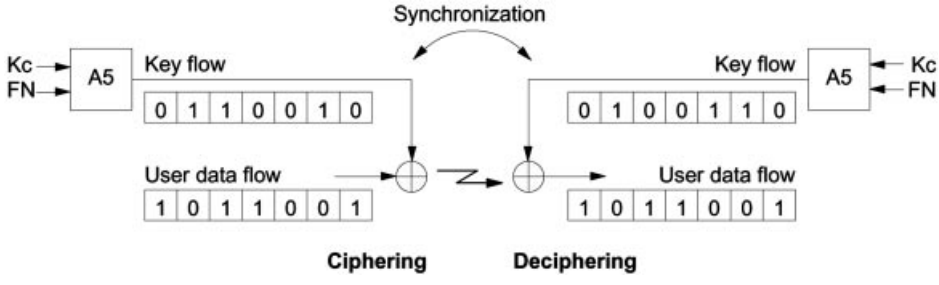

Figure 6.28: Combining payload data stream and ciphering stream

Deciphering consists of performing an additional EXCLUSIVE OR operation of the enciphered data stream with the ciphering stream. The *Frame Number* (FN) of the current TDMA frame within a hyperframe (see Section 5.3.1) is another input for the A5 algorithm besides the key Kc, which is generated anew for each connection or transaction. The current frame number is broadcast on the *Synchronization Channel* (SCH) and is thus available any time to all mobile stations currently in the cell. Synchronization between ciphering and deciphering processes is thus performed through FN.

However, the problem of synchronizing the activation of the ciphering mode has to be solved first: the deciphering mechanism on one side has to be started at precisely the correct moment. This process is started under network control, immediately after the authentication procedure is complete or when the key Kc has been supplied to the base station (BS); see Figure 6.29.

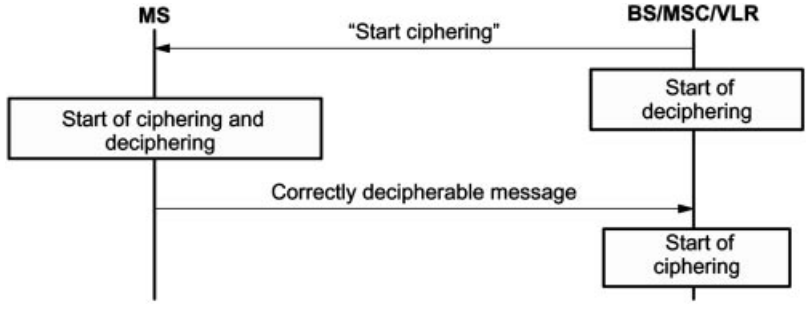

Figure 6.29: Synchronized start of the ciphering process

The network, i.e. the BTS, transmits to the mobile station the request to start its (de)ciphering process, and it starts its own deciphering process. The mobile station then starts its ciphering and deciphering. The first ciphered message from the MS which reaches the network and is correctly deciphered leads to the start of the ciphering process on the network side.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

7 Protocol Architecture

7.1 Protocol Architecture Planes

The various physical aspects of radio transmission across the GSM air interface and the realization of physical and logical channels were explained in Chapter 5. According to the terminology of the OSI Reference Model, these logical channels are at the Service Access Point of Layer 1 (physical layer), where they are visible to the upper layers as transmission channels of the physical layer. The physical layer also includes the forward error correction and the encryption of user data.

The separation of logical channels into the two categories of control channels (signaling channels) and traffic channels (Table 5.1) corresponds to the distinction made in the ISDN Reference Model between user plane and control plane. Figure 7.1 shows a simplified reference model for the GSM *User–Network Interface* (UNI) Um, where the layer-transcending management plane is not elaborated in the following. In the user plane, protocols of the seven OSI layers are defined for the transport of data from a subscriber or a data terminal. User data is transmitted in GSM across the air

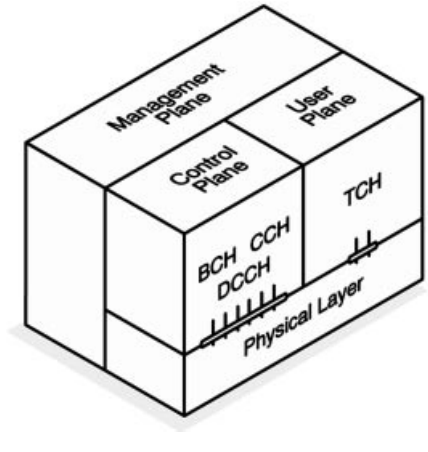

Figure 7.1: Logical channels at the air interface in the ISDN reference model

interface over traffic channels TCH, which therefore belong to Layer 1 of the user plane (Figure 7.1).

Protocols in the signaling plane are used to handle subscriber access to the network and for the control of the user plane (reservation, activation, routing, switching of channels and connections). In addition, signaling protocols between network nodes are needed (network internal signaling). The Dm channels of the air interface in GSM are signaling channels and are therefore realized in the signaling plane (Figure 7.1).

Since signaling channels are physically present but mostly unused during an active user connection, it is obvious to use them also for the transmission of certain user data. In ISDN, packet-switched data communication is therefore permitted on the D channel, i.e. the physical D channel carries multiplexed traffic of signaling data (s-data) and user (payload) data (p-data). The same possibility also exists in GSM. Data transmission without allocation of a dedicated traffic channel is used for the *Short Message Service* (SMS) by using free capacities on signaling channels. For this purpose, a separate SDCCH is allocated, or, if a traffic connection exists, the SMS protocol data units are multiplexed onto the signaling data stream of the SACCH (Figure 7.2).

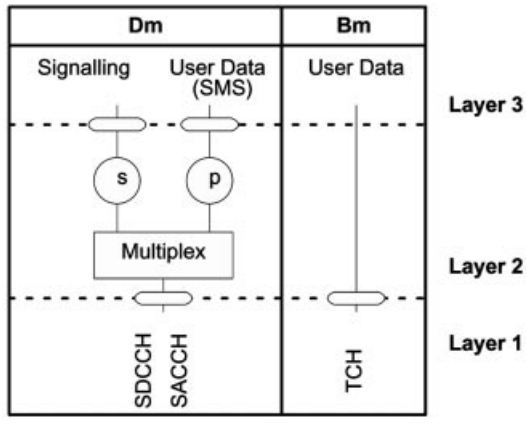

Figure 7.2: User data and control at the air interface

The control (signaling) and user plane can be defined and implemented separately of each other, ignoring for the moment that control and user data have to be transmitted across the same physical medium at the air interface and that signaling procedures initiate and control activities in the user plane. Therefore, for each plane there exists a corresponding separate protocol architecture within the GSM system: the user data protocol architecture (see Section 7.2) and the signaling protocol architecture (see Section 7.3), with an additional separate protocol architecture for the transmission of p-data on the control (signaling) plane (see Section 7.3.2). A protocol architecture comprises not only the protocol entities at the radio interface Um but all protocol entities of the GSM network components.

7.2 Protocol Architecture of the User Plane

A GSM PLMN can be defined by a set of access interfaces (see Section 9.1) and a set of connection types used to realize the various communication services. A connection in GSM is defined between reference points. Connections are constructed from connection elements (Figure 7.3), and the signaling and transmission systems may change from element to element. Two elements therefore exist within a GSM connection: the radio interface connection element and the A interface connection element. The radio interface and the pertinent connection element are defined between the MS and the BSS, whereas the A interface connection element exists between BSS and MSC across the A interface. A GSM-specific signaling system is used at the radio interface, whereas ISDN-compatible signaling and payload transport are used across the A interface. The BSS is subdivided into BTS and BSC. Between them they define the Abis interface, which has no connection element defined; this is because it is usually transparent for user data.

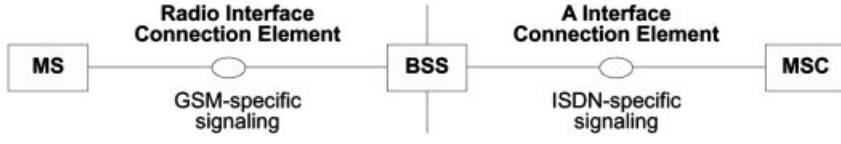

Figure 7.3: Connection elements

A GSM connection type provides a way to describe GSM connections. Connection types represent the capabilities of the lower layers of the GSM PLMN. In the following section, the protocol models are presented as the basis for some of the connection types defined in the GSM standards. These are speech connections and transparent as well as nontransparent data connections. A detailed discussion of the individual connection types can be found in Chapter 9 with a description of how various data services have been realized in GSM.

7.2.1 Speech Transmission

The digital, source-coded speech signal of the mobile station is transmitted across the air interface in error-protected and encrypted form. The signal is then deciphered in the BTS, and the error protection is removed before the signal is passed on. This specially protected speech transmission occurs transparently between mobile station and a *Transcoding and Rate Adaptation Unit* (TRAU) which serves to transform the GSM speech-coded signals to the ISDN standard format (ITU-T A-law). A possible transport path for speech signals is shown in Figure 7.4, where the bit transport plane (encryption and TDMA/FDMA) has been omitted.

A simple GSM speech terminal (MT0, see also Figure 9.1) contains a *GSM Speech Codec* (GSC) for speech coding. Its speech signals are transmitted to the BTS after channel coding (FEC) and encryption, where they are again deciphered, decoded, and

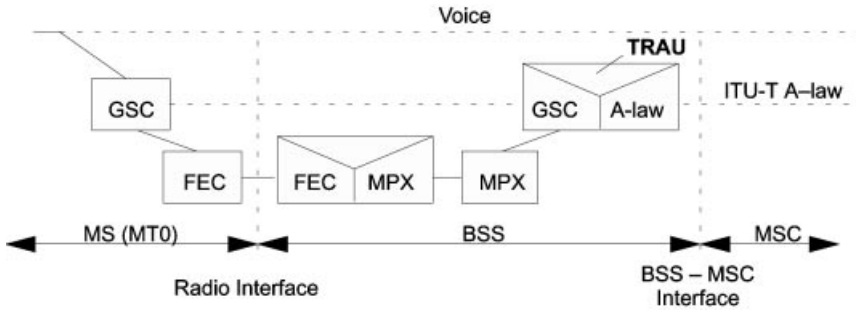

Figure 7.4: Speech transmission in GSM

if necessary, error-corrected. More than one GSM speech signal can be multiplexed onto an ISDN channel, with up to four GSM speech signals (at 13 kbit/s each) per ISDN B channel (64 kbit/s). Before they are passed to the MSC, speech signals are transcoded in the BSS from GSM format to ISDN format (ITU-T A-law).

The BTSs are connected to the BSC over digital fixed lines, usually leased lines or microwave links, with typical transmission rates of 2048 kbit/s (in Europe), 1544 kbit/s (in the USA) or 64 kbit/s (ITU-T G. 703, G. 705, G. 732). For speech transmission, the BSS implements channels of 64 or 16 kbit/s. The physical placement of the *Transcoding and Rate Adaptation Unit* (TRAU) largely determines which kind of speech channel is used in the fixed network. The TRAU performs the conversion of speech data between GSM format (13 kbit/s) and ISDN A-law format (64 kbit/s). In addition, it is responsible for the adaptation of data rates, if necessary, for data services. There are two alternatives for the positioning of the TRAU: the TRAU can be placed into the BTS or outside of the BTS into the BSC. An advantage of placing the TRAU outside of the BTS is that up to four speech signals can be submultiplexed (MPX in Figure 7.4) onto an ISDN B channel, so that less bandwidth is required on the BTS-to-BSC connection. Beyond this consideration, placing the TRAU outside of the BTS allows the TRAU functions to be combined for all BTSs of a BSS in one separate hardware unit, perhaps produced by a separate manufacturer. The TRAU is, however, always considered as part of the BSS and not as an independent network element.

Figure 7.5 shows some variants of TRAU placement. A BTS consists of a *Base Control Function* (BCF) for general control functions like frequency hopping, and several (at least one) *Transceiver Function* (TRX) modules which realize the eight physical TDMA channels on each frequency carrier. The TRX modules are also responsible for channel coding and decoding as well as encryption of speech and data signals. If the TRAU is integrated into the BTS, speech transcoding between GSM and ISDN formats is also done within the BTS.

In the first case, TRAU within the BTS (BTS 1,2,3 in Figure 7.5), the speech signal in the BTS is transcoded into a 64 kbit/s A-law signal, and a single speech signal per B channel (64 kbit/s) is transmitted to the BSC/MSC. For data signals, the bit rates are adapted to 64 kbit/s, or several data channels are submultiplexed over one ISDN channel. The resulting user plane protocol architecture for speech transport is shown in Figure 7.6.

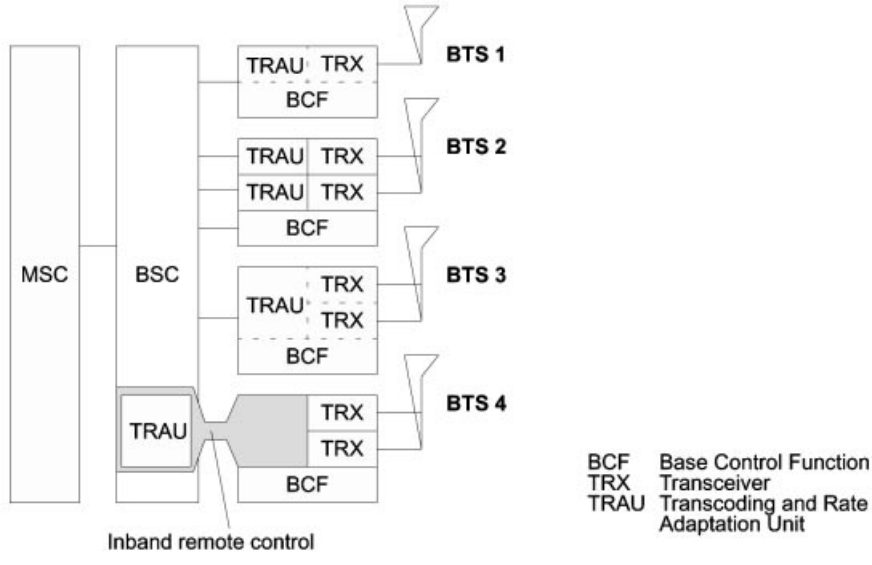

Figure 7.5: BTS architecture variations and TRAU placement

GSM-coded speech (13 kbit/s) is transmitted over the radio interface (Um) in a format that is coded for error protection and encryption. At the BTS site, the GSM signal is transcoded into an ISDN speech signal and transmitted transparently through the ISDN access network of the MSC.

In the second case, the TRAU resides outside of the BTS (BTS 4 in Figure 7.5) and is considered a part of the BSC. However, physically it could also be located at the MSC site, i.e. at the MSC side of the BSC-to-MSC links (Figure 7.7). Channel coding/decoding and encryption are still performed in the TRX module of the BTS, whereas speech transcoding takes place in the BSC. For control purposes, the TRAU needs to receive synchronization and decoding information from the BTS, e.g. *Bad Frame Indication* (BFI) for error concealment (see Section 6.1). If the TRAU does not reside in the BTS, it must be remotely controlled from the BTS by inband signaling. For this purpose, a subchannel of 16 kbit/s is reserved for the GSM speech signal on the BTS-to-BSC link,

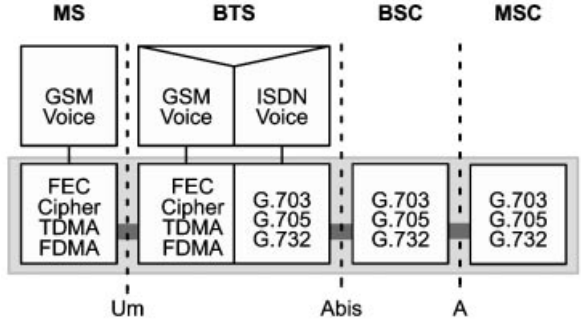

Figure 7.6: GSM protocol architecture for speech (TRAU at BTS site) Um

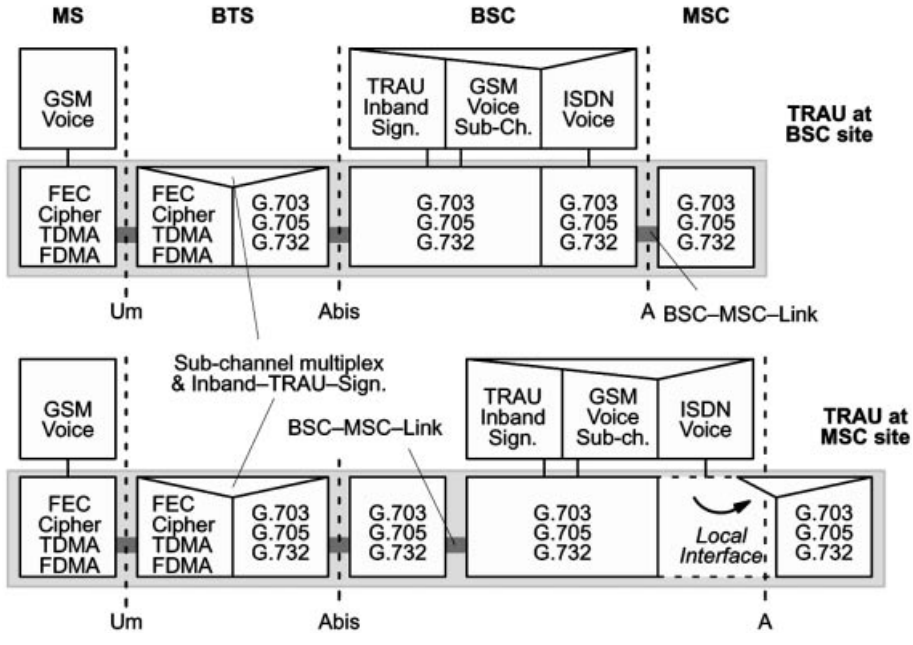

Figure 7.7: GSM protocol architecture for speech

so an additional 3 kbit/s is made available for inband signaling. Alternatively, the GSM speech signal with added inband signaling could also be transmitted in a full ISDN B channel.

7.2.2 Transparent Data Transmission

The digital mobile radio channel is subject to severe quality variations and generates burst errors, which one tries to correct through interleaving and convolutional codes (see Section 6.2). However, if the signal quality is too low due to fading breaks or interference, the resulting errors cannot be corrected. For data transmission across the air interface Um, a residual bit error ratio varying between 10^{-2} and 10^{-5} according to channel conditions can be observed [58]. This kind of variable quality of data transmission at the air interface determines the service quality of transparent data transmission. Transparent data transmission defines a GSM connection type used for the realization of some basic bearer services (transparent asynchronous and synchronous data, Table 4.2). The pertinent protocol architecture is illustrated in Figure 7.8. The main aspect of the transparent connection type is that user data is protected against transmission errors by forward error correction only across the air interface. Further transmission within the GSM network to the next MSC with an interworking function (IWF) to an ISDN or a PSTN occurs unprotected on digital line segments, which have anyway a very low bit error ratio in comparison to the radio channel. The transparent GSM data service offers a constant throughput rate and constant trans-

mission delay; however, the residual error ratio varies with channel quality due to the limited correction capabilities of the FEC.

For example, take a data terminal communicating over a serial interface of type V.24. A transparent bearer service provides access to the GSM network directly at a mobile station or through a terminal adapter (reference point R in Figure 9.1). A data rate of up to 9600 bit/s can be offered based on the transmission capacity of the air interface and using an appropriate bit rate adaptation. The bit rate adaptation also performs the required asynchronous-to-synchronous conversion at the same time. This involves supplementing the tokens arriving asynchronously from the serial interface with fill data, since the channel coder requires a fixed block rate. This way there is a digital synchronous circuit-switched connection between the terminal accessing the service and the IWF in the MSC, which extends across the air interface and the digital ISDN B channel inside the GSM network; this synchronous connection is completely transparent for the asynchronous user data of the terminal equipment (TE).

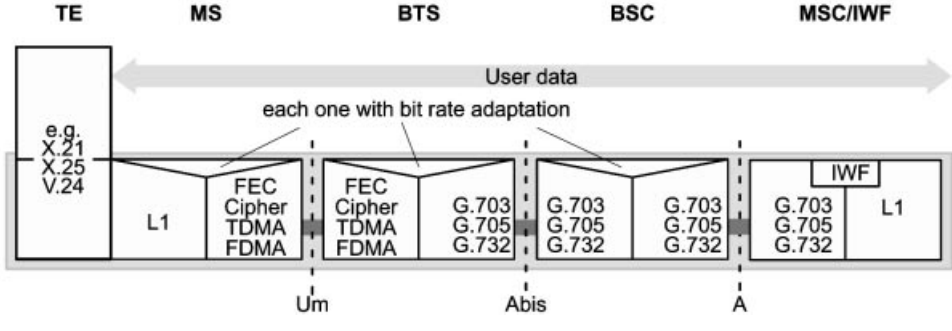

Figure 7.8: GSM protocol architecture for transparent data

7.2.3 Nontransparent Data Transmission

Compared to the bit error ratio of the fixed network, which is on the order of 10^{-6} to 10^{-9}, the quality of transparent data service is often insufficient for many applications, especially under adverse conditions. To provide more protection against transmission errors, more redundancy has to be added to the data stream. Since this redundancy is not always required, but only when there are residual errors in the data stream, forward error correction is inappropriate. Rather, an error detection scheme with automatic retransmission of faulty blocks is used, *Automatic Repeat Request* (ARQ). Such an ARQ scheme which was specifically adapted to the GSM channel, is the *Radio Link Protocol* (RLP). The assumption for RLP is that the underlying forward error correction of the convolutional code realizes a channel with an average block error ratio of less than 10%, with a block corresponding to an RLP protocol frame of length 240 bits. Now the nontransparent channel experiences a constantly lower bit error ratio than the transparent channel, independent of the varying transmission quality of the radio channel; however, due to the

RLP-ARQ procedure both throughput and transmission delay vary with the radio channel quality.

The data transmission between mobile station and interworking function of the next MSC is protected with the data link layer protocol RLP, i.e. the endpoints of RLP terminate in MS and IWF entities, respectively (Figure 7.9). At the interface to the data terminal TE, a *Nontransparent Protocol* (NTP) and an *Interface Protocol* (IFP) are defined, depending on the nature of the data terminal interface. Typically, a V.24 interface is used to carry character-oriented user data. These characters of the NTP are buffered and combined into blocks in the *Layer 2 Relay* (L2R) protocol, which transmits them as RLP frames. The data transport to and from the data terminal is flow-controlled. Therefore, transmission within the PLMN is no longer transparent for the data terminal. At the air interface, a new RLP frame is transmitted every 20 ms; thus L2R may have to insert fill tokens, if a frame cannot be completely filled at transmission time.

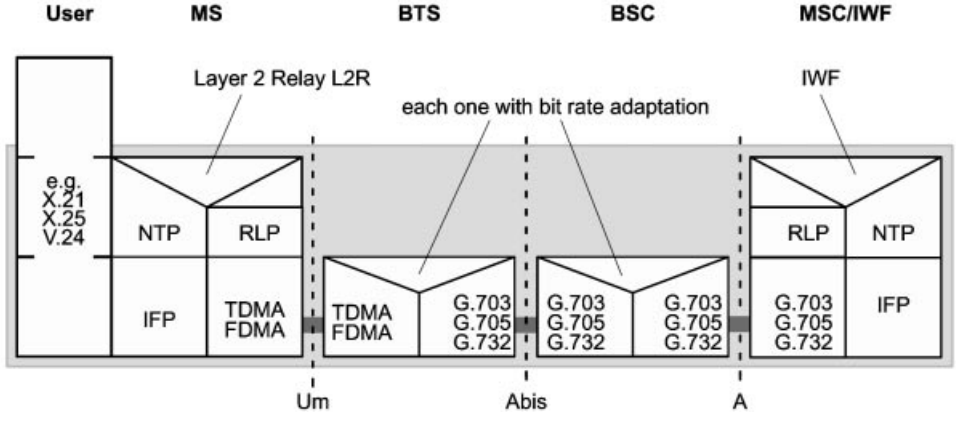

Figure 7.9: GSM protocol architecture for nontransparent data

The RLP protocol is very similar to the HDLC of ISDN with regard to frame structure and protocol procedures, the main difference being the fixed frame length of 240 bits, in contrast to the variable length of HDLC. The frame consists of a 16-bit protocol header, 200-bit information field, and a 24-bit *Frame Check Sequence* (FCS); see Figure 7.10. Because of the fixed frame length, the RLP has no reserved flag pattern, and a special procedure to realize code transparency like bit stuffing in HDLC is not needed. The very short – and hence less error prone – frames are exactly aligned with channel coding blocks. (The probability of frame errors increases with the length of the frame.)

RLP makes use of the services of the lower layers to transport its protocol data units (PDUs). The channel offered to RLP therefore has the main characteristic of a 200 ms transmission delay, besides the possibly occurring residual bit errors. The delay is mostly caused by interleaving and channel coding, since the transmission itself takes only about 25 ms for a data rate of 9600 bit/s. This means it will take at least 400 ms until a positive

acknowledgement is received for an RLP frame, and protocol parameters like transmission window and repeat timers need to be adjusted accordingly.

The RLP header is similar to the one used in HDLC [31], with the difference that the RLP header contains no address information but only control information for which 16 bits are available. One distinguishes between *supervisory frames* and *information frames*. Whereas information frames carry user data, supervisory frames serve to control the connection (initialize, disconnect, reset) as well as the retransmission of information frames during data transfer. The information frames are labeled with a sequence number N(S) for identification, for which 6 bits are available in the RLP header (Figure 7.10). To conserve space, this field is also used to code the frame type. Sequence number values smaller than 62 indicate that the frame carries user data in the information field (information frame). Otherwise the information field is discarded, and only the control information in the header is of interest (supervisory frame). These frames are marked with the reserved values 62 and 63 (Figure 7.10).

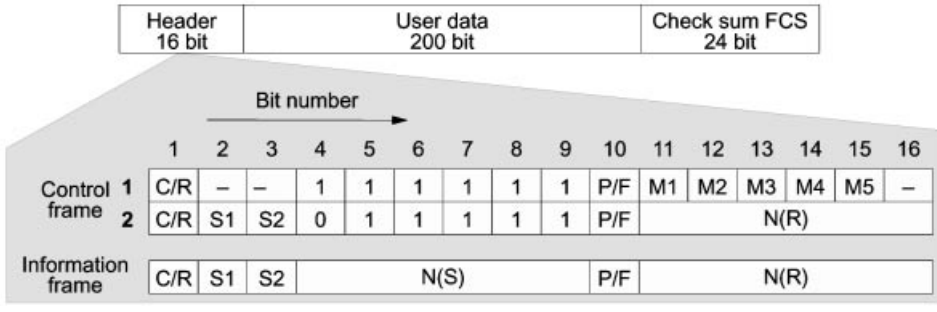

Figure 7.10: Frame structure of the RLP protocol

Due to this header format, information frames can also carry (implicit) control information, a process known as piggybacking. The header information of the second variant can be carried completely within the header of an information frame. This illustrates further how RLP has been adapted to the radio channel, since it makes the transmission of additional control frames unnecessary during information transfer, which reduces the protocol overhead and increases the throughput.

Thus the send sequence number is calculated modulo 62, which amounts to a window of 61 frames, allowing 61 outstanding frames without acknowledgement before the sender has to receive the acknowledgement of the first frame. Positive acknowledgement is used; i.e. the receiver sends an explicit supervisory frame as a receipt or an implicit receipt within an information frame. Such an acknowledgement frame contains a receive frame number N(R) which designates correct reception of all frames, including send sequence number $N(S) = N(R) - 1$.

Each time the last information frame is sent, a timer T1 is started at the sender. If an acknowledgement for some or all sent frames is not received in time, perhaps because the acknowledging RLP frame had errors and was therefore discarded, the timer expires and causes the sender to request an explicit acknowledgement. Such a request may be

repeated N2 times; if this still leads to no acknowledgement, the connection is terminated. If an acknowledgement N(R) is obtained after expiration of timer T1, all sent frames starting from and including N(R) are retransmitted. In the case of an explicitly requested acknowledgement, this corresponds to a modified *Go-back-N* procedure. Such a retransmission is also allowed only up to N2 times. If no receipt can be obtained even after N2 trials, the RLP connection is reset or terminated.

Two procedures are provided in RLP for dealing with faulty frames: *selective reject*, which selects a single information frame without acknowledgement; and *reject*, which causes retransmission with implicit acknowledgement. With *selective reject*, the receiving RLP entity requests retransmission of a faulty frame with sequence number N(R), but this does not acknowledge receipt of other frames. Each RLP implementation must at least include the *reject* method for requesting retransmission of faulty frames. With a *reject*, the receiver asks for retransmission of all frames starting with the first defective received frame with number N(R) (*Go-back-N*). Simultaneously, this implicitly acknowledges correct reception of all frames up to and including N(R) − 1. Realization of *selective reject* is not mandatory in RLP implementations, but it is recommended. The reason is that *Go-back-N* causes retransmission of frames that may have been transmitted correctly and thus deteriorates the throughput that could be achieved with selective reject.

7.3 Protocol Architecture of the Signaling Plane

7.3.1 Overview of the Signaling Architecture

Figure 7.11 shows the essential protocol entities of the GSM signaling architecture (control plane or signaling plane). Three connection elements are distinguished: the radio-interface connection element, the BSS-interface connection element, and the A-interface connection element. This control plane protocol architecture consists of a GSM-specific part with the interfaces Um and Abis and a part based on *Signaling System Number 7* (SS#7) with the interfaces A, B, C, E (Figure 7.11). This change of signaling system corresponds to the change from radio interface connection element to A-interface connection element as discussed above for the user data plane (Figure 7.3).

The radio interface Um is defined between MS and BSS, more exactly between MS and BTS. Within the BSS, the BTS and the BSC cooperate over the Abis interface, whereas the A interface is located between BSC and MSC. The MSC has also signaling interfaces to VLR (B), HLR (C), to other MSCs (E), and to the EIR (F). Further signaling interfaces are defined between VLRs (G) and between VLR and HLR (D). Figure 3.9 gives an overview of the interfaces in a GSM PLMN.

Physical Layer – In the control plane, the lowest layer of the protocol model at the air interface, the *Physical Layer*, implements the logical signaling channels (TDMA/FDMA, multiframes, channel coding, etc.; see Chapter 5, Sections 6.1, 6.2, and 6.3). Like user data, signaling messages are transported over the Abis interface (BTS-BSC) and the A interface (BSC-MSC) on digital lines with data rates of 2048 kbit/s (1544 kbit/s in the USA), or 64 kbit/s (ITU-T G.703, G.705, G.732).

Layer 2: LAPDm – On Layer 2 of the logical signaling channels across the air interface,

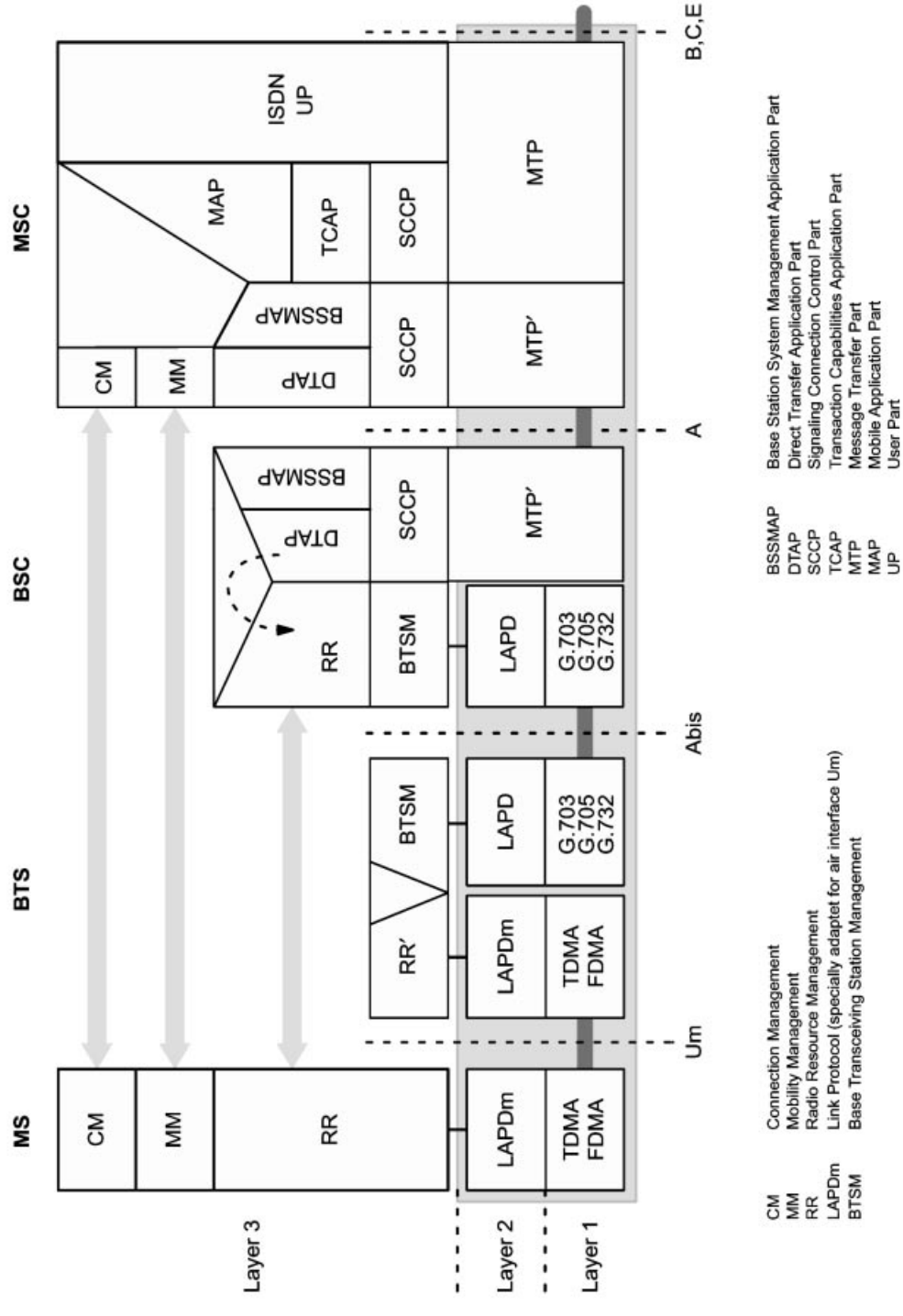

Figure 7.11: GSM protocol architecture for signaling

a data link protocol entity is implemented, the *Link Access Procedure* on *Dm* channels (LAPDm). LAPDm is a derivative of LAPD which is specifically adapted to the air interface. This data link protocol is responsible for the protected transfer of signaling messages between MS and BTS over the air interface, i.e. LAPDm is terminated in mobile station and base station.

In essence, LAPDm is a protocol similar to HDLC which offers a number of services on the various logical Dm channels of Layer 3: connection setup and teardown, protected signaling data transfer. It is based on various link protocols used in fixed networks, such as LAPD in ISDN [7]. The main task of LAPDm is the transparent transport of messages between protocol entities of Layer 3 with special support for:

- Multiple entities in Layer 3 and Layer 2
- Signaling for broadcasting (BCCH)
- Signaling for paging (PCH)
- Signaling for channel assignment (AGCH)
- Signaling on dedicated channels (SDCCH)

A detailed discussion of LAPDm is presented in Section 7.4.2.

Layer 3 – In the mobile station, the LAPDm services are used at Layer 3 of the signaling

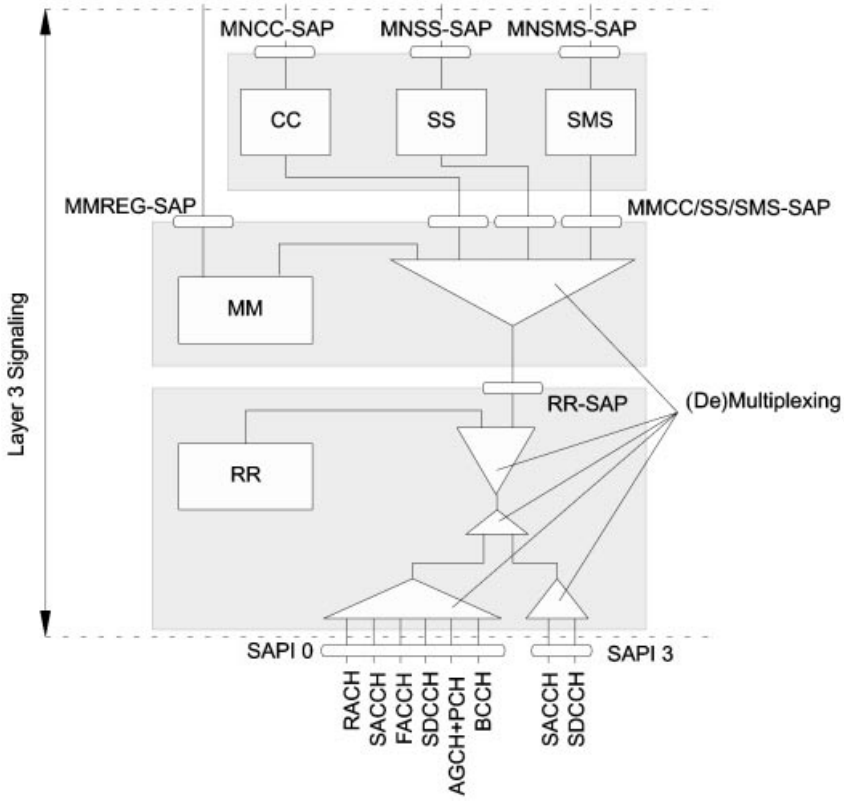

Figure 7.12: Layer 3 protocol architecture at the MS side

protocol architecture. There, Layer 3 is divided into three sublayers: *Radio Resource Management* (RR), *Mobility Management* (MM), and *Connection Management* (CM). The protocol architecture formed by these three sublayers is shown in Figure 7.12. Connection management is further subdivided into three protocol entities: *Call Control* (CC), *Supplementary Services* (SS), and *Short Message Service* (SMS). Additional multiplexing functions within Layer 3 are required between these sublayers.

The call-independent supplementary services and the short message service are offered to higher layers at two *Service Access Points* (SAPs), MNSS and MNSMS. A more detailed look at the services offered by the RR, MM, and CC protocol entities is given in the following.

Radio Resource Management – *Radio Resource Management* (RR) essentially handles the administration of the frequencies and channels. This involves the RR module of the MS communicating with the RR module of the BSC (Figure 7.11). The general objective of the RR is to set up, maintain, and take down RR connections which enable point-to-point communication between MS and network. This also includes cell selection in idle mode and handover procedures. Furthermore, the RR is responsible for monitoring BCCH and CCCH on the downlink when no RR connections are active.

The following functions are realized in the RR module:

- Monitoring of BCCH and PCH (readout of system information and paging messages)
- RACH administration: mobile stations send their requests for connections and replies to paging announcements to the BSS
- Requests for and assignments of data and signaling channels
- Periodic measurement of channel quality (quality monitoring)
- Transmitter power control and synchronization of the MS
- Handover (part of which is sometimes erroneously attributed to roaming functions and mobility management), always initiated by the network
- Synchronization of encryption and decryption on the data channel

The RR sublayer provides several services at the RR-SAP to the MM sublayer. These services are needed to set up and take down signaling connections and to transmit signaling messages.

Mobility Management – *Mobility Management* (MM) encompasses all the tasks resulting from mobility. The MM activities are exclusively performed in cooperation between MS and MSC, and they include

- TMSI assignment
- Localization of the MS
- Location updating of the MS; parts of this are sometimes known as roaming functions
- Identification of the MS (IMSI, IMEI)
- Authentication of the MS
- IMSI attach and detach procedures (e.g. at insertion or removal of SIM)
- Ensuring confidentiality of subscriber identity

Registration services for higher layers are provided by Layer 3 at the MMREG-SAP (Figure 7.12). Registration involves the IMSI attach and detach procedures which are

used by the mobile to report state changes such as power-up or power-down, or SIM card removal or insertion.

The MM sublayer offers its services at the MMCC-SAP, MMSS-SAP, and MMSMS-SAP to the CC, SS, and SMS entities. This is essentially a connection to the network side over which these units can communicate.

Connection Management – *Connection Management* consists of three entities: *Call Control* (CC), *Supplementary Services* (SS), and *Short Message Service* (SMS). Call control handles all tasks related to setting up, maintaining and taking down calls. The services of call control are provided at the MNCC-SAP, and they encompass:

- Establishment of normal calls (MS-originating and MS-terminating)
- Establishment of emergency calls (only MS-originating)
- Termination of calls
- *Dual-Tone Multifrequency* (DTMF) signaling
- Call-related supplementary services
- Incall modification: the service may be changed during a connection (e.g. speech and transparent/nontransparent data are alternating; or speech and fax alternate)

The service primitives at this SAP of the interface to higher layers report reception of incoming messages and effect the sending of messages, essentially ISDN user-network signaling according to Q.931.

RR messages are mainly exchanged between MS and BSS. In contrast, CM and MM functions are handled exclusively between MS and MSC; the exact division of labor between BTS, BSC, and MSC is summarized in Table 7.1. As can be seen, RR messages have to be transported over the Um and Abis interfaces, whereas CM and MM messages need additional transport mechanisms across the A interface.

Message Transfer Part – From a conceptual viewpoint, the A interface in GSM networks is the interface between the MSCs, the ISDN exchanges with mobile network specific extensions, and the BSC, the dedicated mobile network specific control units. Here too is the reference point, where the signaling system changes from GSM-specific to the general ISDN-compatible SS#7. Message transport in the SS#7 network is realized through the *Message Transfer Part* (MTP). In essence, MTP comprises the lower three layers of the OSI Reference Model, i.e. the MTP provides routing and transport of signaling messages.

A slightly modified (reduced) version of the MTP, called MTP′, has been defined for the protected transport of signaling messages across the A interface between BSC and MSC. At the ISDN side of the MSCs, the complete MTP is available. For signaling transactions between MSC and MS (CM, MM), it is necessary to establish and identify distinct logical connections. The *Signaling Connection Control Part* (SCCP) is used for this purpose to facilitate implementation with a slightly reduced range of functions defined in SS#7.

BSS Application Part – For GSM-specific signaling between MSC and BSC, the *Base Station System Application Part* (BSSAP) has been defined. The BSSAP consists of the *Direct Transfer Application Part* (DTAP) and the *Base Station System Management Application Part* (BSSMAP). The DTAP is used to transport messages between MSC and MS. These are the *Call Control* (CC) and *Mobility Management* (MM) messages. At

the A interface, they are transmitted with DTAP and then passed transparently through the BSS across the Abis interface to the MS without interpretation by the BTS.

Table 7.1: Distribution of functions between BTS, BSC, and MSC (according to GSM Rec. 08.02, 08.52)

	BTS	BSC	MSC
Terrestrial channel management			
MSC-BSC-channels			
Channel allocation			X
Blocking indication		X	
BSC-BTS-Channels			
Channel allocation		X	
Blocking indication	X		
Mobility management			
Authentication			X
Location updating			X
Call control			X
Radio channel management			
Channel coding/decoding	X		
Transcoding/rate adaptation	X		
Interworking function			X
Measurements			
Uplink measuring	X		X
Processing of reports from MS/TRX	X	X	X
Traffic measurements			X
Handover			
BSC internal, intracell		X	
BSC internal, intercell		X	
BSC external		X	
Recognition, decision, execution			X
HO access detection	X		
Paging			
Initiation		X	
Execution	X		
Channel configuration management		X	
Frequency hopping			
Management		X	

Table 7.1 (*continued*)

	BTS	BSC	MSC
Execution	X		
TCH management			
Channel allocation		X	
Link supervision		X	
Channel release		X	X
Idle channel observation	X		
Power control determination	X	X	
SDCCH management			
SDCCH allocation		X	
Link supervision		X	
Channel release		X	X
Power control determination	X	X	
BCCH/CCCH management			
Message scheduling management		X	
Message scheduling execution	X		
Random access detection	X		
Immediate assign		X	
Timing Advance			
Calculation	X		
Signaling to MS at random access		X	
Signaling to MS at handover/during call	X		
Radio resource indication			
Report status of idle channels	X		
LAPDm functions	X		
Encryption			
Management		X	
Execution	X		

The BSSMAP is the protocol definition part which is responsible for all of the administration and control of the radio resources of the BSS. RR is one of the main functions of a BSS. Therefore, the RR entities terminate in the mobile station and the BTS or BSC respectively. Some functions of RR however, require involvement of the MSC (e.g. some handover situations, or release of connections or channels). Such actions should be initiated and controlled by the MSC (e.g. handover and channel assignment). This control is the responsibility of BSSMAP. RR messages are mapped and converted within the BSC into procedures and messages of BSSMAP and

vice versa. BSSMAP offers the functions which are required at the A interface between BSS and MSC for RR of the BSS. Accordingly, RR messages initiate BSSMAP functions, and BSSMAP functions control RR protocol functions.

BTS Management – A similar situation exists at the Abis interface. Most of the RR messages are passed transparently by the BTS between MS and BSC. Certain RR information, however, must be interpreted by the BTS, e.g. in situations like random access of the MS, the start of the ciphering process, or paging to localize an MS for connection setup. The *Base Transceiver Station Management* (BTSM) contains functions for the treatment of these messages and other procedures for BTS management. Besides, a mapping occurs in the BTS from BTSM onto the RR messages relevant at the air interface (RR', Figure 7.11).

Mobile Application Part – The MSC is equipped with the *Mobile Application Part* (MAP), a mobile network specific extension of SS#7, for communication with the other components of the GSM network (the HLR and VLR registers, other MSCs) and other PLMNs. Among the MAP functions are all signaling functions among MSCs as well as between MSC and the registers (Figure 7.13). These functions include

- Updating of residence information in the VLR
- Cancellation of residence information in the VLR
- Storage of routing information in the HLR
- Updating and supplementing of user profiles in HLR and VLR
- Inquiry of routing information from the HLR
- Handover of connections between MSC

The exchange of MAP messages, e.g. with other MSCs, HLR, or VLR, occurs over the transport and transaction protocol of the SS#7. The SS#7 transaction protocol is the

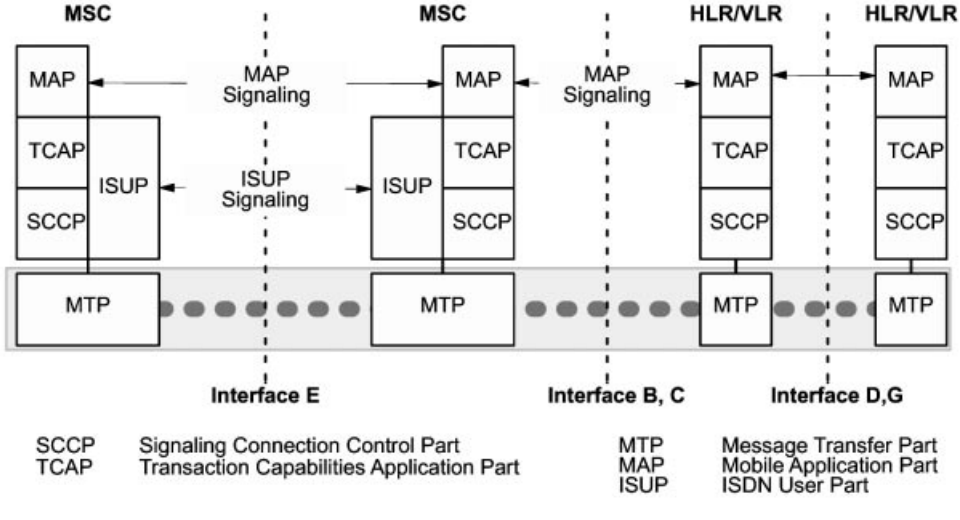

Figure 7.13: Protocol interfaces in the mobile network

Transaction Capabilities Application Part (TCAP). A connectionless transport service is offered by the *Signaling Connection Control Part* (SCCP).

The MAP functions require channels for signaling between different PLMNs which are provided by the international SS#7. Access to SS#7 occurs through the fixed ISDN.

Connection to the fixed network is typically done through leased lines; in the case of the German GSM network operators, it is through lines with a rate of 2 Mbit/s from Deutsche Telekom [25]. Often the majority of the MSC in a PLMN has such an access to the fixed network. On these lines, both user data and signaling data is transported. From the viewpoint of a fixed network, an MSC is integrated into the network like a normal ISDN exchange node. Outside of a PLMN, starting with the GMSC, calls for mobile stations are treated like calls for subscribers of the fixed network, i.e. the mobility of a subscriber with an MSISDN becomes ''visible'' only beyond the GMSC. For CC, the MSC has the same interface as an ISDN switching node. Connection-oriented signaling of GSM networks is mapped at the fixed network side (interface to ISDN) into the *ISDN User Part* (ISUP) used to connect ISDN channels through the network (Figure 7.14). The mobile-specific signaling of the MAP is routed over a gateway of the PLMN (GMSC) and the *International Switching Center* (ISC) of the national ISDN network into the international SS#7 network [25]. In this way, transport of signaling data between different GSM networks is also guaranteed without problems.

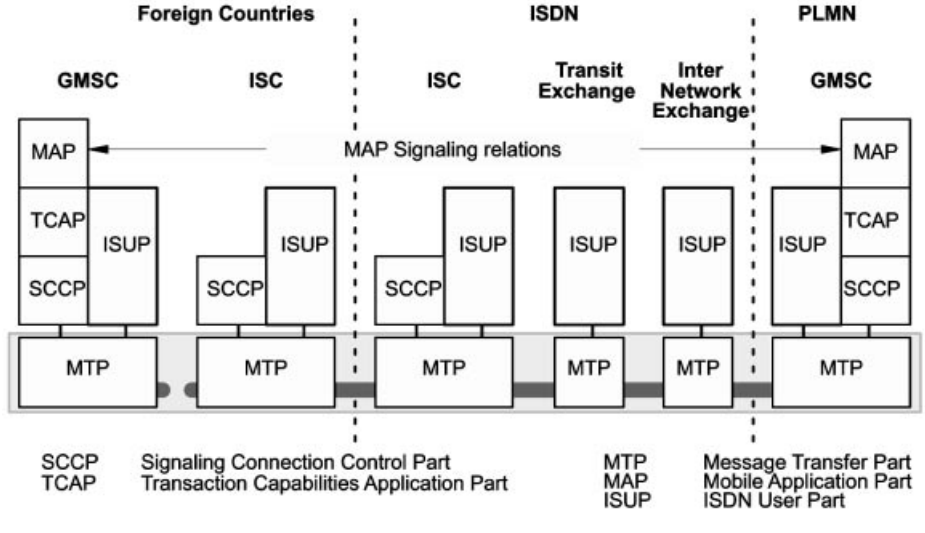

Figure 7.14: International signaling relations via ISDN [25]

7.3.2 Transport of User Data in the Signaling Plane

In the signaling plane (control plane) of the GSM architecture, one can also transport packet-oriented user data from or to mobile stations. This occurs for the point-to-point

SMS (see Section 4.2). Short messages are always transmitted in store-and-forward mode through a *Short Message Service Center* (SMS-SC). The service center accepts these messages, which can be up to 160 characters long, and forwards them to the recipients (other mobile stations or fax, email, etc.). In principle, GSM defines a separate protocol architecture for the realization of this service.

Between mobile station and service center, short messages are transmitted using a connectionless transport protocol: *Short Message Transport Protocol* (SM-TP) which uses the services of the signaling protocols within the GSM network. Transport of these messages outside of the GSM network is not defined. For example, the SMS-SC could be directly connected to the gateway switching center (SMS-GMSC), or it could be connected to a *Short Message Service Interworking MSC* (SMS-IWMSC) through an X.25 connection (Figure 7.15). Within the GSM network between MSCs, a short message is transferred with the MAP and the lower layers of SS#7. Finally, between a mobile station and its local MSC, two protocol layers are responsible for the transfer of transport protocol units of SMS. First, there is the SMS entity in the CM sublayer of Layer 3 at the user-network interface (see Figure 7.12) which realizes the *Short Message Control Protocol* (SM-CP) and its connection-oriented service. Second, there is the relay layer, in which the *Short Message Relay Protocol* (SM-RP) is defined, which offers a connectionless service for transfer of SMS transport PDUs between MS and MSC. This, however, uses services at the service access point MMSMS-SAP (see Figure 7.12) and thus a connection of the MM sublayer.

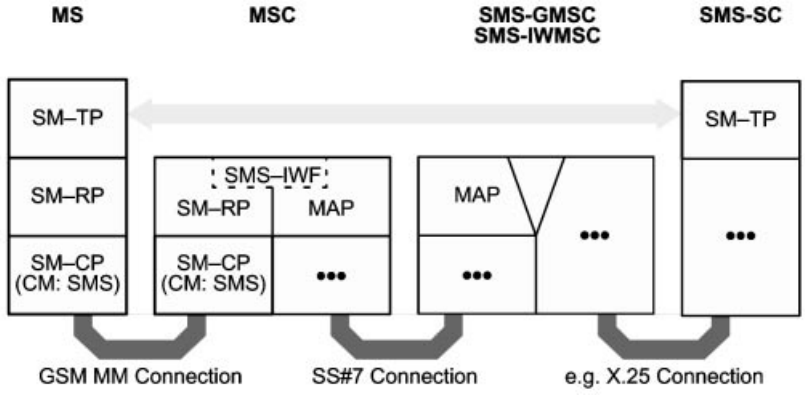

Figure 7.15: Protocol architecture for SMS transfer

In addition to the SM-CP, the relay protocol SM-RP was introduced above the CM sublayer (Figures 7.12 and 7.15) to realize an acknowledged transmission of short messages, but with minimal overhead for the radio channel. A short message sent by a mobile station is passed over the signaling network until it reaches the service center SMS-SC. If the service center determines the error-free reception of a message, an acknowledgement message is returned on the reverse path, which finally causes sending of an acknowledgement message from the SM-RP entity in the MSC to the mobile station. Until this acknowledgement message arrives, the connection in the MM sublayer

can be taken down, and thus also the reserved radio channel. In this way, radio resources across the air interface are only occupied during the actual transmission of SM-RP messages. And each successful transmission of an SM-CP PDU across the MM connection, which includes the error-prone air interface, is immediately acknowledged, or else errors are immediately reported to the sending SM-CP entity. So if a message is damaged at the radio interface, this avoids it being transmitted to the service center.

7.4 Signaling at the Air Interface (Um)

Signaling at the user–network interface in GSM is essentially concentrated in Layer 3. Layers 1 and 2 provide the mechanisms for the protected transmission of signaling messages across the air interface. Besides the local interface, they contain functionality and procedures for the interface to the BTS.

The signaling of Layer 3 at the user–network interface is very complex and comprises protocol entities in the mobile station and in all functional entities of the GSM network (BTS, BSC, and MSC).

7.4.1 Layer 1 of the MS-BTS Interface

Layer 1 of the OSI Reference Model (physical layer) contains all the functions necessary for the transmission of bit streams over the physical medium, in this case the radio channel. GSM Layer 1 defines a series of logical channels based on the channel access

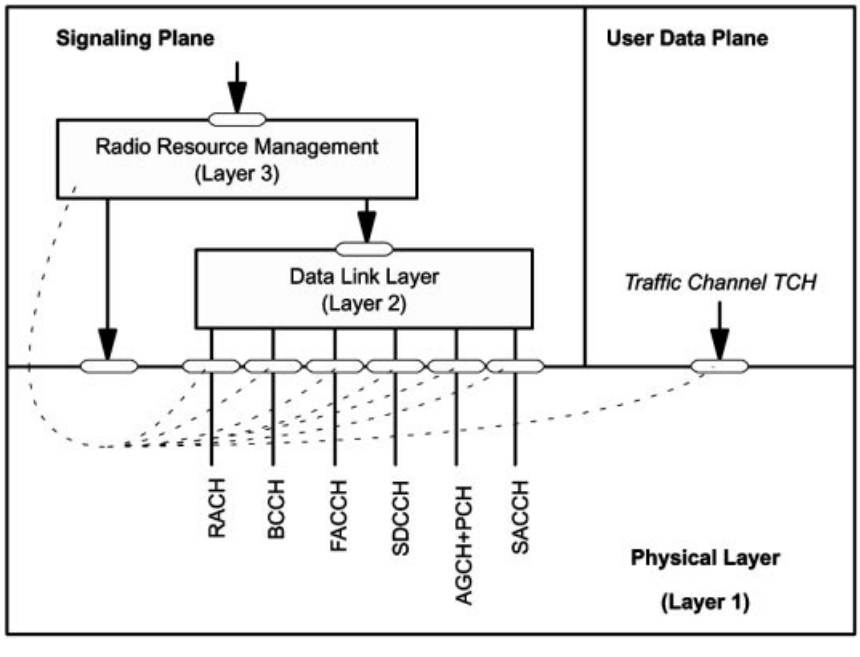

Figure 7.16: Layer 1 service interfaces

procedures with their physical channels. The higher layer protocols access these services at the Layer 1 service interface. The three interfaces of Layer 1 are schematically illustrated in Figure 7.16.

LAPDm protocol frames are transmitted across the service mechanisms of the data link layer interface, and the establishment of logical channels is reported to Layer 2. The communication across this interface is defined by abstract physical layer service primitives. A separate *Service Access Point* (SAP) is defined for each logical control channel (BCCH, PCH + AGCH, RACH, SDCCH, SACCH, FACCH).

Between Layer 1 and the RR sublayer of Layer 3 there is a direct interface. The abstract service primitives exchanged at this interface mostly concern channel assignment and Layer 1 system information, including measurement results of channel monitoring. At the third Layer 1 interface, the traffic channels for user (payload) data are provided.

The service access points (SAP) of Layer 1 as defined in GSM are not genuine service access points in the spirit of OSI. They differ from the PHY-SAPs of the OSI Reference Model insofar as these SAPs are controlled by Layer 3 RR sublayer (layer management, establishment and release of channels) rather than by control procedures in the link layer. Control of Layer 1 SAPs by RR comprises activation and deactivation, configuration, routing and disconnection of physical and logical channels. Furthermore, exchange of measurement and control information for channel monitoring occurs through service primitives.

7.4.1.1 Layer 1 Services

Layer 1 services of the GSM user–network interface are divided into three groups:

- Access capabilities
- Error detection
- Encryption

Layer 1 provides a bit transport service for the logical channels. These are transmitted in multiplexed format over physical channels which consist of elements defined for the transmission on the radio channel (frequency, time slot, hopping sequence, etc.; see Section 7.1). Some physical channels are provided for common (shared) use (BCCH and CCCH), whereas others are assigned to dedicated connections with single mobile stations (dedicated physical channels). The combination of logical channels used on a physical channel can vary over time, e.g. TCH + SACCH/FACCH replaced by SDCCH + SACCH (see Table 5.4).

The GSM standard distinguishes explicitly between access capabilities for dedicated physical channels and for common physical channels BCCH/CCCHs. Dedicated physical channels are established and controlled by Layer 3 RR management. During the operation of a dedicated physical channel, Layer 1 continuously measures the signal quality of the used channel and the quality of the BCCH channels of the neighboring base stations. This measurement information is passed to Layer 3 in measurement service primitives MPH. In idle mode, Layer 1 selects the cell with the best signal quality in cooperation with the RR sublayer based on the quality of the BCCH/CCCH (cell selection).

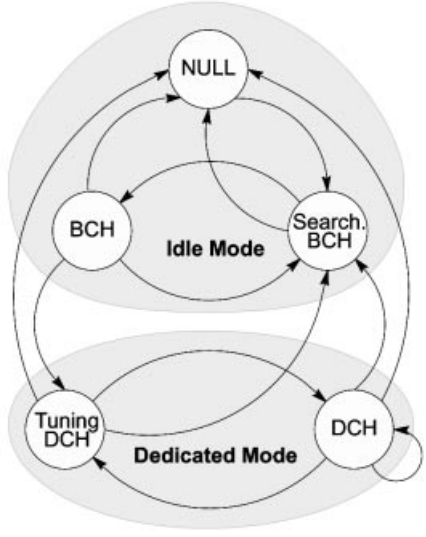

Figure 7.17: State diagram of a mobile station's physical layer

GSM Layer 1 offers an error-protected bit transport service and therefore also error detection and correction mechanisms. To do this, error-correcting and error-detecting coding mechanisms are provided (see Section 6.2). Frames recognized as faulty are not passed up to Layer 2. Furthermore, security-relevant functions like encryption of user data is implemented in Layer 1 (see Section 6.3).

7.4.1.2 Layer 1: Procedures and Peer-to-Peer Signaling

GSM defines and distinguishes between two operational modes of a mobile station: *idle mode* and *dedicated mode* (Figure 7.17). In idle mode, the mobile station is either powered off (state NULL) or it searches for or measures the BCCH with the best signal quality (state SEARCHING BCH), or is synchronized to a specific base station's BCCH and ready to perform a random access procedure on the RACH for requesting a dedicated channel in state BCH (see Section 5.5.4).

In state TUNING DCH of the *dedicated mode*, the mobile station occupies a physical channel and tries to synchronize with it, which will eventually result in transition to state

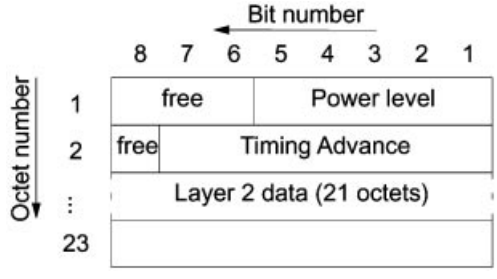

Figure 7.18: Format of an SACCH block

DCH. In this state, the MS is finally ready to establish logical channels and switch them through. The state transitions of Layer 1 are controlled by MPH service primitives of the RR interface, i.e. directly from the Layer 3 RR sublayer of the signaling protocol stack.

Layer 1 defines its own frame structure for the transport of signaling messages, which occur as LAPDm frames at the SAP of the respective logical channel. Figure 7.18 shows the format of an SACCH block as an example, which essentially contains 21 octets of LAPDm data.

Furthermore, the SACCH frame contains a kind of protocol header which carries the current power level and the value of the timing advance. This header is omitted in the other logical channels (FACCH, SDCCH, CCCH, BCCH) which contain only LAPDm PDUs.

7.4.2 Layer 2 Signaling

The LAPDm protocol is the data link protocol for signaling channels at the air interface. It is similar to HDLC. It provides two operational modes:

- Unacknowledged operation
- Acknowledged operation

In the *unacknowledged operation* mode, data is transmitted in UI frames (unnumbered information) without acknowledgement; there is no flow control or L2 error correction. This operational mode is allowed for all signaling channels, except for the RACH which is accessed in multiple access mode without reservation or protection.

The *acknowledged operation* mode provides protected data service. Data is transmitted in I frames (information) with positive acknowledgement. Error protection through retransmission (ARQ) and flow control are specified and activated in this mode. This mode is only used on DCCH channels.

In LAPDm, the *Connection End Points* (CEPs) of L2 connections are labeled with *Data Link Connection Identifiers* (DLCIs), which consist of two elements:

- The Layer 2 *Service Access Point Identifier* (SAPI) is transmitted in the header of the L2 protocol frame.
- The physical channel identifier on which the L2 connection is or will be established, is the real Layer 2 *Connection End Point Identifier* (CEPI). The CEPI is locally administered and not communicated to the L2 peer entity. (The terminology of the GSM standard is somewhat inconsistent in this case – what is really meant is the respective logical channel. The physical channels from the viewpoint of LAPDm are the logical channels of GSM, rather than the physical channels defined by frequency/time slot/hopping sequence.)

When a Layer 3 message is transmitted, the sending entity chooses the appropriate SAP and CEP. When the service data unit SDU is handed over at the SAP, the chosen CEP is given to the L2 entity. Conversely, when receiving an L2 frame, the appropriate L2-CEPI can be determined from the physical/logical channel identity and the SAPI in the frame header.

Table 7.2: Logical channels, operational modes and Layer 2 SAPIs

Logical channel	SAPI = 0	SAPI = 3
BCCH	Unacknowledged	–
CCCH	Unacknowledged	–
SDCCH	Unacknowledged and acknowledged	Unacknowledged and acknowledged
SACCH assoc. with SDCCH	Unacknowledged	–
SACCH assoc. with TCH	Unacknowledged	Unacknowledged and acknowledged
FACCH	Unacknowledged and acknowledged	–

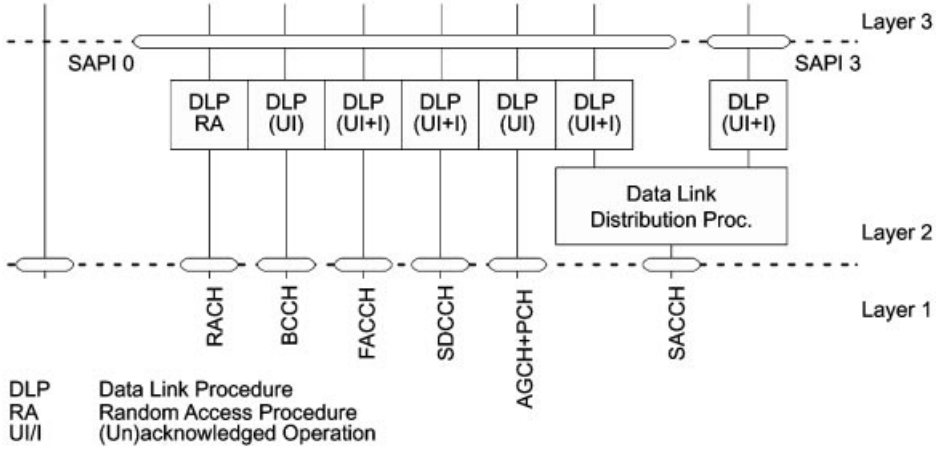

Figure 7.19: Sample configuration of the MS data link layer

Specific SAPI values are reserved for the certain functions:

- SAPI = 0 for signaling (CM, MM, RR)
- SAPI = 3 for SMS

In the control plane, these two SAPI values serve to separate signaling messages from packet-oriented user data (short messages). Further functions needing a new SAPI value can be defined in future versions of the GSM standard.

An LAPDm entity is established for each of the pertinent physical/logical channels. For some of the channel/SAPI combinations only a subset of the LAPDm protocol is needed (e.g. unacknowledged operation), and some channel/SAPI combinations are not supported (Table 7.2). These LAPDm entities perform the *Data Link* procedure, i.e. the functions of the L2 peer-to-peer communication as well as the service primitives between adjacent layers. Segmentation and reassembly of Layer 3 messages is also included.

Further Layer 2 procedures are the *Distribution Procedure* and the *Random Access* (RA) procedure. The distribution procedure is needed if multiple SAPs are associated with one physical/logical channel. It performs the distribution of the L2 frames received on one channel to the respective data link procedure, or the priority-controlled multiplexing of L2 frames from multiple SAPs onto one channel. The random access procedure is used on the random access channel (RACH); it deals with the random controlled retransmission of random access bursts, but it does not perform any error protection on the unidirectional RACH.

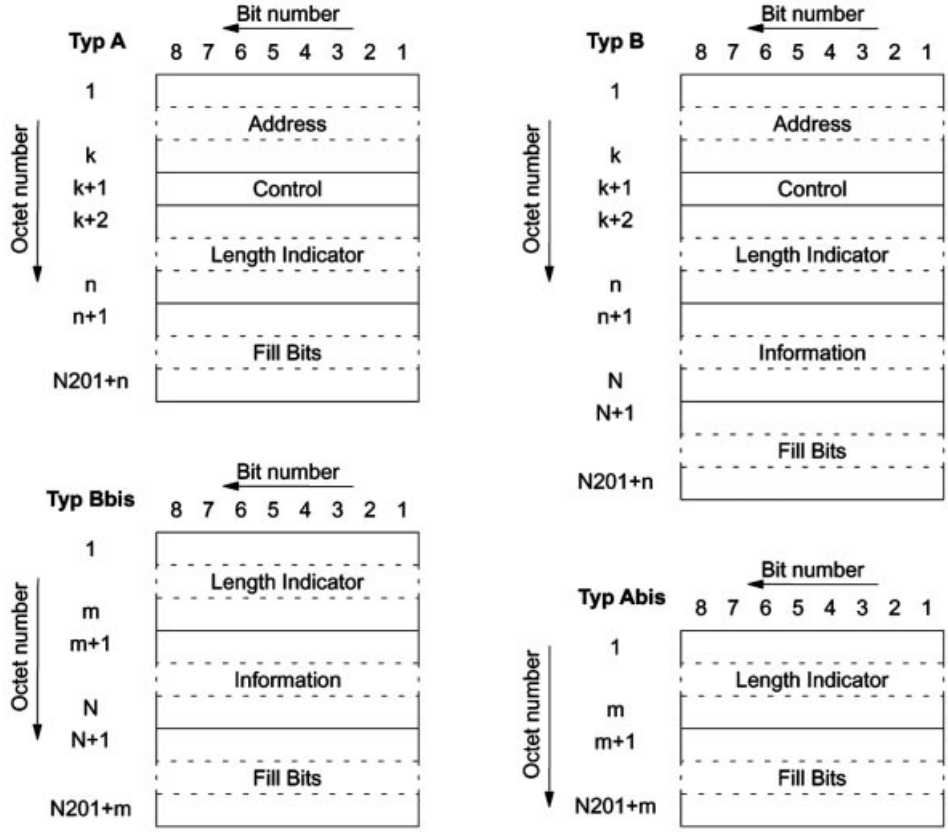

Figure 7.20: LAPDm frame formats

For certain aspects of RR, the protocol logic of Layer 3 has to have direct access to the services of Layer 1. Especially, this is needed for functions of *Radio Subsystem Link Control*, i.e. for channel measurement, transmitter power control, and timing advance.

A possible link layer configuration of an MS is shown in Figure 7.19. The base station has a similar configuration with one PCH + AGCH, SDCCH and SACCH/FACCH for each active mobile station.

Figure 7.20 shows the different types of protocol data frames used for communication between L2 peer entities in MS and BTS. Frame formats A and B are used on the

SACCH, FACCH and SDCCH channels, depending upon whether the frame has an information field (Type B) or not (Type A). For unacknowledged operation (BCCH, PCH, AGCH), format types Abis and Bbis are used on channels with SAPI = 0. The Abis format is used when there is no information to be transmitted on the respective logical channel.

In contrast to HDLC, LAPDm frames have no flag to designate beginning and end of a frame, rather the delineation of frames is done as in RLP at the link level (see Section 7.2.3) through the fixed-length block structure of Layer 1. The maximum number of octets N201 per information field depends on the type of logical channel (Table 7.3). The end of the information field is given by a *Length Indicator*, a value of less than N201 indicates that the frame has to be supplemented with fill bits to the full length. In the case of an SACCH channel, for example, this yields a fixed-length LAPDm packet of 21 octets. Combined with the fields for transmitter power control and timing advance, an SACCH block of Layer 1 is thus 23 octets long.

The address field may have a variable length, however; for use on control channels it consists of exactly one octet. Besides other fields, this octet contains an SAPI (3 bits) and the *Command/Response* (C/R) flag known from HDLC. In LAPDm, the coding of the control field with sending and receiving sequence numbers and the state diagram describing the protocol procedures are almost identical to HDLC [31]. Some additional parameters are required at the service interface to Layer 3; for example, a parameter CEP designating the desired logical channel. Furthermore, the LAPDm protocol has some simplifications or peculiarities with regard to HDLC:

- The sending window size is restricted to $k = 1$.
- The protocol entities should be implemented in such a way that the state RECEIVER BUSY is never reached. Thus RNR packets can be safely ignored. The HDLC polling procedure for state inquiry of the partner station need not be implemented in LAPDm.
- Connections to SAPI = 0 are always initiated by the mobile station.

In addition, the repetition timer T200 and the maximum number of allowed repetitions N200 have been adapted to the special needs of the mobile channel. In particular, they have their own value determined by the type of logical channel.

Table 7.3: Logical channels and the maximum length of the LAPDm information field

Logical channel	N201
SACCH	18 octets
SDCCH, FACCH	20 octets
BCCH, AGCH, PCH	22 octets

7.4.3 Radio Resource Management

The procedures for *Radio Resource Management* (RR) are the basic signaling and control procedures at the air interface. They handle the assignment, allocation and administration

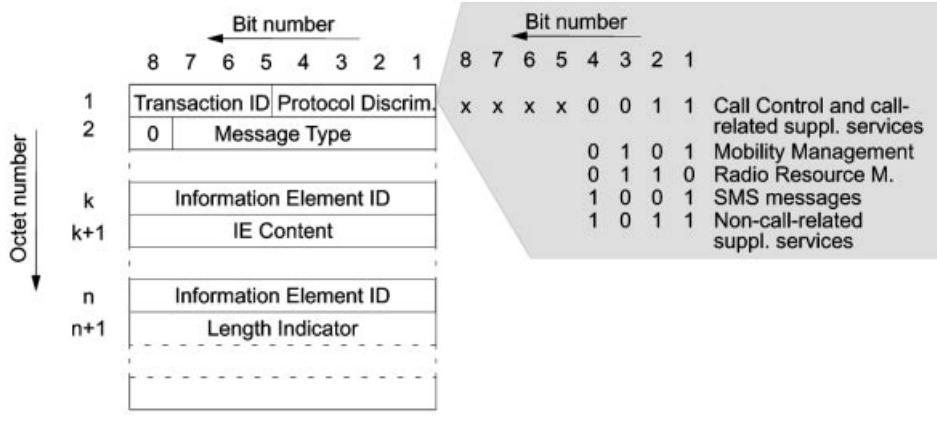

Figure 7.21: Format of a Um signaling message (Layer 3)

of radio resources, the acquisition of system information from broadcast channels (BCCH) and the selection of the cell with the best signal reception (see cell selection in Section 5.5.4.1). Accordingly, the RR procedures and pertinent messages (Table 7.4) are defined for idle mode as well as for setting up, maintaining, and taking down of RR connections.

Figure 7.21 shows the format of RR messages, which is uniform for all three Layer 3 signaling sublayers (CM, MM, RR). Each Layer 3 message contains a protocol discriminator in the first octet, which allows association of messages with the respective sublayer or service access point (Figure 7.12). The uppermost four bits of the first octet also contain a *Transaction ID*, which enables an MS to perform several signaling transactions in parallel. The *Message Type* (MT) is registered in the lower seven bits of the second octet (see also Tables 7.4–7.6). Otherwise, Layer 3 messages consist of *Information Elements* (IEs) of fixed or variable length; a *Length Indicator* (LI) is added for variable-length messages.

In idle mode, the MS is reading continuously the BCCH information and conducts periodic measurements of the signaling strength of the BCCH carriers in order to be able to select the current cell (see Section 5.5.4). In this state, there is no exchange of signaling messages with the network. The data required for RR and other signaling procedures is collected and stored: the list of neighboring BCCH carriers, thresholds for RR algorithms, CCCH configurations, information about the use of RACH and PCH, etc. This information is broadcast by the BSS on the BCCH (SYSTEM INFORMATION, Types 1–4) and therefore is available to all mobile stations currently in the cell. Also important is the periodic monitoring of the paging channel (PCH) so that paging calls are not lost. For this purpose, the BSS is sending on all paging channels of a cell continuously valid Layer 3 messages (PAGING REQUEST) which the MS can decode and recognize if its address is paged.

Connection Setup and Release – Each exchange of signaling messages with the network (BSS, MSC) requires an RR connection and the establishment of an LAPDm connection between MS and BTS. Setting up the RR connection can be initiated by the

Table 7.4: RR messages

Category	Message	Logical channel	Direction	MT-code
Channel establishment	Additional assignment	DCCH	N → MS	00111011
	Immediate assignment	CCCH	N → MS	00111111
	Immediate assignment extended	CCCH	N → MS	00111001
	Immediate assignment rejected	CCCH	N → MS	00111010
Ciphering	Ciphering mode command	DCCH	N → MS	00110101
	Ciphering mode complete	DCCH	MS → N	00110010
Handover	Assignment command	DCCH	N → MS	00101110
	Assignment complete	DCCH	MS → N	00101001
	Assignment failure	DCCH	MS → N	00101111
	Handover access	DCCH	MS → N	–
	Handover command	DCCH	N → MS	00101011
	Handover complete	DCCH	MS → N	00101100
	Handover failure	DCCH	MS → N	00101000
	Physical information	DCCH	N → MS	00101101
Channel release	Channel release	DCCH	N → MS	00001101
	Partial release	DCCH	N → MS	00001010
	Partial release complete	DCCH	MS → N	00001111
Paging	Paging request, Type 1/2/3	PCH	N → MS	00100xxx
	Paging response	DCCH	MS → N	00100111
System information	System information Type 1/2/3/4	BCCH	N → MS	00011xxx
	System information Type 5/6	SACCH	N → MS	00011xxx
Miscellaneous	Channel mode modify	DCCH	N → MS	00010000
	Channel mode modify acknowledge	DCCH	MS → N	00010111
	Channel request	RACH	MS → N	–
	Classmark change	DCCH	MS → N	00010110
	Frequency redefinition	DCCH	N → MS	00010100
	Measurement report	SACCH	MS → N	00010101
	Synchronization channel information	SCH	N → MS	–
	RR-status	DCCH	MS ↔ N	00010010

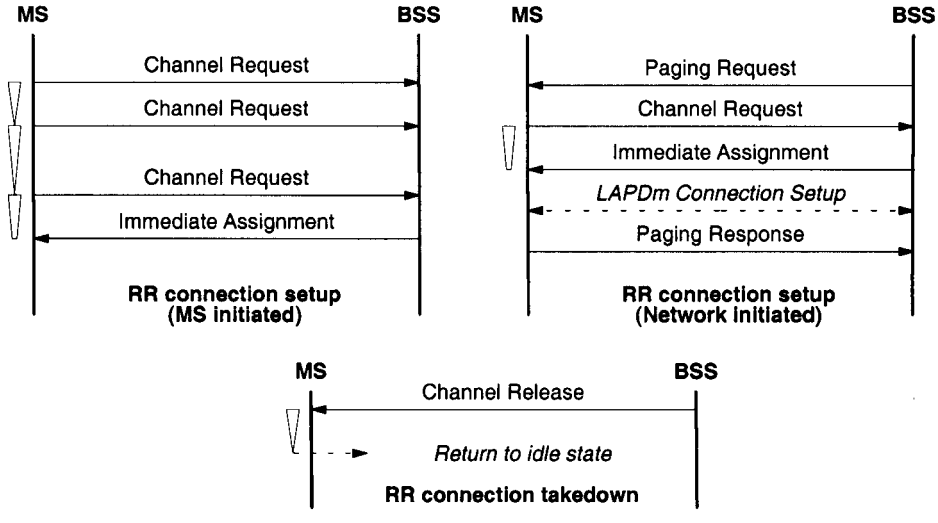

Figure 7.22: RR connection setup and takedown

network or the MS (Figure 7.22). In either case, the MS sends a channel request (CHAN-QUEST) on the RACH in order to get a channel assigned on the AGCH (immediate assignment procedure). There is also a procedure to deny a channel request (immediate assignment reject).

If the network does not immediately answer to the channel request, the request is repeated using the Aloha method with a random number controlled timer (Figure 7.22). In the case of a network-initiated connection, this procedure is preceded by a paging call (PAGING REQUEST) to be answered by the mobile station (PAGING RESPONSE). After an RR connection has been successfully completed, the higher protocol layers (CM, MM) can receive and transmit signaling messages at SAPI 0.

In contrast to the setup of connections, the release is always initiated by the network (CHANNEL RELEASE). Reasons for the release of the channel could be end of the signaling

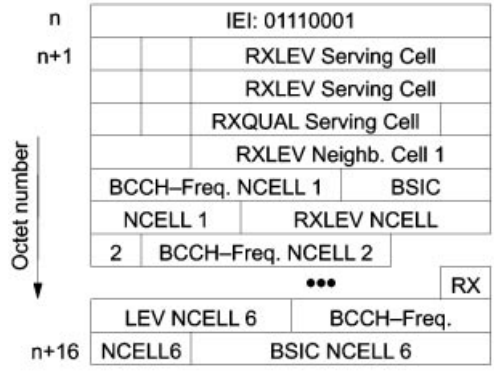

Figure 7.23: Measurement result (information element)

transaction, too many errors, removal of the channel in favor of a higher priority call (e.g. emergency call), or end of a call. After receiving the channel release command, the mobile station assumes the idle state following a brief waiting period (Figure 7.22).

Once an RR connection has been set up, the mobile station has either an SDCCH or a TCH with associated SACCH/FACCH available for exclusive bidirectional use. On the SACCH, data must be sent continuously (see also Section 5.5.3), i.e. the MS keeps sending current channel measurements (MEASUREMENT REPORT, see Section 5.5.1.2) if no other signaling messages need to be sent. In the other direction, the BSS keeps sending system information (SYSTEM INFORMATION, alternating between Type 5 and Type 6). The information element with the coded measurement results contains the following among other data: RXLEV and RXQUAL of the serving cell as well as RXLEV and carrier frequency of up to six neighboring cells as well as their BSICs (Figure 7.23). The system information sent by the BSS on the SACCH contains first information about the neighbor cells and their BCCH (Type 5), and second, information about the current cell (Type 6) such as *Cell Identity* (CI) and the current *Location Area Identifier* (LAI).

Channel Change – For established RR connections, a channel change within the cell can be performed (dedicated channel assignment, Figure 7.24) to change the configuration of the physical channel in use. Such a channel change can be requested by higher protocol layers, or it can be requested by the RR sublayer; however, it is always initiated by the network. When the mobile station receives an ASSIGNMENT COMMAND, the transmission of all signaling messages is suspended, the LAPDm connection is taken down, the traffic channel, if existent, is switched off, and the old channel is deactivated. After activation of

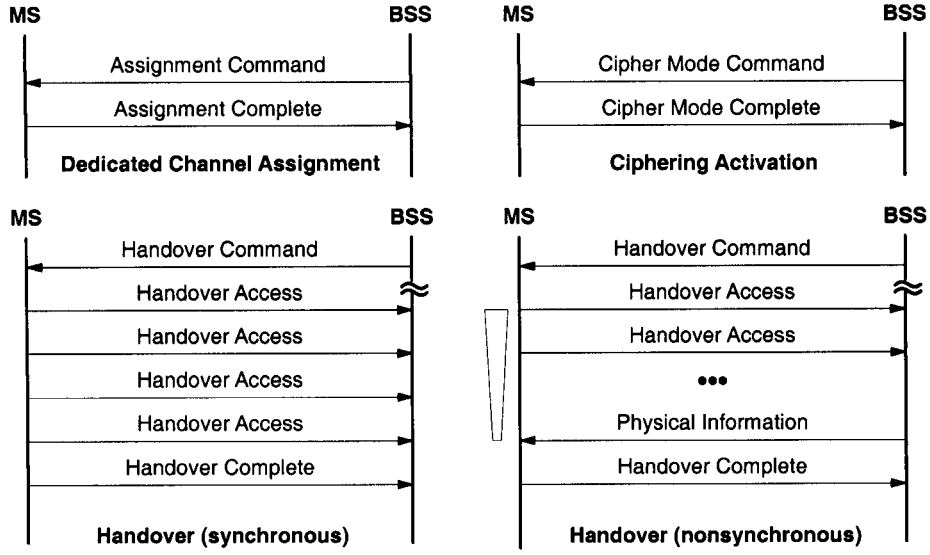

Figure 7.24: Channel change, encryption, and handover

the new physical channel and a successful establishment of a new LAPDm connection (Layer 2), the held-back signaling messages can be transmitted.

Handover – A second signaling procedure to change the physical channel configuration of an established RR connection is the handover procedure, which is also initiated only from the network side and, for example, becomes necessary if the current cell is left. In contrast to ASSIGNMENT COMMAND, a HANDOVER COMMAND contains not only the new channel configuration but also information about the new cell (e.g. BSIC and BCCH frequency), the procedure variant to establish a physical channel (asynchronous or synchronous handover, Figure 7.24), and a handover reference number.

Having received a HANDOVER COMMAND on the FACCH, the mobile station terminates the LAPDm connection on the old channel, interrupts the connection, deactivates the old physical channel, and finally switches over to the channel newly assigned in the HANDOVER COMMAND. On the main DCCH (in this case FACCH), the mobile station sends the unencrypted message HANDOVER ACCESS in an access burst (Figure 5.6, coding as on the RACCH, see Section 6.2) to the base station. Even though this is a message on the FACCH, an access burst is used because the mobile station at this time does not yet know the complete synchronization information. The eight data bits of the access burst contain the handover reference of the handover command. The way in which these access bursts are transmitted depends on whether both cells have synchronized their TDMA transmission or not.

In the case of existing synchronization, the access burst (HANDOVER ACCESS) is sent in exactly four subsequent time slots of the main DCCH (FACCH). Thereafter, the mobile station activates the new physical channel in both directions, establishes an LAPDm connection, activates encryption, and finally sends a message HANDOVER COMPLETE to the BSS. In the nonsynchronous case, the mobile station repeats the access burst until either a timer expires (handover failed) or until the base station answers with an RR message PHYSICAL INFORMATION which contains the currently needed timing advance and this way enables the establishment of the new RR connection.

Activation of Ciphering – Another important RR procedure is the activation of ciphering. This is done by the BSS with the CIPHER MODE COMMAND, which also indicates that the BTS has activated its deciphering function. Having received the CIPHER MODE COMMAND, the mobile station activates ciphering as well as deciphering and sends the answer CIPHER MODE COMPLETE already in enciphered form. If the BTS is able to correctly decipher this message, the ciphering mode has been successfully established.

Other Signaling Procedures – In addition, there are a number of less significant signaling procedures defined, such as *Frequency Redefinition*, *Additional Assignment*, *Partial Release*, or *Classmark Change*. The first one concerns the change of the MA; see Section 5.2.3. The next two deal with a change of the physical channel configuration. With the last message, CLASSMARK CHANGE, the mobile station reports that it now belongs to a new power class (see Table 5.8), which can be achieved by installing a commercially available power booster kit, for example.

Table 7.5: MM messages

Category	Message	Direction	MT
Registration	IMSI detach indication	MS → N	0x000001
	Location updating accept	N → MS	0x000010
	Location updating reject	N → MS	0x000100
	Location updating request	MS → N	0x001000
Security	Authentication reject	N → MS	0x010001
	Authentication request	N → MS	0x010010
	Authentication response	MS → N	0x010100
	Identity request	N → MS	0x001000
	Identity response	MS → N	0x001001
	TMSI reallocation command	N → MS	0x001010
	TMSI reallocation complete	MS → N	0x001011
Connection	CM service accept	MS ↔ N	0x100001
management	CM service reject	N → MS	0x100010
	CM service request	MS → N	0x100100
	CM reestablishment request	MS → N	0x101000
Miscellaneous	MM-status	MS ↔ N	0x110001

7.4.4 Mobility Management

The main task of *Mobility Management* (MM) is to support the mobility of the mobile station; for example, by reporting the current location to the network or verifying the subscriber identity. Another task of the MM sublayer is to offer MM connections and associated services to the CM sublayer above. The message format for MM messages is the uniform Layer 3 signaling message format (Figure 7.21). MM has its own protocol discriminator, and the MM messages are marked with a type code (MT, Table 7.5).

All MM procedures presume an established RR connection, i.e. a dedicated logical channel must be assigned with an established LAPDm connection in place, before MM transactions can be performed. These transactions occur between MS and MSC, i.e. messages are passed through the BSS transparently without interpretation and forwarded to the MSC with the DTAP transport mechanism. The MM procedures are divided into three categories: *common*, *specific*, and *MM Connection Management*. Whereas *common* procedures can always be initiated and executed as soon as an RR connection exists, Specific procedures exclude one another, i.e. they cannot be processed simultaneously or during an MM connection. Conversely, an MM connection can only be set up if no Specific procedure is running.

7.4.4.1 Common MM Procedures

The *common* MM procedures are summarized in Figure 7.25. Besides the *IMSI Detach* procedure, they are all initiated from the network side. An important role for the protection of subscriber identity is held by the *TMSI Reallocation* procedure. If the confidentiality of a subscriber's identity IMSI is to be protected (an optional network service), the signaling procedures across the air interface use the TMSI instead of the IMSI. This TMSI has only local significance within a *Location Area* and must be used together with the LAI for the unique identification of a subscriber.

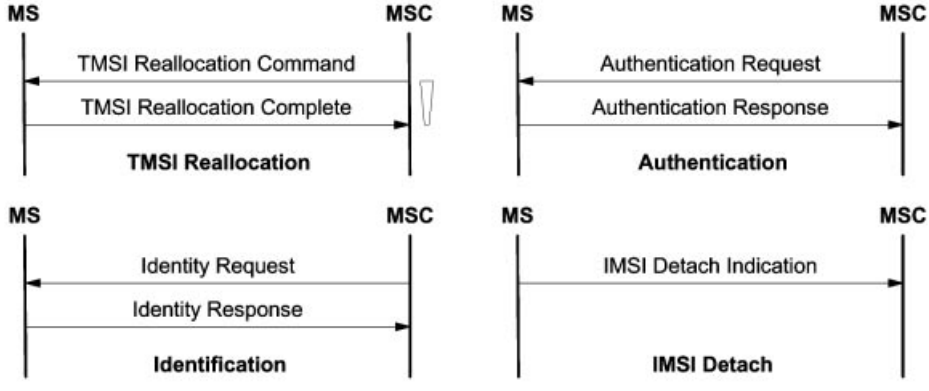

Figure 7.25: MM signaling procedures of category *common*

For further protection, the TMSI can also be repeatedly reallocated (TMSI reallocation) which must be done at the latest when the location area changes. Otherwise this TMSI change is left as an option to the network operator, but it can be performed any time after an encrypted RR connection to the mobile station has been set up. The TMSI reallocation is either executed explicitly as a standalone procedure, or implicitly from other procedures using the TMSI, e.g. the location update. In the case of explicit TMSI reallocation, the network sends a TMSI REALLOCATION COMMAND with the new TMSI and the current LAI on an encrypted RR connection to the mobile station (Figure 7.25).

The MS stores the TMSI and LAI in nonvolatile SIM storage and acknowledges it with the message TMSI REALLOCATION COMPLETE. If this message reaches the MSC before the timer expires, the timer is cancelled, and the TMSI is valid. However, if the timer expires before the acknowledgement arrives, the procedure is repeated. If it fails a second time, the old as well as the new TMSI are barred for a certain time interval, and the IMSI is used for paging the mobile station. If the mobile station answers a paging call, TMSI reallocation is started again. Furthermore, the TMSI is assumed valid in spite of failed reallocation if it is used by the MS in subsequent transactions.

Two more *common* procedures are used for the identification of a mobile station or a subscriber (*identification* procedure) and for the verification of the respective identity (*authentication* procedure). For the identification of a mobile station, there is the equipment identity IMEI as well as the subscriber identity IMSI which is assigned to the MS

through the SIM card. The network may request these two identity parameters at any time from the mobile station with an IDENTITY REQUEST. Therefore the mobile station must be able at any time to supply these identity parameters to the network with an IDENTITY RESPONSE message.

Authentication also assigns a new key for encryption of user payload data. This procedure is started from the network with an AUTHENTICATION REQUEST message. A mobile station must be able to process this request at any time during an RR connection. The MS calculates the new key Kc for the encryption of user data from the information obtained during authentication which is locally stored, and it also calculates authentication information to prove its identity without doubt. This authentication data is transmitted with an AUTHENTICATION RESPONSE message to the MSC which evaluates them. If the answer is not valid and the authentication has therefore failed, further processing depends on whether the IMSI or TMSI was used. In the case of TMSI, the network can start the identification procedure. If the implied IMSI is not identical to the one associated with the TMSI by the network, the authentication is restarted with new correct parameters. If the two IMSIs agree, or the IMSI was used a priori by the MS, the authentication has failed, which is indicated to the MS with an AUTHENTICATION REJECT message. This forces the MS to cancel all the assigned identity and security parameters (TMSI, LAI, Kc) and to enter idle mode, so that only simple cell selection and emergency calls are enabled.

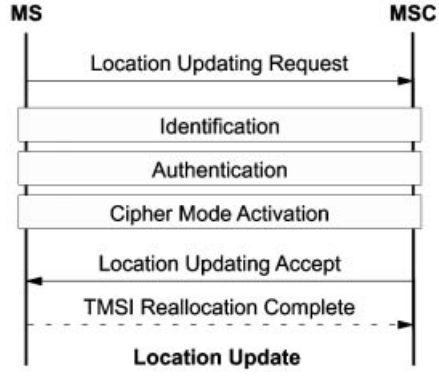

Figure 7.26: MM signaling procedures of category *specific*

If the mobile station is powered off or the SIM has been removed, the MS is not reachable because the MS does not monitor the paging channel, and calls cannot be delivered. In order to relieve the paging load on the BSS caused by unnecessary paging calls, a network operator can optionally request an explicit deregistration message from the mobile station, which is not normally required. This option is signaled by setting a flag on the BCCH (SYSTEM INFORMATION Type 3) and on the SACCH (SYSTEM INFORMATION Type 8). If the flag is set, the MS sends an IMSI DETACH INDICATION message when it powers off or when the SIM is removed, which allows the network to mark the MS as inactive. The IMSI detach procedure is the only *common* procedure that cannot be started

at an arbitrary time even during a *specific* procedure. Is start has to be delayed until any *specific* procedure has ended.

7.4.4.2 Specific MM Procedures

In GSM systems, updating of current location information is the sole responsibility of the mobile station. Using the information broadcast on BCCH channels, it has to recognize any change in the current location area and report it to the network, so that the databases HLR and VLR can be kept up to date. The generic structure of a location update is shown in Figure 7.26: The mobile station requests to update its current location information in the network with a LOCATION UPDATING REQUEST. If this can be done successfully, the network acknowledges this with a message LOCATION UPDATING ACCEPT. In the course of a location update, the network can ask for the mobile station's identity and check it out (identification and authentication). If the service "confidential subscriber identity" has been activated, a new TMSI assignment is a permanent component of the location update. In this case, enciphering of user data on the RR connection is activated, and the new TMSI is transmitted together with the message LOCATION UPDATING ACCEPT and is acknowledged with the message REALLOCATION COMPLETE. Periodic updating of location information can be used to indicate the presence of the mobile station in the network. For this purpose, the mobile station keeps a timer which periodically triggers a *location update* procedure. If this option is in use, the timer interval to be used is broadcast on the BCCH (SYSTEM INFORMATION Type 3). The procedure *IMSI Attach* is the converse of the procedure *IMSI Detach* (see Figure 7.25) and is executed as a special variant of the location update if the network requires this. However, the mobile station executes an *IMSI Detach* only if the LAI broadcast on the BCCH agrees with the LAI stored in the MS. If the stored LAI and received LAI differ, a normal *location update* procedure is executed.

7.4.4.3 MM Connection Management

Finally, there is a third category of MM procedures which are needed for the establishment and the operation of MM connections (Figure 7.27). An MM connection is established on request from the CM sublayer above and serves for the exchange of messages between CM entities, where each CM entity has its own MM connection (Figure 7.12).

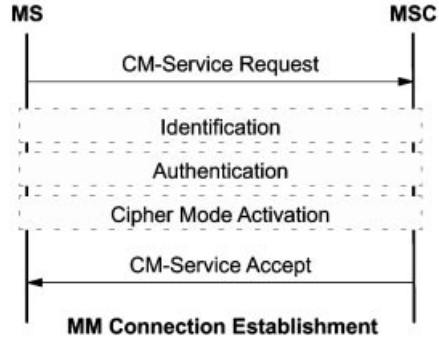

Figure 7.27: MM signaling procedures of category *MM connection management*

The procedures for the setup of MM connections are different depending on whether initiation occurs from the network or the mobile station.

Setting up an MM connection from the side of the mobile station presumes the existence of an RR connection, but a single RR connection can be used by multiple MM connections. The MM connection can only be established if the mobile station has executed a successful location update in the current location area. An exception is an emergency call, which is possible at any time. If there is a request from the CM sublayer for an MM connection, it may be delayed or rejected if there are *specific* procedures active, depending on implementation. If the MM connection can be established, the mobile station

Table 7.6: CC messages for circuit-switched connections

Category	Message	Direction	MT
Call establishment	Alerting	N → MS	0x000001
	Call confirmed	MS → N	0x001000
	Call proceeding	N → MS	0x000010
	Connect	N ↔ MS	0x000111
	Connect acknowledge	N ↔ MS	0x001111
	Emergency setup	MS → N	0x001110
	Progress	N → MS	0x000011
	Setup	N ↔ MS	0x000101
Call Information Phase	Modify	N ↔ MS	0x010111
	Modify complete	N ↔ MS	0x011111
	Modify reject	N ↔ MS	0x010011
	User information	N ↔ MS	0x010000
Call Clearing	Disconnect	N ↔ MS	0x100101
	Release	N ↔ MS	0x101101
	Release complete	N ↔ MS	0x101010
Miscellaneous	Congestion control	N ↔ MS	0x111001
	Notify	N ↔ MS	0x111110
	Start DTMF	MS → N	0x110101
	Start DTMF acknowledge	N → MS	0x110010
	Start DTMF Reject	N → MS	0x110111
	Status	N ↔ MS	0x111101
	Status enquiry	N ↔ MS	0x110100
	Stop DTMF	MS → N	0x110001
	Stop DTMF acknowledge	N → MS	0x110010

sends the message CM-SERVICE REQUEST to the network. This message contains information about the mobile subscriber (IMSI or TMSI) as well as information about the requested service (outgoing voice call, SMS transfer, activation or registration of a supplementary service, etc.). Depending on these parameters, the network can execute any *common* MM procedure (except *IMSI Detach*) or activate enciphering of user data. If the mobile station receives the message CM-SERVICE ACCEPT or the local message from

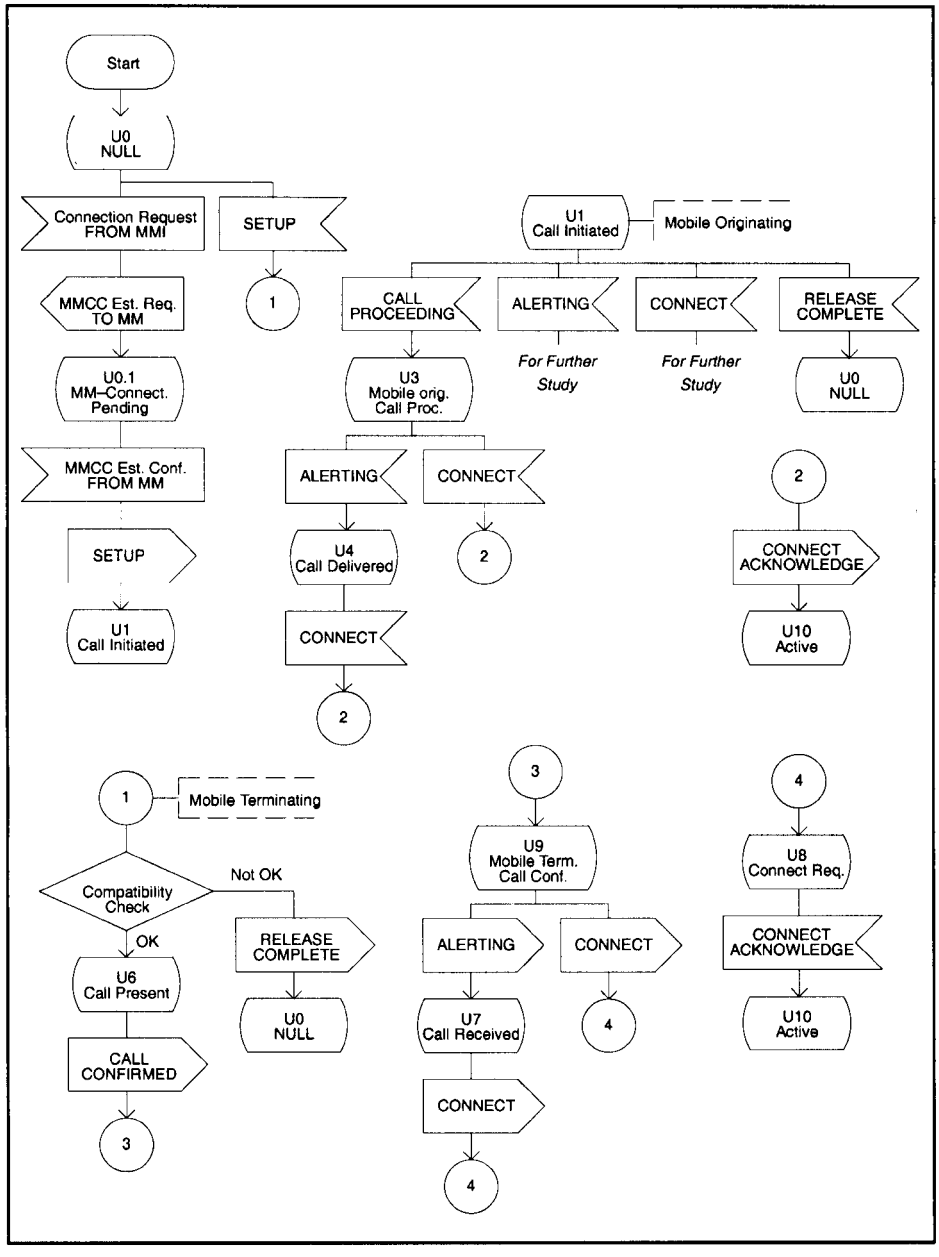

Figure 7.28: Call setup (mobile station): mobile-originating and mobile-terminating

the RR sublayer that enciphering was activated, it treats this as an acceptance of the service request, and the requesting CM entity is informed about the successful setup of an MM connection. Otherwise, if the service request has been rejected by the network, the MS receives a message CM-SERVICE REJECT, and the MM connection cannot be established.

The network-initiated setup of an MM connection does not require an exchange of CM service messages. After successful paging, an RR connection is established, and the sublayer on the network side executes one of the MM procedures if necessary (mostly *location update*) and requests from the RR sublayer the activation of user data encryption. If these transactions are successful, the service requesting CM entity is informed, and the MM connection is established.

7.4.5 Connection Management

Call Control (CC) is one of the entities of *Connection Management* (CM); the CM sublayer is shown in Figure 7.12. It comprises procedures to establish, control, and terminate calls. Several parallel CC entities are provided, such that several parallel calls on different MM connections can be processed. For CC, finite state models are defined both on the mobile side as well as on the network side. The two entities at the MS and MSC sites each instantiate a protocol automaton, and these communicate with each other using the messages in Table 7.6 and the uniform Layer 3 signaling message format (Figure 7.21).

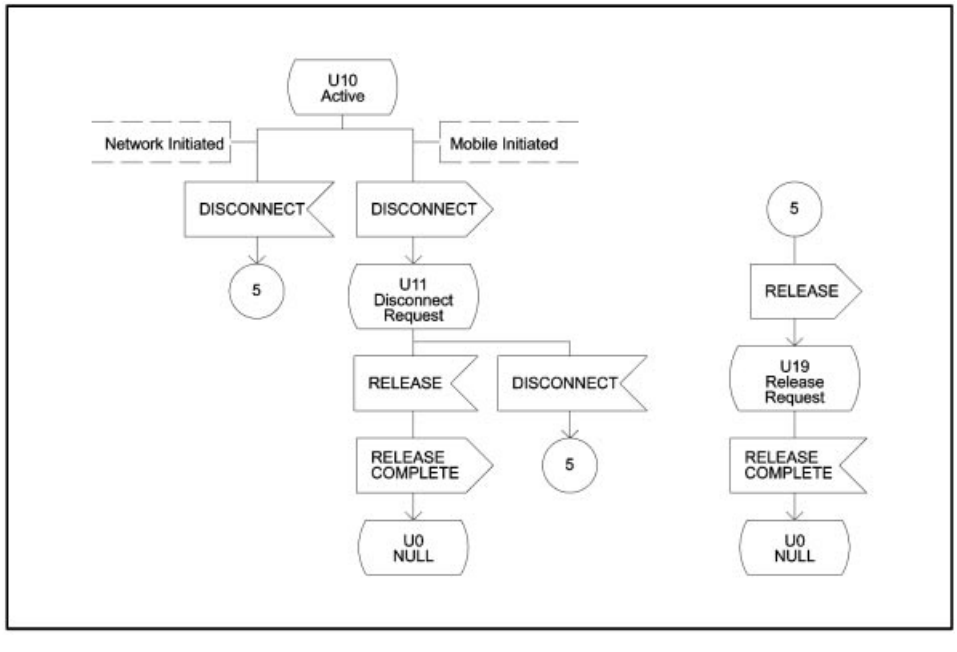

Figure 7.29: Call termination at the mobile station: mobile-initiated and network-initiated

Parts of CC in the mobile station are presented schematically in Figures 7.28 and 7.29. They show the mobile-originating and mobile-terminating setup of a call and the mobile/network initiated call takedown. If there is a desire to call from the mobile station (mobile-originating call), the CC entity first requests an MM connection from the local MM entity, also indicating whether it is a normal or emergency call (MMCC ESTABLISHMENT REQUEST, Figure 7.28). The call to be established on this MM connection requires a special service quality of the MM sublayer.

For a simple call, the mobile station must be registered with the network, whereas this is only optionally required with an emergency call, i.e. an emergency call can also be established on an unenciphered RR connection from a mobile station that is not registered.

After successful establishment of this MM connection and activation of the user data encryption, the service-requesting CC entity is informed (further interactions with the MM entity are not shown in Figure 7.28 for call establishment). The mobile station signals on this connection the desire to connect to the CC entity in the MSC (SETUP). An emergency call is initiated with the message EMERGENCY SETUP; the remaining call setup is identical to the one used for single calls.

The MSC can respond to this connection request in several ways: it can indicate with a message CALL PROCEEDING that the call request has been accepted and that all the necessary information for the setup of the call is available. Otherwise the call request is declined with RELEASE COMPLETE. As soon as the called party is being signaled, the MS receives an ALERTING message; once the called party accepts the call, a CONNECT message is returned which is acknowledged with a CONNECT ACKNOWLEDGE message, thus switching through the call and the associated user data connection. If the call request need not be signaled to the called party and the call can be accepted directly, the ALERT message is omitted. Essentially CC signaling in GSM corresponds to the call setup according to Q.931 in ISDN. In addition, CC in GSM has a number of peculiarities, especially to account for the limited resources and properties of the radio channel. In particular, the call request of the MS can be entered into a queue (call queuing), if there is no immediately free traffic channel (TCH) for the establishment of the call. The maximum waiting time a call may have to wait for assignment of a traffic channel can be adjusted according to operator needs. Furthermore, the point at which the traffic channel is actually assigned can be chosen. For example, the traffic channel can be assigned immediately after acknowledging the call request (CALL PROCEEDING); this is *early assignment*. On the other hand, the call can be first processed completely and the assignment occurs only after the targeted subscriber is being called; this is *late assignment* or *Off-Air Call Setup* (OACSU). The variant OACSU avoids unnecessary allocation of a traffic channel if the called subscriber is not available. The blocking probability for call arrivals at the air interface can be reduced this way. On the other hand, there is the probability that after a successful call request signaling procedure, no traffic channel can be allocated for the calling party before the called party accepts the call, and thus the call cannot be completely switched through and must be broken off.

If a call arrives at the mobile station (mobile-terminating call), an RR connection with the mobile station is established within the MM connection setup (inclusive of paging). Once the MM connection is successfully completed and the encryption is activated, the

call request is signaled to the mobile station with a SETUP message. This message includes information about the requested service, and the mobile station examines first whether it can satisfy the requested service profile (compatibility check). If affirmative, it accepts the call request and signals this to the local subscriber (local generation of call signal). This is finally communicated to the MSC with a CALL CONFIRMED message and an ALERTING message. If the mobile subscriber eventually accepts the call, the call is switched through completely with handshake messages CONNECT and CONNECT ACKNOWLEDGE. If because of the selected service there is no necessity for call request signaling to the called subscriber and the call can be switched through immediately (e.g. with fax call), the mobile station signals the call acceptance (CONNECT) immediately after the message CALL CONFIRMED. Call queuing and OACSU can also be used for mobile-terminating calls. The OACSU variant for mobile-terminating calls allocates a traffic channel only after the call has been accepted by the mobile subscriber with a CONNECT message.

The release of a connection is started with a DISCONNECT message either from the mobile or the network side (mobile-/network-initiated) and is completed with handshake messages RELEASE and RELEASE COMPLETE. If there is a collision of DISCONNECT messages, i.e. if both CC entities send a DISCONNECT simultaneously, they also answer it with a RELEASE so that a secure termination is ensured.

During an established call, two more CC procedures can be employed: *Dual-Tone Multifrequency* (DTMF) signaling and *Incall Modification*. DTMF signaling is an inband signaling procedure, which allows terminals (here mobile stations) to communicate with the respective other side, e.g. answering machines, or configuring special network services, e.g. voice mailboxes in the network. In GSM, DTMF can only be used during a voice connection. With a message START DTMF on the FACCH, the network is told that a key has been pressed, and the release of the key is signaled with a STOP DTMF message (Figure 7.30). Each of these messages is acknowledged by the network (MSC). A

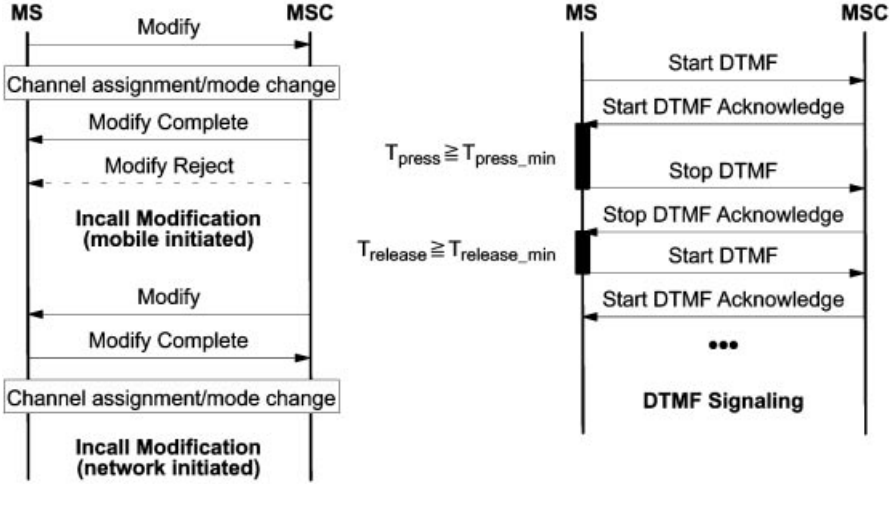

Figure 7.30: DTMF signaling and service change

minimum interval must be maintained between the START/STOP messages (T_{press_min}, $T_{release\_min}$). While a key is depressed at the mobile station, the MSC generates a DTMF tone corresponding to the key code signaled with the START DTMF command. The DTMF tones must be generated within the MSC, since the speech coding in the GSM codec does not permit the pure transmission of DTMF tones in the voice band, and thus DTMF tones generated by the MS would arrive at the other side in distorted form.

Using the incall modification procedure, a service change can be performed, e.g. when speech and fax data is sent in sequence and are alternating during a call (see Chapter 4). A service change is started either from the mobile station or from the network by sending a MODIFY request. This request contains the service and kind of change (returning or non-returning). After sending the MODIFY request, the transmission of user data is halted. If the service change can be performed, this is signaled with a MODIFY COMPLETE message; otherwise the request is denied with a MODIFY REJECT. The service change may necessitate a change in the current physical channel configuration or operating mode. For this

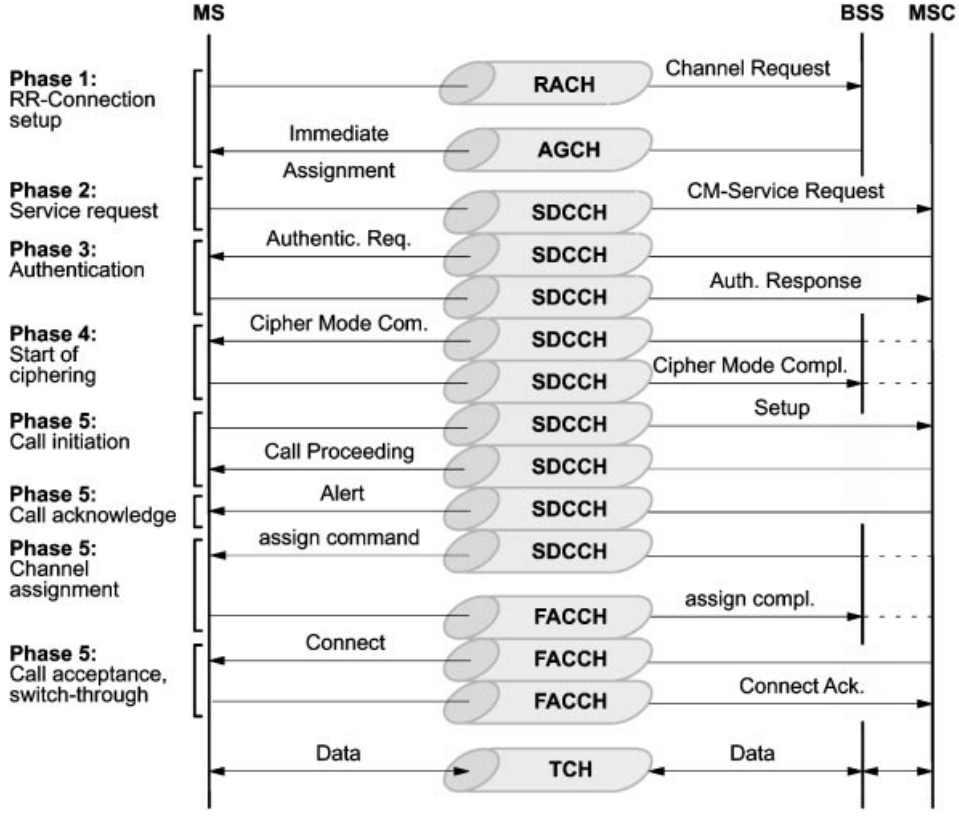

Figure 7.31: Mobile-initiated call setup with OACSU (*late assignment*)

purpose, the MSC will use the respective channel assignment (ASSIGNMENT COMMAND, see Figure 7.24).

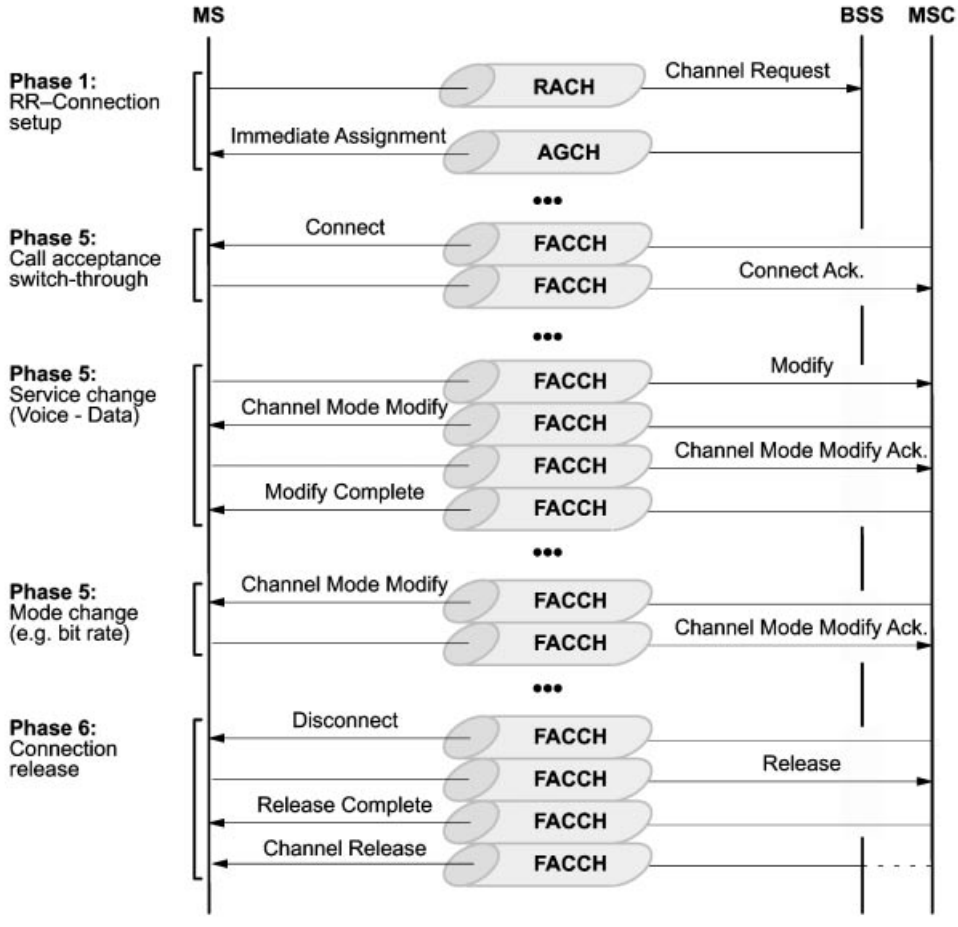

Figure 7.32: In-call modification and call release

7.4.6 Structured Signaling Procedures

The preceding sections have presented the basic signaling procedures of the three sublayers RR, MM, and CM. These procedures have to cooperate in the form of structured procedures for the different transactions. The elements of a structured signaling procedure are

- Phase 1: paging, channel request, assignment of a signaling channel
- Phase 2: service request and collision resolution
- Phase 3: authentication
- Phase 4: activation of user data encryption
- Phase 5: transaction phase
- Phase 6: release and deallocation of the channel

Two examples of structured signaling procedures are presented in Figures 7.31 and 7.32. They show the phases executed for the structured transaction, the terminating entities (MS, BSS, MSC), the respective message, and the logical channel used for the transport of the message. The first example (Figure 7.31) is a mobile-initiated call setup with OACSU – a traffic channel is only assigned after the subscriber of the called station is presented with the call request (ALERTING). The second example (Figure 7.32) shows a service change from voice to data and the modification of the selected data service. Such a modification could for example be the change of transmission rate. Finally the call is released and the traffic channel is deallocated.

7.4.7 Signaling Procedures for Supplementary Services

As can be seen in Figure 7.21, signaling messages to control *Supplementary Services* (SS) are coded with special protocol discriminators: 0011 for call related; 1011 for noncall related. A special set of signaling messages has been defined for their control (Table 7.7). The category *CC Messages* (Table 7.6) consists of the subcategories *Call Information Phase* (message type MT = 0x01tttt) and *Miscellaneous* (message type MT = 0x11tttt). These two message categories are used in two categories of SS procedures: the *separate message approach* and the *common information element procedure*. Whereas the *separate message approach* uses its own messages (HOLD/RETRIEVE, Table 7.7) to activate specific functions, the functions of the *common information element procedure* are handled with a generic FACILITY message. Functions of the first category need synchronization between network and mobile station. The FACILITY category, however, is only used for supplementary services which do not require synchronization. This distinction becomes obvious in the examples of realized supplementary services, which is presented in the following.

Table 7.7: CC messages for supplementary services

Category	Message	Direction	MT
Call information phase	Hold	N ↔ MS	0x011000
	Hold acknowledge	N ↔ MS	0x011001
	Hold reject	N ↔ MS	0x011010
	Retrieve	N ↔ MS	0x011100
	Retrieve acknowledge	N ↔ MS	0x011101
	Retrieve reject	N ↔ MS	0x011110
Miscellaneous	Facility	N ↔ MS	0x111010
	Register	N ↔ MS	0x111011

The messages of the *separate message approach* can be used during the call information phase to realize supplementary services like hold, callback, or call waiting. Figure 7.33

shows examples. A completely established call (call reference CR: 1 in Figure 7.33) can be put into the hold state from either one of the two partner entities.

To perform this supplementary service, it is initiated with a HOLD message. The MSC interrupts the connection and indicates with a HOLD message to the partner entity that the call is in the hold state. On each call segment this fact is acknowledged with a HOLD ACKNOWLEDGE message, which leads to both the requesting mobile station and the MSC to cut the traffic channel. The mobile station which caused the hold state to be entered can now establish another call (CR: 2 in Figure 7.33) or accept a call that may be coming in. Using another handshake HOLD/HOLD ACKNOWLEDGE, this call could be put into hold state too, and there could be switching between both held calls (brokering). For this purpose, a held call (CR: 1 in Figure 7.33) can be reactivated with a RETRIEVE message

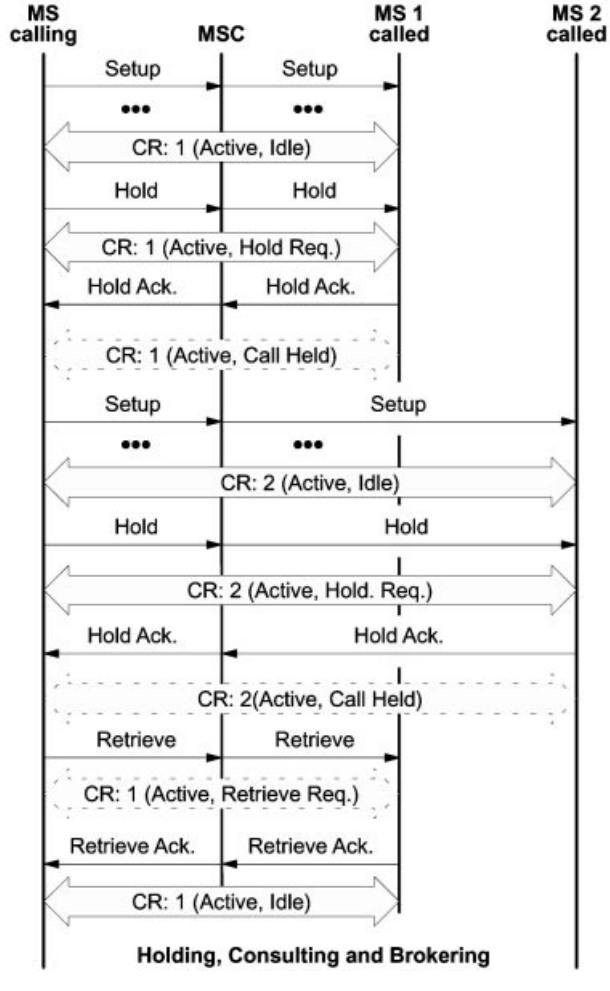

Figure 7.33: Call holding and associated procedures

and reconnected to the call at each side of the traffic channel after the reactivation of the call has been acknowledged with a RETRIEVE ACKNOWLEDGE message.

These call-related signaling procedures modify the call state and define an extended state diagram with an auxiliary state. The participating calls all remain in the active state whereas the auxiliary state changes between hold and idle. In a two-dimensional state space, for example, call CR: 1 changes from (*active*, *idle*) through the state (*active*, *hold request*) into the state (*active*, *call held*) and back through the state (*active*, *retrieve request*).

If outgoing or incoming calls are barred (Figure 7.34), a call request is immediately refused by giving a RELEASE COMPLETE message with a reason in a FACILITY information element (BAOC, BAIC). A state change of the call in the extended state space does not occur. The assumption is, of course, that the calling or called subscriber has activated call barring. The MSC receiving the call request from the calling subscriber must verify the activation of this supplementary service. This requires an inquiry of the HLR of the calling subscriber (for BAOC) or the called subscriber (for BAIC), since the HLRs store the service profiles of the respective subscribers. In this case the HLR acts not only as a database but also as a participant in controlling intelligent network services.

Another call-related supplementary service uses the FACILITY message of the *common* information element procedure: *Call Forwarding Unconditional* (CFU); see Figure 7.34. With this supplementary service, a regular call setup is performed, however, not to the called subscriber but to the forwarding target selected when the service was activated (in Figure 7.34 it is another MS). The calling subscriber is informed about the change of the called number with a FACILITY message. Likewise the target of the forwarding is informed with a FACILITY message that the incoming call is a forwarded call. In this case, there is no necessity for a change in the extended state diagram nor is synchronization between network and mobile station required. It is only necessary that the involved

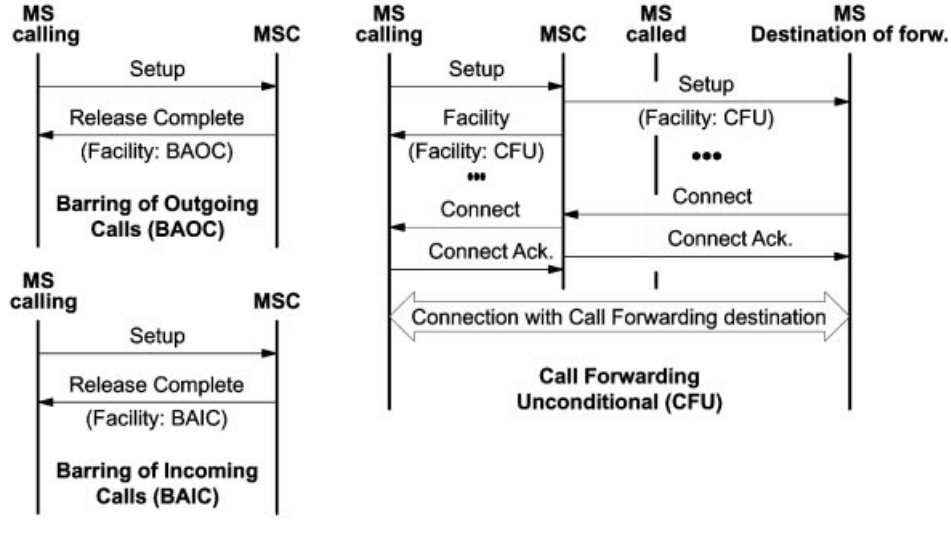

Figure 7.34: Barring and forwarding of calls

mobile stations are informed about the occurrence of forwarding. In this case, the target of the call forwarding is also stored in the HLR of the subscriber who activated the service (the called MS in Figure 7.34). Thus the call processing in the MSC of the called subscriber must be interrupted and the HLR must be informed about the call request. If the called subscriber has activated unconditional call forwarding, the HLR returns the new call target to the MSC, which can continue call processing with the changed target number.

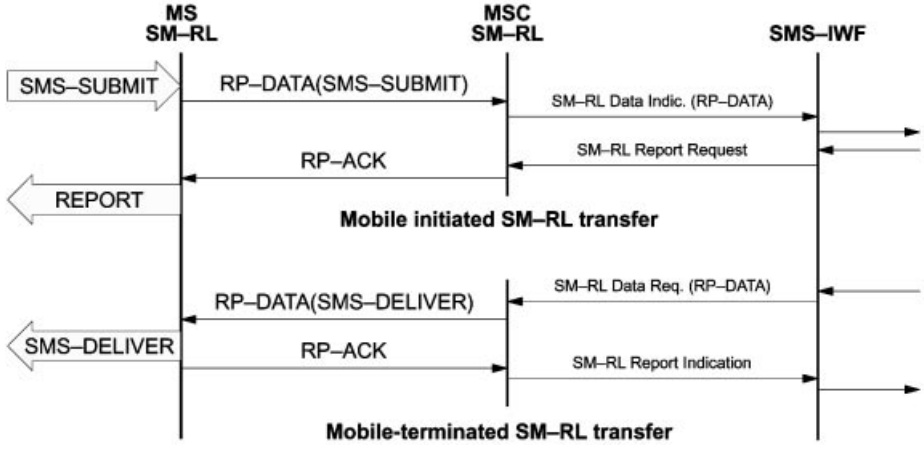

Figure 7.35: Short message transfer between SMR entities

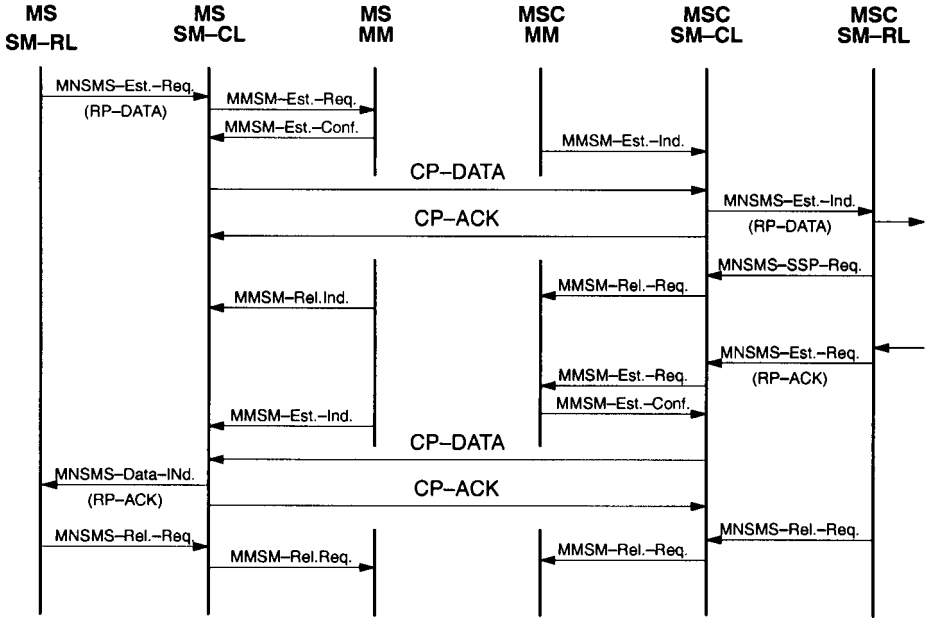

Figure 7.36: Short message transfer on the CM plane between MS and MSC

7.4.8 Realization of Short Message Services

The procedures for the transport of point-to-point short messages reside in the CM sublayer, also called the *Short Message Control Layer* (SM-CL) and in the *Short Message Relay Layer* (SM-RL) directly above. Accordingly, the protocol entities are called the *Short Message Control* entity (SMC) and the *Short Message Relay* entity (SMR). A complete established MM connection is needed for the transport of short messages, which again presumes an existing RR connection with LAPDm protection on an SDCCH or SACCH channel. To distinguish among these packet-switched user data connections, SMS messages are transported across SAPI = 3 of the LAPDm entity.

An SMS transport PDU (SMS-SUBMIT or SMS-DELIVER, Figure 7.35) is transmitted with an RP-DATA message between MSC and MS using the *Short Message Relay Protocol* (SM-RP); see Section 7.3.2. Correct reception is acknowledged with an RP-ACK message either from the SMS service center (mobile-initiated SMS transfer) or from the MS (mobile-terminated SMS transfer).

For the transfer of short messages between SMR entities in MS and MSC, the CM sublayer provides a service to the SM-RL layer above. The SMR entity requests this service for the transfer of RP-DATA or RP-ACK (MNSMS-ESTABLISH-REQUEST, Figure 7.36). Following the SMR service request, the SMC entity itself requests an MM connection on which it then transfers the short message inside a CP-DATA message. The appropriate service primitives between protocol layers are also illustrated in Figure 7.36. The correct reception of CP-DATA is acknowledged with CP-ACK. In these SMC messages, one protocol data unit (PDU) transports a service data unit (SDU) from the SMR sublayer above. This SMC–SDU is the SMS relay message RP-DATA and its acknowledgement RP-ACK which are used to signal the transfer of short messages. The *Short Message Transport Layer* SM-TL above the SM-RL provides end-to-end transport of short messages between mobile station and *SMS Service Center* (SMS-SC).

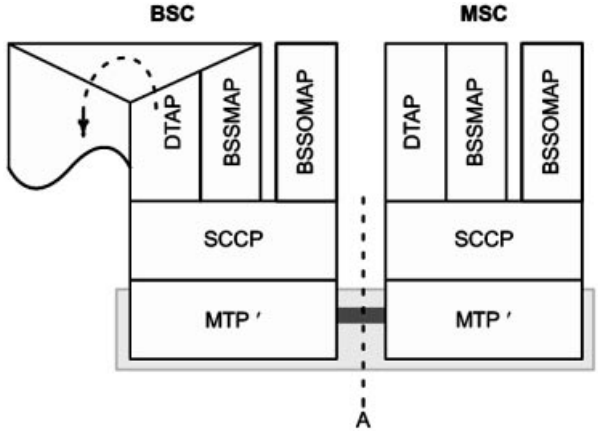

Figure 7.37: Protocols at the A interface between MSC and BSS

7.5 Signaling at the A and Abis Interfaces

Whereas the transport of user data between MSC and BSC occurs across standard connections of the fixed network with 64 kbit/s or 2048 kbit/s (or 1544 kbit/s), the transport of signaling messages between MSC and BSC runs over the SS#7 network. The MTP and SCCP parts of SS#7 are used for this purpose. A protocol function using the services of the SCCP is defined at the A interface. This is the *Base Station Application Part* (BSSAP), which is further subdivided into *Direct Transfer Application Part* (DTAP) and *Base Station System Management Part* (BSSMAP); see Figure 7.37. In addition, the *Base Station System Operation and Maintenance Part* (BSSOMAP) was introduced, which is needed for the transport of network management information from OMC via the MSC to the BSC.

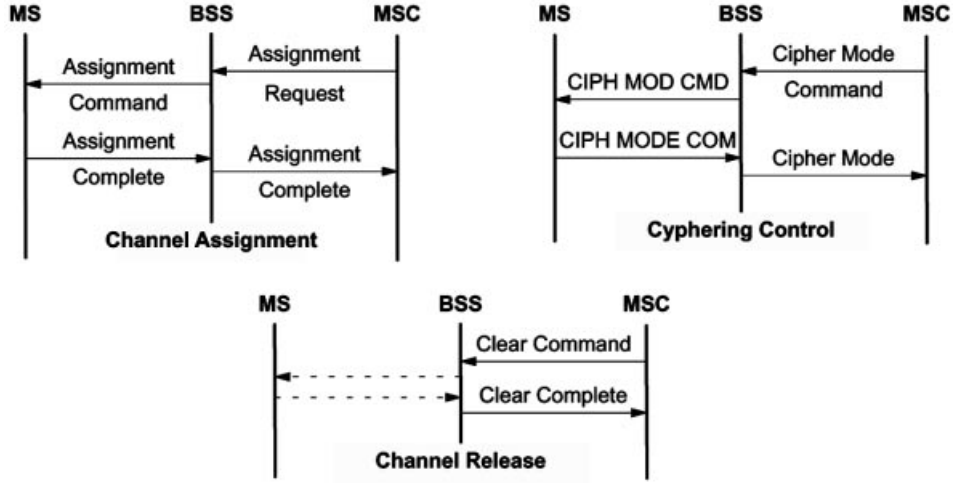

Figure 7.38: Examples of dedicated BSSMAP procedures

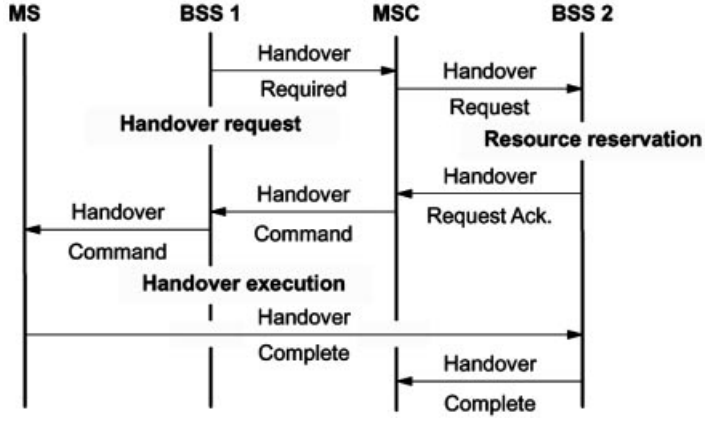

Figure 7.39: Dedicated BSSMAP procedures for internal handover

At the A interface, one can distinguish between two signaling message streams: one between MSC and MS and another between MSC and BSS. The messages to the mobile station (CM, MM) are passed on transparently through the BSS using the DTAP protocol part of SS#7. BSC and BTS do not interpret them. The SCCP protocol part provides a connection-oriented and a connectionless transfer service for signaling messages. For DTAP messages, only connection-oriented service is offered. The DTAP of the BSSAP uses one signaling connection for each active mobile station with one or more transactions per connection. A new connection is established each time when messages of a new transaction with a mobile station are to be transported between MSC and BSS.

Two cases of setting up a new SCCP connection are distinguished. First, in the case of location update and connection setup (outgoing or incoming), the BSS requests an SCCP

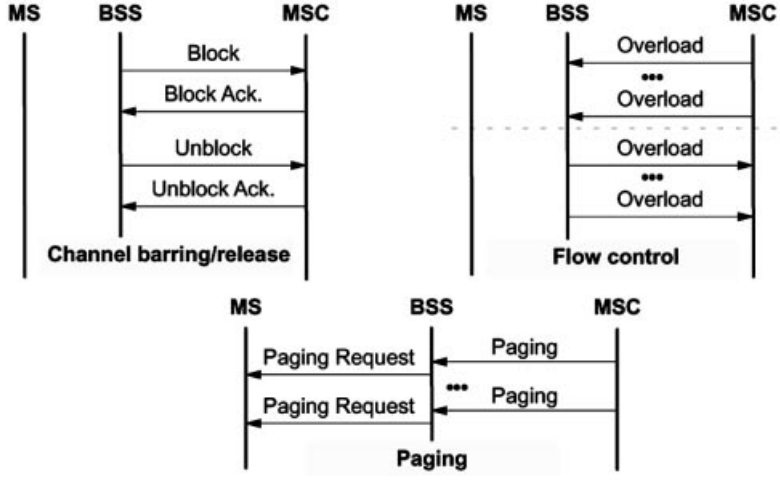

Figure 7.40: Examples of global BSSMAP procedures

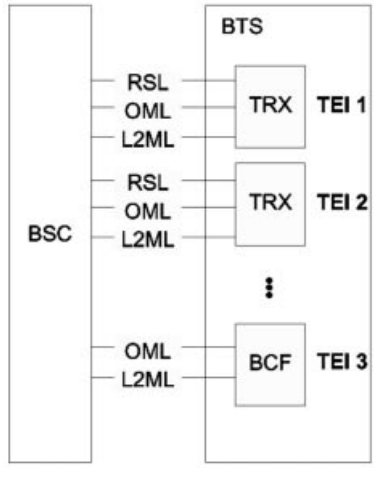

Figure 7.41: Logical connections at Layer 2 of the Abis interface

connection after the channel request (access burst on RACH) from the mobile station has been satisfied with an SDCCH or TCH and after an LAPDm connection has been set up on the SDCCH or FACCH. The second situation for setting up an SCCP connection is a handover to an other BSS, in which case the MSC initiates the connection setup. Most of the signaling messages at the air interface (CM and MM, Tables 7.5 and 7.6) are passed transparently through the BSS and packaged into DTAP–PDUs at the A interface, with the exception of some RR messages.

The BSSMAP implements two more kinds of signaling procedures between MSC and BSS, first those concerning one mobile station or single physical channels at the air interface, and second, global procedures for the control of all the resources of a BSS or cell. In the first case, the BSSMAP also uses connection-oriented SCCP services, whereas, in the second case, global procedures are performed with connectionless SCCP services. Among the BSSMAP procedures for a dedicated resource of the air interface are functions of resource management (channel assignment and release, start of ciphering) and of handover control (Figures 7.38 and 7.39).

Among the global procedures of BSSMAP are paging, flow control to prevent overloading protocol processors or CCCH channels, closing and opening of channels, and parts of handover control (Figure 7.40).

The transmission layer at the Abis interface between BTS and BSC is usually realized as a primary multiplexed line with 2048 kbit/s (1544 kbit/s in North America) or 64 kbit/s. This may include one physical connection per BTS or one for each connection between TRX/BCF module and BTS (Figure 7.5). On these digital paths, traffic or signaling channels of 16 or 64 kbit/s are established. The Layer 2 protocol at the Abis interface is LAPDm, whose *Terminal Equipment Identifier* (TEI) is used to address the TRX and/or BCF of a BTS (Figure 7.41).

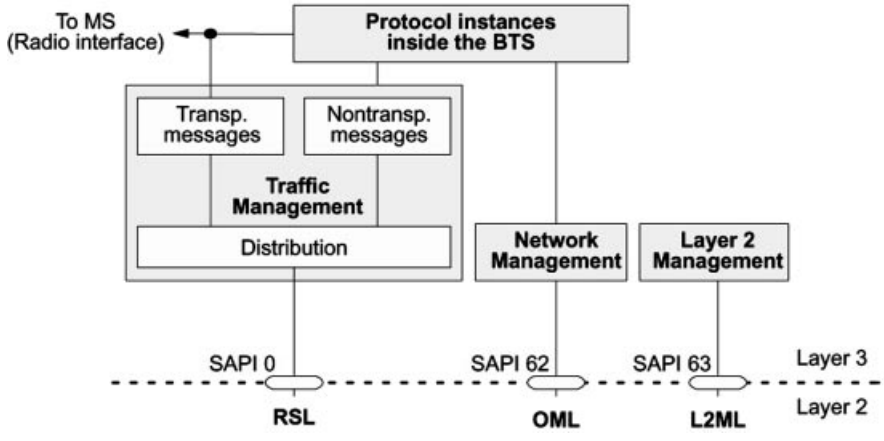

Figure 7.42: Protocol Layer 3 of the BTS at the Abis interface (BTSM)

Several LAPDm connections are established for each TEI: the *Radio Signaling Link* (RSL), SAPI = 0; the *Operation and Maintenance Link* (OML), SAPI = 62; and the *Layer 2 Management Link* (L2ML), SAPI = 63. Traffic management is handled on the

RSL, operation and maintenance on the OML, and management messages of Layer 2 are sent on the L2ML to the TRX or BCF. The RSL is the most important of these three links for the control of radio resources and connections for communication between MS and network. Two types of messages are distinguished on this signaling link: transparent and nontransparent messages (Figure 7.42). Whereas the BTS passes transparent messages on from/to the LAPDm entity of an MS without interpreting or changing them, nontransparent messages are exchanged between BTS and BSC.

In addition, one distinguishes between four groups of *traffic management* messages of the BTS:

- *Radio link layer management*: Procedures to establish, modify and release connections of the link layer (LAPDm) to the mobile station at the air interface Um.
- *Dedicated channel management*: Procedures to start ciphering, transfer of channel measurement reports of an MS, transmitter power control of MS and BTS, handover detection, and modification of a dedicated channel of the BTS for a specific MS which can then receive the channel in an other message (assign, handover command).
- *Common channel management*: Procedures for transferring channel requests from MS

Figure 7.43: Transfer of transparent signaling messages

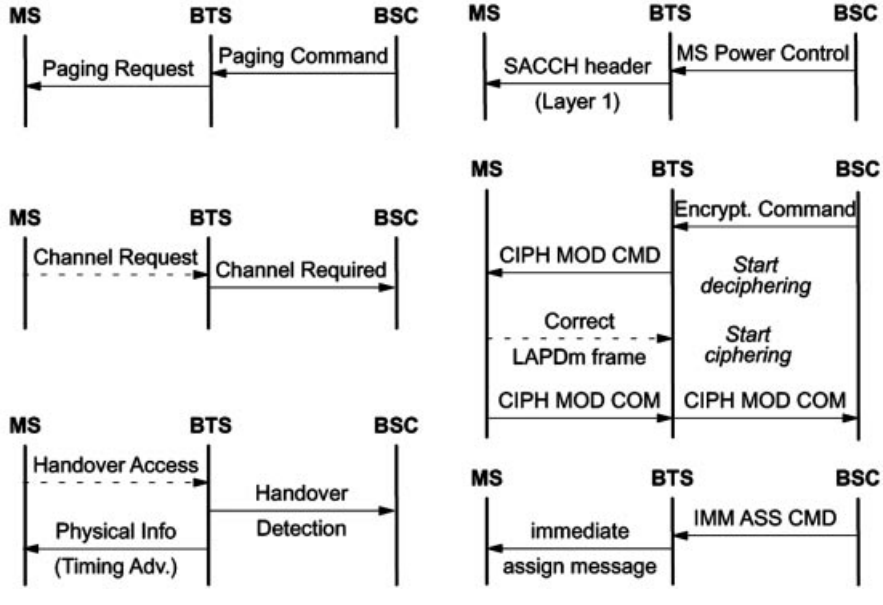

Figure 7.44: Examples of nontransparent signaling between BTS and BSC

(received on RACH), start of paging calls, measurement and transfer of CCCH traffic load measurements, modification of BCCH broadcast information, channel assignments to the MS (on the AGCH), and transmission of Cell Broadcast Short Messages (SMSCB).

- *TRX management*: Procedures for the transfer of measurements of free traffic channels of a TRX to the BSC or for flow control In the case of overloaded TRX processors or overload on the downlink CCCH/ACCH.

In this way, all the RR functions in the BTS can be controlled. The majority of RR messages (Table 7.4) are passed on transparently and do not terminate in the BTS. These messages are transported between BTS and BSC (Figure 7.43) in special messages (DATA REQUEST/INDICATION) packaged into LAPDm frames (Layer 2 at the radio interface).

All protocol messages received by the BTS on the uplink from the MS in LAPDm I/UI frames, except for channel measurement reports of the MS, are passed on as transparent messages in a DATA INDICATION.

Except for the link protocol LAPDm which is completely implemented in the BTS, there are some functions which are also handled by the BTS, and the pertinent messages from or to the MS are transformed by the BTS into the appropriate RR messages. This includes channel assignment, ciphering, assembly of channel measurements from MS and TRX, and their transfer to the BSC (possibly with processing in the BTS), power control commands from the BTS for the MS, and channel requests from the MS (on the RACH) as well as channel assignments (Figure 7.44). Thus four of the RR messages on the downlink direction to the MS (Table 7.4) cannot be treated as transparent messages: CIPHERING MODE COMMAND, PAGING REQUEST, SYSTEM INFORMATION, and the three IMMEDIATE ASSIGN messages. All the other RR messages to the MS are sent transparently within a DATA REQUEST to the BTS.

Figure 7.45 shows the format of a BTSM message (Layer 3 between BSC and BTS). Transparent and nontransparent messages are distinguished with a *message discriminator*

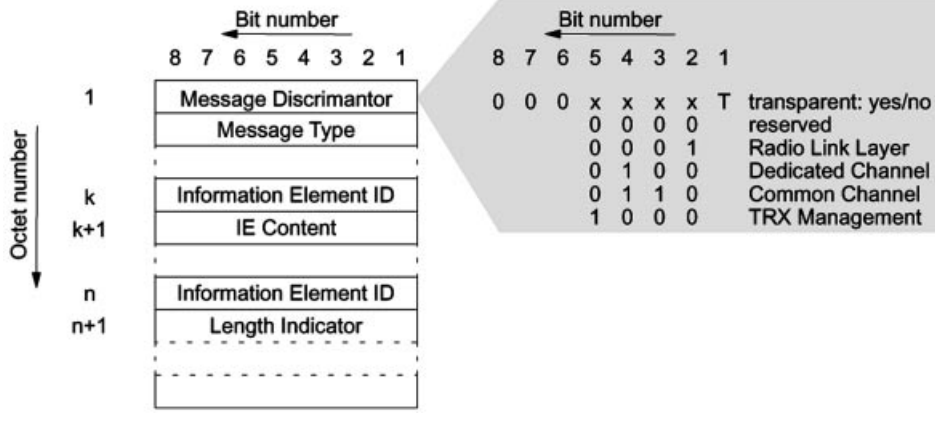

Figure 7.45: Format of BTSM-RSL protocol messages

Table 7.8: Input format of some MMI commands

Function	MMI procedure
Activate	*nn(n)*Si#
Deactivate	#nn(n)*Si#
Status enquiry	*#nn(n)*Si#
Registration	**nn(n)*Si#
Delete	##nn(n)*Si#

in the first octet. For this purpose, the T-bit (bit 1 of octet 1) is set to logical 1 for messages which the BTS is supposed to handle transparently or has recognized as transparent. Bits 2 to 5 serve to assign the messages to one of the four groups defined on the *Radio Signaling Link* (RSL). Including the *Message Type* (MT) defines the message completely (Figure 7.45). The remainder of the BTSM message contains mandatory and optional *Information Elements* (IEs) which have a fixed length of mostly two octets or which contain an additional length indicator in the case of variable length.

7.6 Signaling at the User Interface

Another often neglected but nevertheless very important interface in a mobile system is the user interface of the mobile station equipment. This *Man–Machine Interface* (MMI) can be realized freely and therefore in many different ways by the mobile equipment manufacturers. In order to keep a set of standardized service control functions in spite of this variety, the MMI commands have been introduced. These MMI commands define procedures mainly for the control of basic and supplementary services. The control procedures are constructed around the input of command token strings which are delineated and formatted with the tokens * and #. In order to avoid a user having to learn and memorize a certain number of service control procedures before being able to use the mobile phone, a small set of basic required commands for the MMI interface has been defined; this is the *basic public MMI* which must be satisfied by all mobile stations.

Table 7.9: Some basic MMI commands

Function	MMI procedure
Mobile phone IMEI enquiry	*#06#
Change password for call barring	**03*330*old_PWD*new_PWD*new_PWD#
Change PIN in SIM	**04*old_PIN*new_PIN*new_PIN#
Select SIM number storage	n(n)(n)#

The specification *basic public MMI* outlines the basic functions which must be implemented as a minimum at the MMI of a mobile station. This includes the arrangement of a 12-key keyboard with the numbers 0 through 9 and the keys for * and # as well as SEND and END keys, which also serve to initiate a desired call or accept or terminate a call.

Table 7.10: MMI service codes for supplementary services[a]

Abbreviation	Service	MMI code	Sia	Sib
	All call forwarding, only for (de)activation	002	–	–
	All conditional call forwarding (not CFU), only for (de)activation	004	–	–
CFU	Call forwarding unconditional	21	DN	BS
CFB	Call forwarding on mobile subscriber busy	67	DN	BS
CFNRy	Call forwarding on no reply	61	DN	BS
CFNRc	Call forwarding on mobile subscriber not reachable	62	DN	BS
	All call barring (only for deactivation)	330	PW	BS
BAOC	Barring of all outgoing calls	33	PW	BS
BOI	Barring of outgoing international calls	331	PW	BS
BOIC-exHC	Barring of outgoing international calls except those to home PLMN	332	PW	BS
BAIC	Barring of all incoming calls	35	PW	BS
BIC-Roam	Barring of incoming calls when roaming outside the home PLMN	351	PW	BS
CLIP	Calling line identification presentation	30	–	BS
CLIR	Calling line identification restriction	31	–	BS
CW	Call waiting	43	–	BS
COLP	Connected line identification presentation	76	–	BS
COLR	Connected line identification restriction	77	–	BS

[a] BS, basic service (see Table 7.11); DN, destination number; PW, password.

Some basic operational sequences for making or taking a call are also defined. These requirements are so general that they can be easily satisfied by all mobile equipment.

The MMI commands for the control of supplementary services and the enquiry and configuration of parameters are much more extensive. Using a set of MMI commands which are uniform for all mobile stations allows control functions to be performed which are often hidden in equipment-specific user guidance menus. For certain functional areas, a mobile station can thus be operated in a manufacturer-independent way, if one forgoes the sometimes very comfortable possibilities of user-guiding menus and instead learns the control sequences for the respective functions. These sequences are mapped onto the respective signaling procedures within the mobile station.

An MMI command is always constructed according to the same pattern. Five basic formats are distinguished (Table 7.8), which all start with a combination of the tokens * and #: activation (*), deactivation (#), status inquiry (*#), registration (**), and cancellation (##).

In addition, the MMI command must contain an MMI service code of two or three tokens, which selects the function to be performed. In certain cases, the MMI procedure requires additional arguments or parameters, which are separated by *, as *Supplementary*

Table 7.11: MMI codes for basic services

Category	Service	MMI code BS
Telematic service	All telematic services	10
	Telephone	11
	All data services	12
	Facsimile	13
	Videotex	14
	Teletext	15
	SMS	16
	All data services except SMS	18
	All telematic services except SMS	19
Bearer service	All bearer services	20
	All asynchronous services	21
	All synchronous services	22
	All connection-oriented synchronous data services	24
	All connection-oriented asynchronous data services	25
	All packet-oriented synchronous data services	26
	All PAD-access services	27

Information (Si). The MMI command is always terminated with # and may also require depressing the SEND key, if the command is not executed locally within the mobile station but must be transmitted to the network. Table 7.9 contains some basic examples of MMI commands, e.g. the inquiry for the IMEI of the MS (*# 06#) or the change of the PIN (**04*old_PIN*new_PIN*new_PIN#) used to protect the SIM card against misuse. This example also shows how supplementary information is embedded in the command.

With MMI commands, it is also possible to configure and use *supplementary services* (see Section 4.3). For this purpose, each supplementary service is designated with a two- or three-digit MMI service code to select the respective supplementary service (Table 7.10). In some cases, supplementary information is mandatory for the activation of the service, e.g. one needs the target *Destination Number* (DN) for the call forwarding functions, or the *Activation Password* (PW) for the supplementary service of barring incoming or outgoing calls (*Sia* in Table 7.10).

The example of unconditional call forwarding also illustrates the difference between registration and activation of a service. With the command **21*call_number# the forwarding function is registered, the target number configured, and the unconditional forwarding activated. Later, the unconditional forwarding can be deactivated any time with #21# and reactivated with *21#. The target number call_number remains stored, unless it is cancelled with the command ##21#. After cancellation, if call forwarding is desired again, it must first be registered again using the **21… command. For some basic services, characteristics can also be activated selectively. The MMI command can contain a second parameter field with supplementary information (*Sib*, Table 7.10) which is again delineated with *. This field contains the service code BS of the basic service for which the supplementary service is to become effective.

An overview of MMI codes for basic services is given in Table 7.11. For example, one can bar incoming calls except short messages with the command **35*PW*18#, or one can forward incoming fax calls to the number fax_number with the command **21*fax_number*13# (the other teleservices remain unaffected).

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

8 Roaming and Switching

8.1 Mobile Application Part Interfaces

The main benefit for the mobile subscribers that the international standardization of GSM has brought is that they can move freely not only within their home networks but also in international GSM networks and that at the same time they can even get access to the special services they subscribed to at home – provided there are agreements between the operators. The functions needed for this free roaming are called roaming or mobility functions. They rely mostly on the GSM-specific extension of the *Signalling System Number 7* (SS#7). The *Mobile Application Part* (MAP) procedures relevant for roaming are first the *Location Registration/Update*, *IMSI Attach/Detach*, requesting subscriber data for call setup, and paging. In addition, the MAP contains functions and procedures for the control of supplementary services and handover, for subscriber management, for IMEI management, for authentication and identification management, as well as for the user data transport of the *Short Message Service*. MAP entities for roaming services reside in the MSC, HLR, and VLR. The corresponding MAP interfaces are defined as B (MSC-VLR), C (MSC-HLR), D (HLR-VLR), E (MSC-MSC), and G (VLR-VLR) (Figure 3.9). At the subscriber interface, the MAP functions correspond to the functions of *Mobility Management* (MM), i.e. the MM messages and procedures of the Um interface are translated into the MAP protocols in the MSC.

The most important functions of GSM *Mobility Management* are *Location Registration* with the PLMN and *Location Updating* to report the current location of an MS, as well as the identification and authentication of subscribers. These actions are closely interrelated. During registration into a GSM network, during the location updating procedure, and also during the setup of a connection, the identity of a mobile subscriber must be determined and verified (authentication).

The mobility management data are the foundation for creating the functions needed for routing and switching of user connections and for the associated services. For example, they are requested for routing an incoming call to the current MSC or for localizing an MS before paging is started. In addition to mobility data management, information about the configuration of supplementary services is requested or changed, e.g. the currently valid target number for unconditional call forwarding in the HLR or VLR registers.

8.2 Location Registration and Location Update

Before a mobile station can be called or gets access to services, the subscriber has to register with the mobile network (PLMN). This is usually the home network where the subscriber has a service contract. However, the subscriber can equally register with a foreign network provider in whose service area he or she is currently visiting, provided there is a roaming agreement between the two network operators. Registration is only required if there is a change of networks, and therefore a VLR of the current network has not yet issued a TMSI to the subscriber. This means the subscriber has to report to the current network with his IMSI and receives a new TMSI by executing a *Location Registration* procedure. This TMSI is stored by the MS in its nonvolatile SIM storage, such that even after a powerdown and subsequent power-up only a normal *Location Updating* procedure is required.

The sequence of operations for registration is presented schematically in Figure 8.1. After a subscriber has requested registration at his or her current location by sending a LOCATION UPDATE REQUEST with his or her IMSI and the current location area (LAI), first the MSC instructs the VLR with a MAP message UPDATE LOCATION AREA to register the MS with its current LAI. In order for this registration to be valid, the identity of the subscriber has to be checked first, i.e. the authentication procedure is executed. For this purpose, the authentication parameters have to be requested from the AUC through the HLR. The precalculated sets of security parameters (Kc, RAND, SRES) are usually not transmitted individually to the respective VLR. In most cases, several complete sets are kept at hand for several authentications. Each set of parameters, however, can only be used once, i.e. the VLR must continually update its supply of security parameters (AUTHENTICATION PARAMETER REQUEST).

After successful authentication (see Section 6.3.2), the subscriber is assigned a new MSRN, which is stored with the LAI in the HLR, and a new TMSI is also reserved for this subscriber; this is *TMSI Reallocation* (see Figure 7.25). To encrypt the user data, the base station needs the ciphering key Kc, which it receives from the VLR by way of the MSC with the command START CIPHERING. After ciphering of the user data has begun, the TMSI is sent in encrypted form to the mobile station. Simultaneously with the TMSI assignment, the correct and successful registration into the PLMN is acknowledged (LOCAPDATE ACCEPT). Finally, the mobile station acknowledges the correct reception of the TMSI (TMSI REALLOCATION COMPLETE, see Figure 7.26).

While the location information is being updated, the VLR is obtaining additional information about the subscriber, e.g. the MS category or configuration parameters for supplementary services. For this purpose, the *Insert Subscriber Data Procedure* is defined (INSERT SUBSCRIBER DATA message in Figure 8.1). It is used for registration or location updating in the HLR to transmit the current data of the subscriber profile to the VLR. In general, this MAP procedure can always be used when the profile parameters are changed, e.g. if the subscriber reconfigures a supplementary service such as unconditional forwarding. The changes are communicated immediately to the VLR with the *Insert Subscriber Data Procedure*.

The location update procedure is executed, if the mobile station recognizes by reading the LAI broadcast on the BCCH that it is in a new location area, which leads to updating the

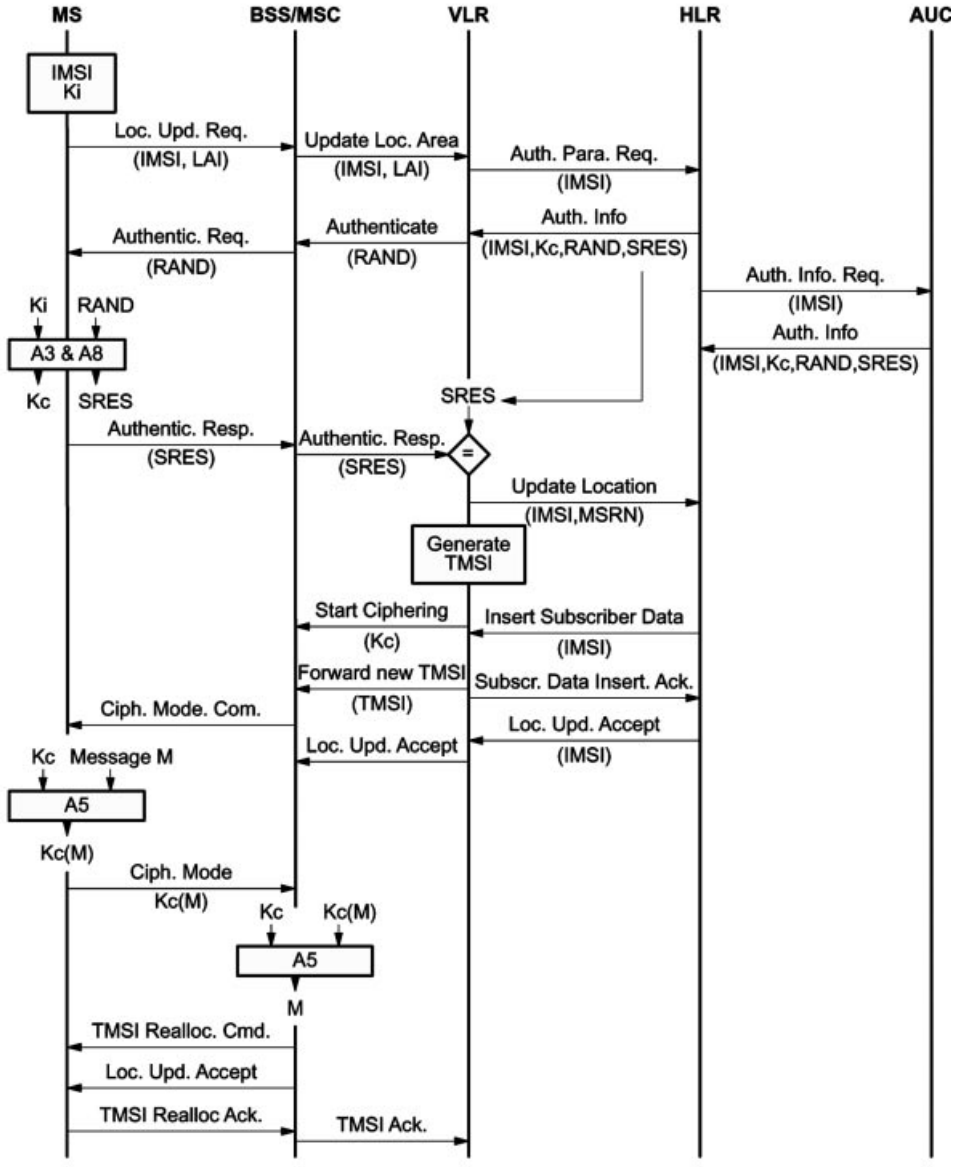

Figure 8.1: Overview of the location registration procedure

location information in the HLR record. Alternatively, the location update can also occur periodically, independent of the current location. For this purpose, a time interval value is broadcast on the BCCH, which prescribes the time between location updates. The main objective of this location update is to know the current location for incoming calls or short messages, so that the call or message can be directed to the current location of the mobile station. The difference between the location update procedure and the location registration procedure is that in the first case the mobile station has already been assigned a TMSI. The

TMSI is unique only in connection with an LAI, and both are kept together in the non-volatile storage of the SIM card. With a valid TMSI, the MS also keeps a current ciphering key Kc for encryption of user data (Figure 8.2), although this key is renewed during the location update procedure. This key is recalculated by the MS based on the random number RAND used for authentication, whereas on the network side it is calculated in the AUC and made available in the VLR.

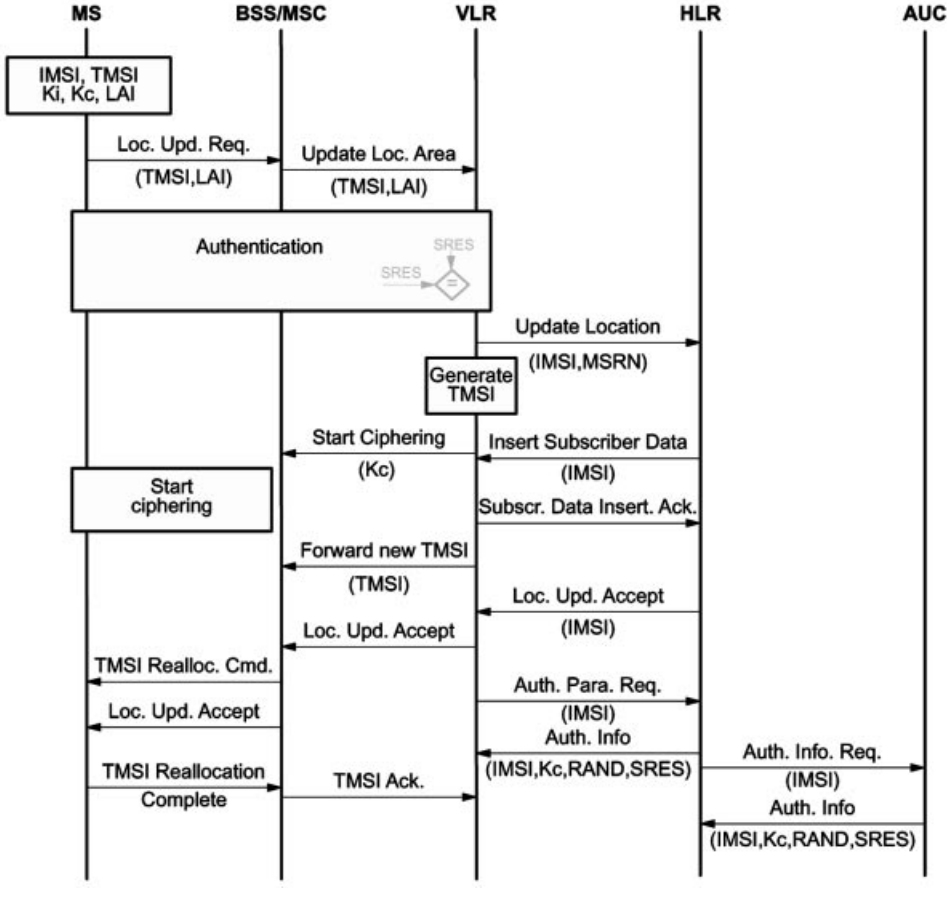

Figure 8.2: Overview of the location updating procedure

Corresponding to the location update procedure, there is an MM procedure at the air interface of the MM-category *specific*. Besides the location updating proper, it contains three blocks which are realized at the air interface by three procedures of the category *common* (see Figure 7.26): the identification of the subscriber, the authentication, and the start of ciphering on the radio channel. In the course of location updating, the mobile station also receives a new TMSI, and the current location is updated in the HLR. Figure 8.2 illustrates the standard case of a location update. The MS has entered a new LA, or the timer for periodic location updating has expired, and the MS requests to update its location information. It is assumed that the new LA still belongs to the same VLR as the previous

one, so only a new TMSI needs to be assigned. This is the most frequent case. But if its not quite so crucial to keep the subscriber identity confidential, it is possible to avoid assigning a new TMSI. In this case, only the location information is updated in the HLR/VLR.

The new TMSI is transmitted to the MS in enciphered form together with the acknowledgement of the successful location update. The location update is complete after acknowledgement by the mobile station. After execution of the authentication, the VLR can complete its database and replace the "consumed" 3-tuple (RAND, SRES, Kc) by another one requested from the HLR/AUC.

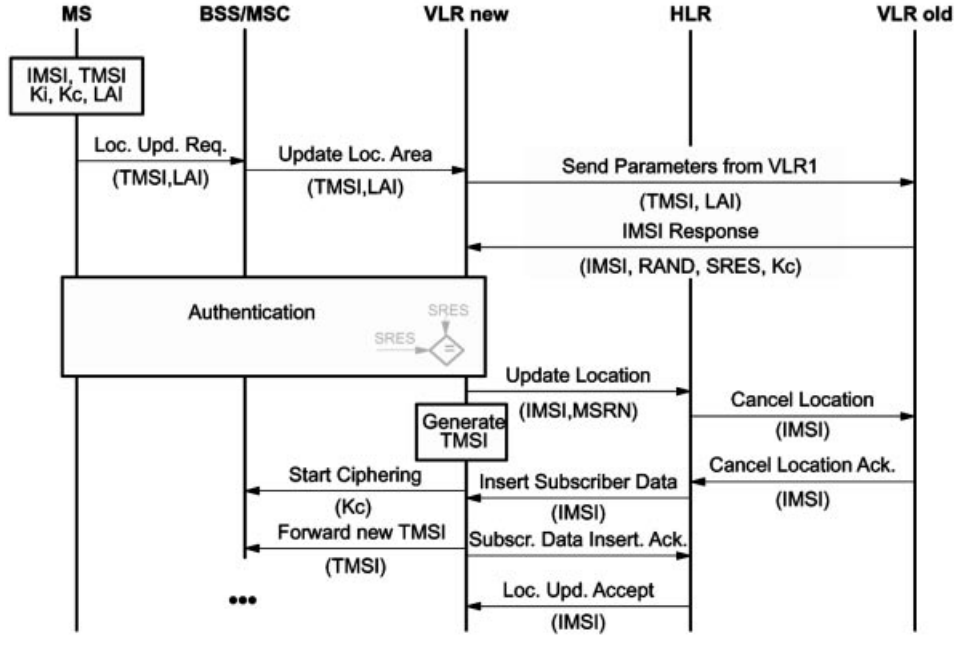

Figure 8.3: Location update after changing the VLR area

If location change involves both LA and VLR, the location update procedure is somewhat more complicated (Figure 8.3). In this case, the new VLR has to request the identification and security data for the MS from the old VLR and store them locally. Only in emergency cases, if the old VLR cannot be determined from the old LAI or if the old TMSI is not known in the VLR, the new VLR may request the IMSI directly from the MS (identification procedure). Only after a mobile station has been identified through the IMSI from the old VLR and after the security parameters are available in the new VLR, is it possible for the mobile station to be authenticated and registered in the new VLR, for a new TMSI to be assigned, and for the location information in the HLR to be actualized. After successful registration in the new VLR (LOCATION UPDATE ACCEPT) the HLR instructs the old VLR to cancel the invalid location data in the old VLR (CANCEL LOCATION).

In the examples shown (Figures 8.1–8.3), the location information is stored as MSRN in the HLR. The MSRN contains the routing information for incoming calls and this infor-

mation is used to route incoming calls to the current MSC. In this case, the HLR receives the routing information already at the time of the location update. Alternatively, at location update time, the HLR may just store the current MSC and/or VLR number in connection with an LMSI, such that routing information is only determined at the time of an incoming call.

8.3 Connection Establishment and Termination

8.3.1 Routing Calls to Mobile Stations

The number dialed to reach a mobile subscriber (MSISDN) contains no information at all about the current location of the subscriber. In order to establish a complete connection to a mobile subscriber, however, one must determine the current location and the locally responsible switch (MSC). In order to be able to route the call to this switch, the routing address to this subscriber (MSRN) has to be obtained. This routing address is assigned temporarily to a subscriber by its currently associated VLR. At the arrival of a call at the GMSC, the HLR is the only entity in the GSM network which can supply this information, and therefore it must be interrogated for each connection setup to a mobile subscriber. The principal sequence of operations for routing to a mobile subscriber is shown in Figure 8.4. An ISDN switch recognizes from the MSISDN that the called subscriber is a mobile subscriber, and therefore can forward the call to the GMSC of the subscriber's home PLMN based on the CC and NDC in the MSISDN (1). This GMSC can now request the current routing address (MSRN) for the mobile subscriber from the HLR using the MAP (2,3). By way of the MSRN the call is forwarded to the local MSC (4), which determines the TMSI of the subscriber (5,6) and initiates the paging procedure in the relevant location area (7). After the mobile station has responded to the paging call (8), the connection can be switched through.

Several variants for determining the route and interrogating the HLR exist, depending on how the MSRN was assigned and stored, whether the call is national or international, and depending on the capabilities of the associated switching centers.

8.3.1.1 Effect of the MSRN Assignment on Routing

There are two ways to obtain the MSRN:

- obtaining the MSRN at location update
- obtaining the MSRN on a per call basis

For the first variant, an MSRN for the mobile station is assigned at the time of each location update which is stored in the HLR. This way the HLR is in a position to supply immediately the routing information needed to switch a call through to the local MSC.

The second variant requires that the HLR has at least an identification for the currently responsible VLR. In this case, when routing information is requested from the HLR, the HLR first has to obtain the MSRN from the VLR. This MSRN is assigned on a per call basis, i.e. each call involves a new MSRN assignment.

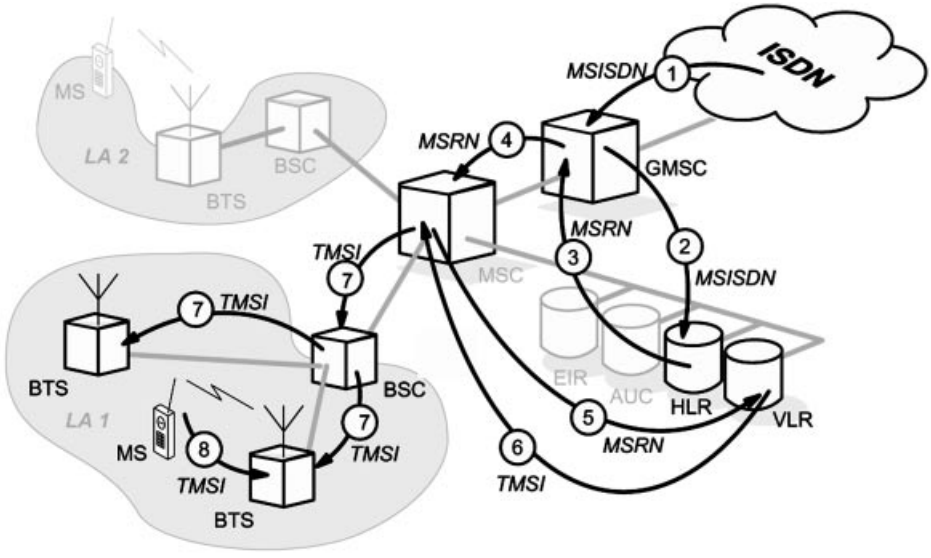

Figure 8.4: Principle of routing calls to mobile subscribers

8.3.1.2 Placement of the Protocol Entities for HLR Interrogation

Depending on the capabilities of the associated switches and the called target (national or international MSISDN), there are different routing procedures. In general, the local switching center analyzes the MSISDN. Due to the NDC, this analysis of the MSISDN allows the separation of the mobile traffic from other traffic. The case that mobile call numbers are integrated into the numbering plan of the fixed network is currently not provided.

In the case of a national number, the local exchange recognizes from the NDC that the number is a mobile ISDN number. The fixed network and home PLMN of the called subscriber reside in the same country. In the ideal case, the local switch can interrogate the HLR responsible for this MSISDN (HLR in the home PLMN of the subscriber) and obtain the routing information (Figure 8.5a). The connection can then be switched through via fixed connections of the ISDN directly to the MSC.

If the local exchange does not have the required protocol intelligence for the interrogation of the HLR, the connection can be passed on preliminarily to a transit exchange, which then assumes the HLR interrogation and routing determination to the current MSC (Figure 8.5b). If the fixed network is not at all capable of performing an HLR interrogation, the connection has to be directed through a GMSC. This GMSC connects through to the current MSC (Figure 8.5c). For all three cases, the mobile station could also reside in a foreign PLMN (roaming); the connection is then made through international lines to the current MSC after interrogating the HLR of the home PLMN.

In the case of an international call number, the local exchange recognizes only the international CC and directs the call to an *International Switching Center* (ISC). Then the ISC can recognize the NDC of the mobile network and process the call accordingly. Figures 8.6 and 8.7 show examples for the processing of routing information. An inter-

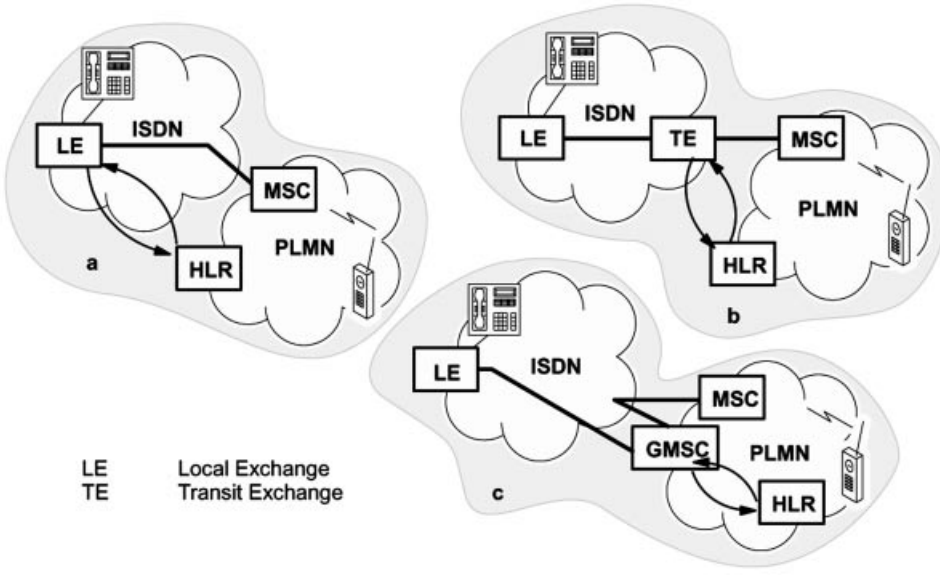

Figure 8.5: Routing variants for national MSISDN

national call to a mobile subscriber involves at least three networks: the country from which the call originates; the country with the home PLMN of the subscriber, *Home PLMN* (H-PLMN); and the country in which the mobile subscriber is currently roaming, *Visited PLMN* (V-PLMN). The traffic between countries is routed through ISCs. Depending on the capabilities of the ISC, there are several routing variants for international calls to mobile subscribers. The difference is determined by the entity that performs the HLR interrogation, resulting in differently occupied line capacities.

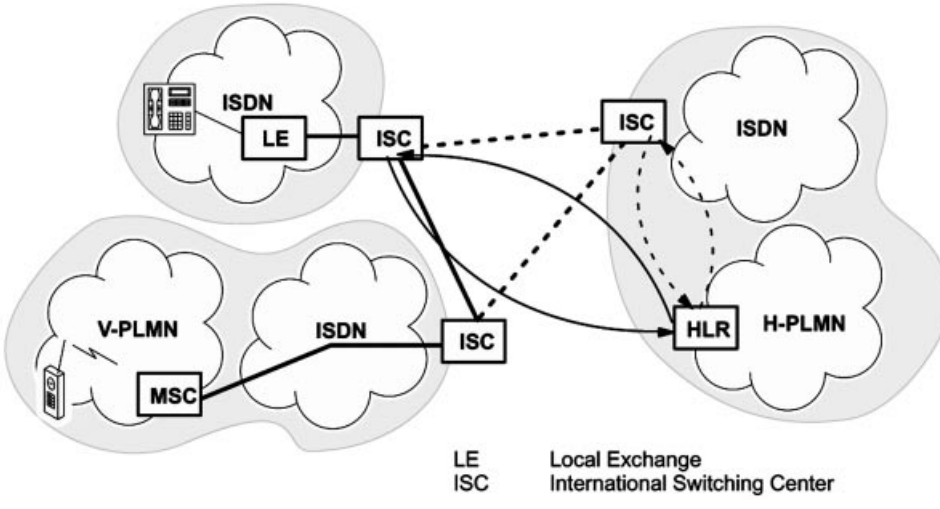

Figure 8.6: Routing for international MSISDN (HLR interrogation from ISC)

If the ISC performs the HLR interrogation, the routing to the current MSC is performed either by the ISC of the originating call or by the ISC of the mobile subscriber's H-PLMN (Figure 8.6). If no ISC can process the routing, again a GMSC has to get involved, either a GMSC in the country where the call originates or the GMSC of the H-PLMN (Figure 8.7).

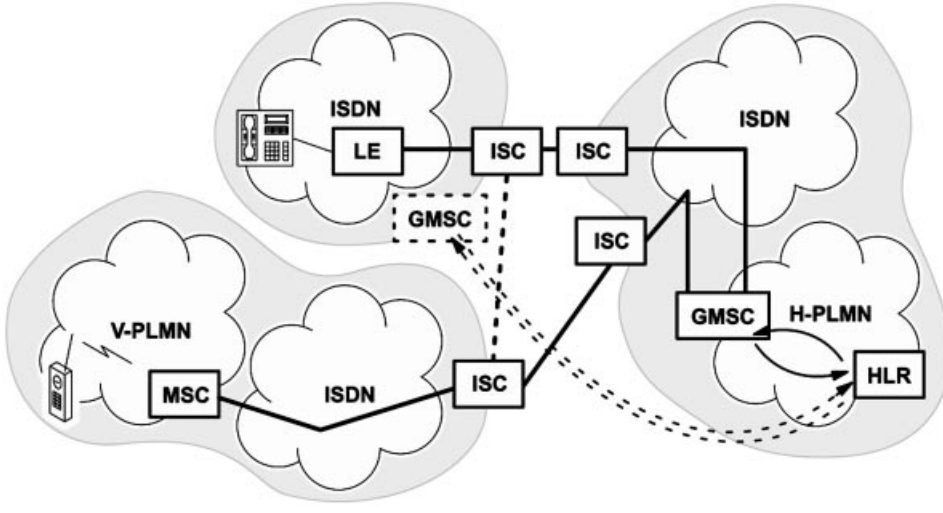

Figure 8.7: Routing through GMSC for international MSISDN

For the routing procedures explained here, it does not matter which kind of subscriber is calling, i.e. the subscriber may be in the fixed network or in the mobile network. However, for calls from mobile subscribers, the HLR interrogation is usually performed at the local exchange (MSC).

8.3.2 Call Establishment and Corresponding MAP Procedures

Call establishment in GSM at the air interface is similar to ISDN call establishment at the user network interface (Q.931) [7]. The procedure is supplemented by several functions: random access to establish a signaling channel (SDCCH) for call setup signaling, the authentication part, the start of ciphering, and the assignment of a radio channel.

The establishment of a connection always contains a verification of user identity (authentication) independent of whether it is a *mobile-originated* call setup or *a mobile-terminated* call setup. The authentication is performed in the same way as for location updating. The VLR supplements its database entry for this mobile station with a set of security data, which replaces the "consumed" 3-tuple (RAND, SRES, Kc). After successful authentication, the ciphering process for the encryption of user data is started.

8.3.2.1 Outgoing Connection Setup

For outgoing connection setup (Figure 8.8), first the mobile station announces its connec-

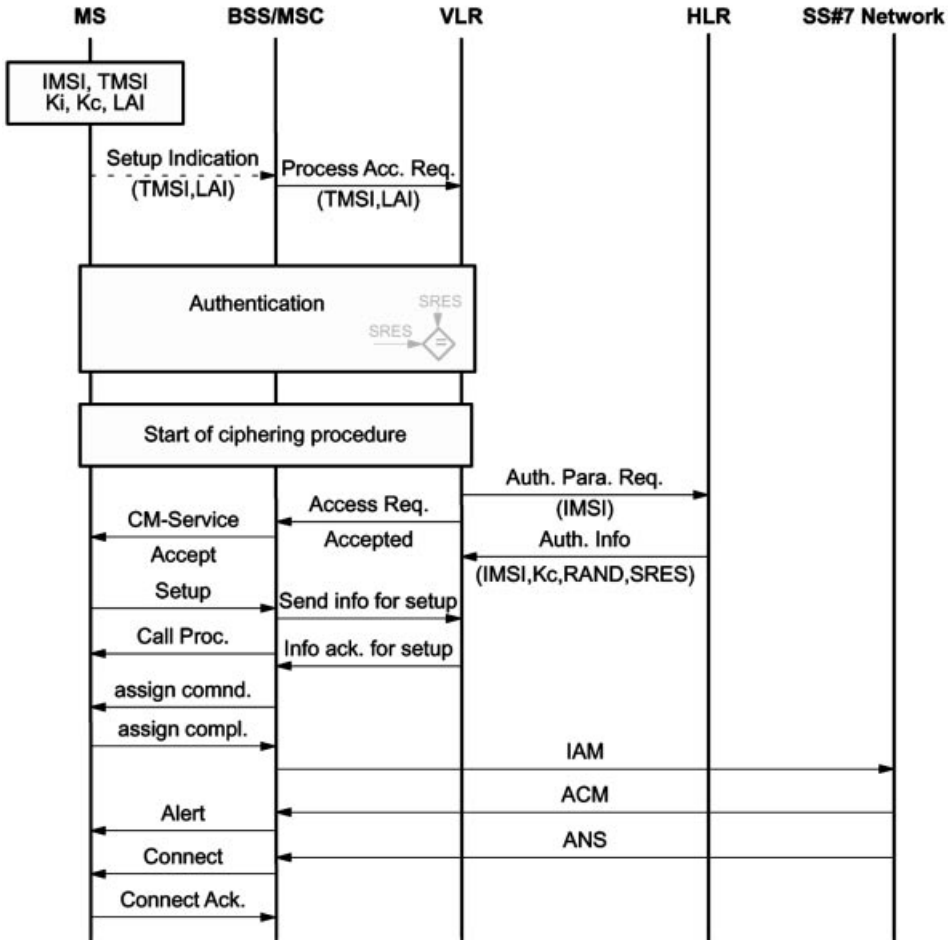

Figure 8.8: Overview of outgoing call setup

tion request to the MSC with a SETUP INDICATION message, which is a pseudo-message. It is generated between the MM entity of the MSC and the MAP entity, when the MSC receives the message CM-SERVICE REQUEST from the MS, which indicates in this way the request for an MM connection (see Figure 7.27). Then the MSC signals to the VLR that the mobile station identified by the temporary TMSI in the location area LAI has requested service access (PROCESS ACCESS REQUEST) which is an implicit request for a random number RAND from the VLR, to be able to start the authentication of the MS. This random number is transmitted to the mobile station, its response with authentication result SRES is returned to the VLR, which now examines the authenticity of the mobile station's identity (compare authentication at registration, Figure 8.1).

After successful authentication, the ciphering process is started on the air interface, and this way the MM connection between MS and MSC has been completely established (CM-SERVICE ACCEPT). Subsequently, all signaling messages can be sent in encrypted form. Only now the MS reports the desired calling target. While the MS is informed with a CALL

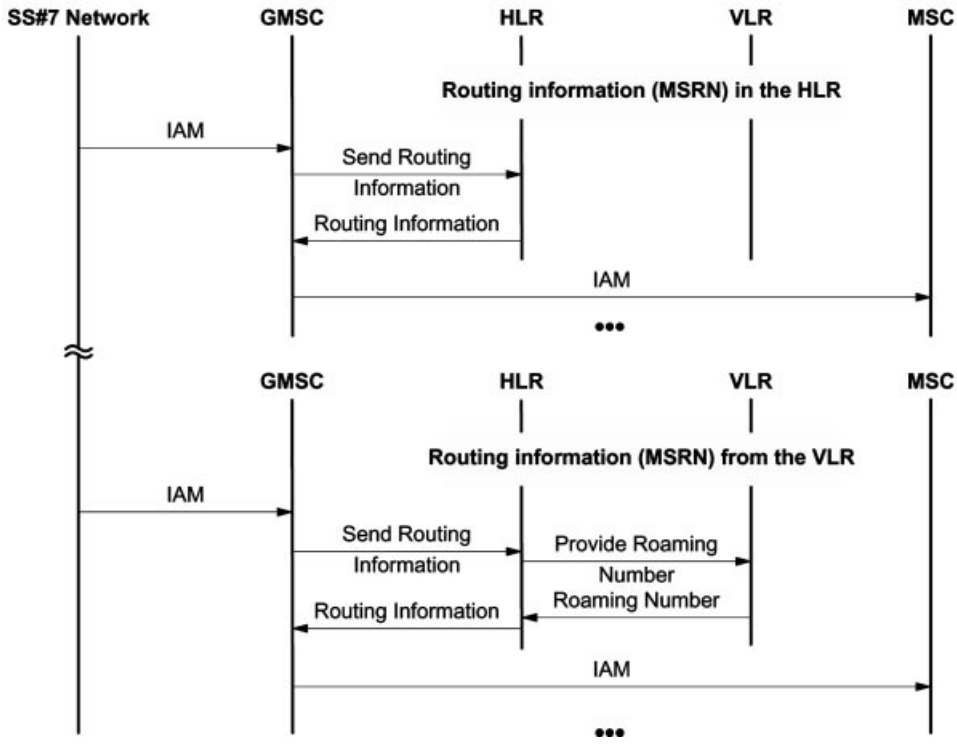

Figure 8.9: Interrogation of routing information for incoming call

PROCEEDING message that processing of its connection request has been started, the MSC reserves a channel for the conversation and assigns it to the MS (ASSIGN). The connection request is signaled to the remote network exchange through the signaling system SS#7 with the *ISDN User Part* (ISUP) message IAM [7]. When the remote exchange answers (ACM), the delivery of the call can be indicated to the mobile station (ALERT). Finally, when the called partner goes off-hook, the connection can be switched through (CONNECT, ANS, CONNECT ACKNOWLEDGE).

8.3.2.2 Incoming Connection Setup

For incoming connection setup, it is necessary to determine the exact location of an MS in order to route the call to the currently responsible MSC. A call to a mobile station is therefore always routed to an entity which is able to interrogate the HLR for temporary routing information and to use it to forward the call. Usually, this entity is a GMSC of the home PLMN of the MS (see Section 8.3.1.2). Through this HLR interrogation, the GMSC obtains the current MSRN of the mobile station and forwards it to the current MSC (Figure 8.9).

Depending on whether the MSRN is stored in the HLR or first has to be obtained from the VLR, two variants of the HLR interrogation exist. In the first case, the interrogated HLR can supply the MSRN immediately (ROUTING INFORMATION). In the second case, the HLR

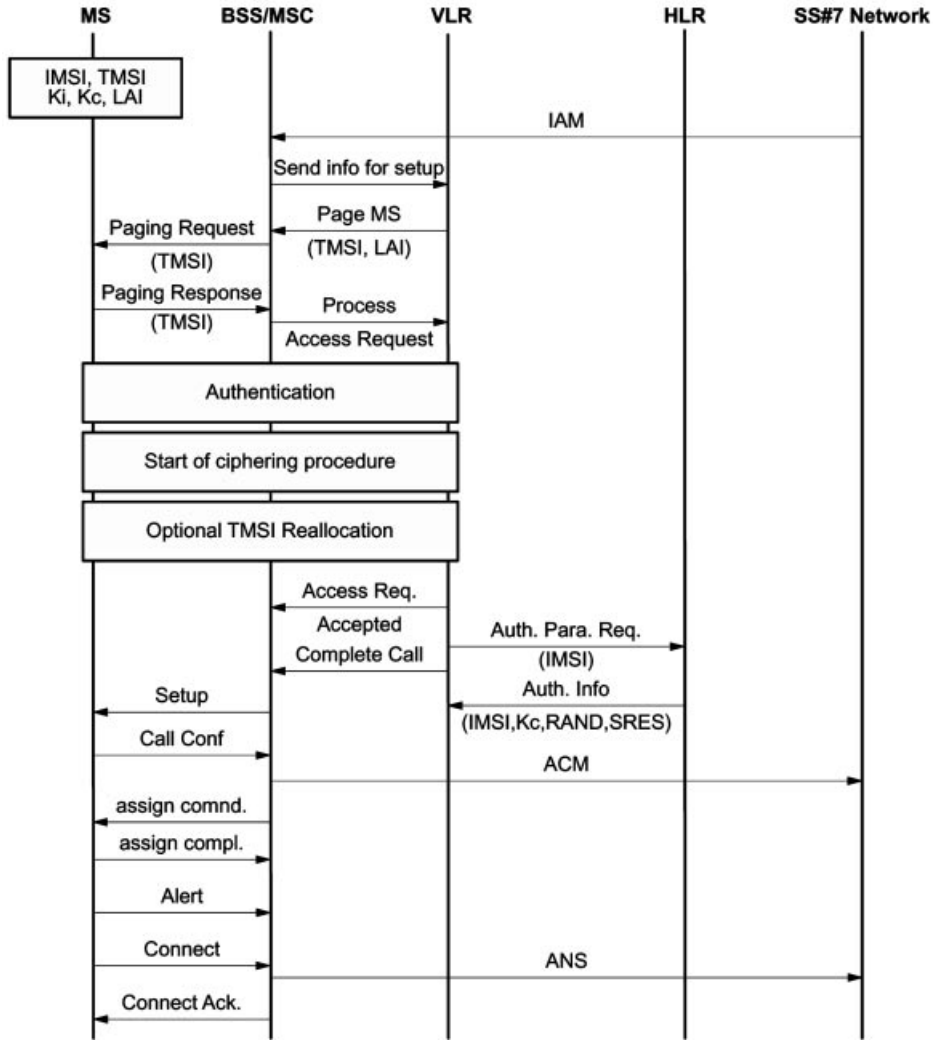

Figure 8.10: Overview of incoming call setup

has only received and stored the current VLR address during location update. Therefore, the HLR first has to request the current routing information from the VLR before the call can be switched through to the local MSC.

Call processing is interrupted again in the local MSC in order to determine the exact location of the mobile station within the MSC area (SEND INFO FOR SETUP, Figure 8.10). The current LAI is stored in the location registers, but an LA can comprise several cells. Therefore, a broadcast (paging call) in all cells of the LA is used to determine the exact location, i.e. cell, of the MS. Paging is initiated from the VLR using the MAP (PAGE MS) and transformed by the MSC into the paging procedure at the air interface. When an MS receives a paging call, it responds directly and thus allows determination of the current cell.

Thereafter, the VLR instructs the MSC to authenticate the MS and to start ciphering on the signaling channel. Optionally, the VLR can execute a reallocation of the TMSI (TMSI reallocation procedure) during call setup. Only at this point, after the network internal connection has been established (see Section 7.4.4), the connection setup proper can be processed (command COMPLETE CALL from VLR to MSC). The MS is told about the connection request with a SETUP message, and after answering CALL COMPLETE it receives a channel. After ringing (ALERT) and going off-hook, the connection is switched through CONNECT, CONNECT, ACKNOWLEDGE), and this fact is also signaled to the remote exchange (ACM, ANS).

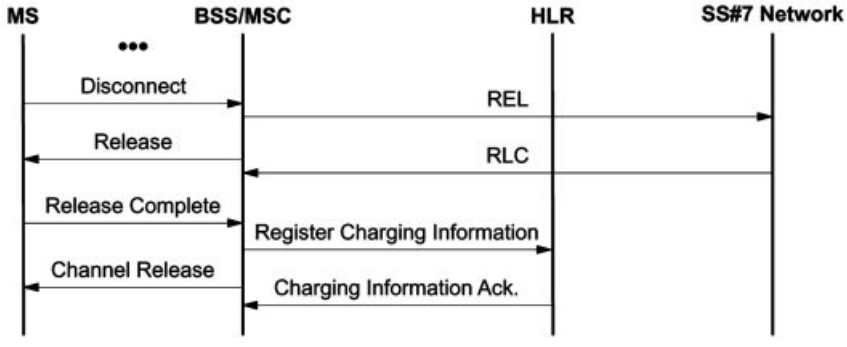

Figure 8.11: Mobile-initiated call termination and storing of charging information

8.3.3 Call Termination

At the air interface, a given call can be terminated either by the mobile equipment or by the network. The taking down of the connection is initiated at the Um interface by means of the CC messages DISCONNECT, RELEASE, and RELEASE COMPLETE. This is followed by an explicit release of occupied radio resources (CHANNEL RELEASE). On the network side, the connection between the involved switching centers (MSC, etc.) is terminated using the ISUP messages REL and RLC in the SS#7 network (Figure 8.11).

After taking down of the connection, information about charges (CHARGING INFORMATION) is stored in the VLR or HLR using the MAP. This charging data can also be required for an incoming call, e.g. if roaming charges are due because the called subscriber is not in his or her home PLMN.

8.3.4 MAP Procedures and Routing for Short Messages

A connectionless relay protocol has been defined for the transport of short messages (see Section 7.4.8) at the air interface, which has a counterpart in the network in a store-and-forward operation for short messages. This forwarding of transport PDUs of the SMS uses MAP procedures. For an incoming short message which arrives from the *Short Message Service Center* (SMS-SC) at a *Short Message Gateway MSC* (SMS-GMSC), the exact location of the MS is the first item that needs to be determined just as for an incoming call. The current MSC of the MS is first obtained with an HLR interrogation (SHORT

MESSAGE ROUTING INFORMATION, Figure 8.12a). The short message is then passed to this MSC (FORWARD SHORT MESSAGE) and is locally delivered after paging and SMS connection setup. Success or failure are reported to the SMS-GMSC in another MAP message (FORWARD ACKNOWLEDGEMENT/ERROR INDICATION) which then informs the service center.

In the reverse case, for an outgoing short message, no routing interrogation is needed, since the SMS-GMSC is known to all MSC, so the message can be passed immediately to the SMS-GMSC (Figure 8.12b).

8.4 Handover

8.4.1 Overview

Handover is the transfer of an existing voice connection to a new base station. There are different reasons for the handover to become necessary. In GSM, a handover decision is made by the network, not the mobile station, and it is based on BSS criteria (received signal level, channel quality, distance between MS and BTS) and on network operation criteria (e.g. current traffic load of the cell and ongoing maintenance work).

The functions for preparation of handover are part of the *Radio Subsystem Link Control*. Above all, this includes the measurement of the channel. Periodically, a mobile station checks the signal field strength of its current downlinks as well as those of the neighboring base stations, including their BSICs. The MS sends measurement reports to its current base station (quality monitoring); see Section 5.5.1. On the network side, the signal quality of the uplink is monitored, the measurement reports are evaluated, and handover decisions are made.

As a matter of principle, handovers are only performed between base stations of the same PLMN. Handovers between BSS in different networks are not allowed. Two kinds of handover are distinguished (Figure 8.13):

- *Intracell Handover*: for administrative reasons or because of channel quality (channel-selective interferences), the mobile station is assigned a new channel within the same cell. This decision is made locally by the *Radio Resource Management* (RR) of the BSS and is also executed within the BSS.
- *Intercell Handover*: the connection to an MS is transferred over the cell boundary to a new BTS. The decision about the time of handover is made by the RR protocol module of the network based on measurement data from MS and BSS. The MSC, however, can participate in the selection of the new cell or BTS. The intercell handover occurs most often when it is recognized from weak signal field strength and bad channel quality (high bit error ratio) that a mobile station is moving near the cell boundary. However, an intercell handover can also occur due to administrative reasons, say for traffic load balancing. The decision about such a *network-directed handover* is made by the MSC, which instructs the BSS to select candidates for such a handover.

Two cases need to be distinguished with regard to participation of network components in the handover, depending on whether the signaling sequences of a handover execution also

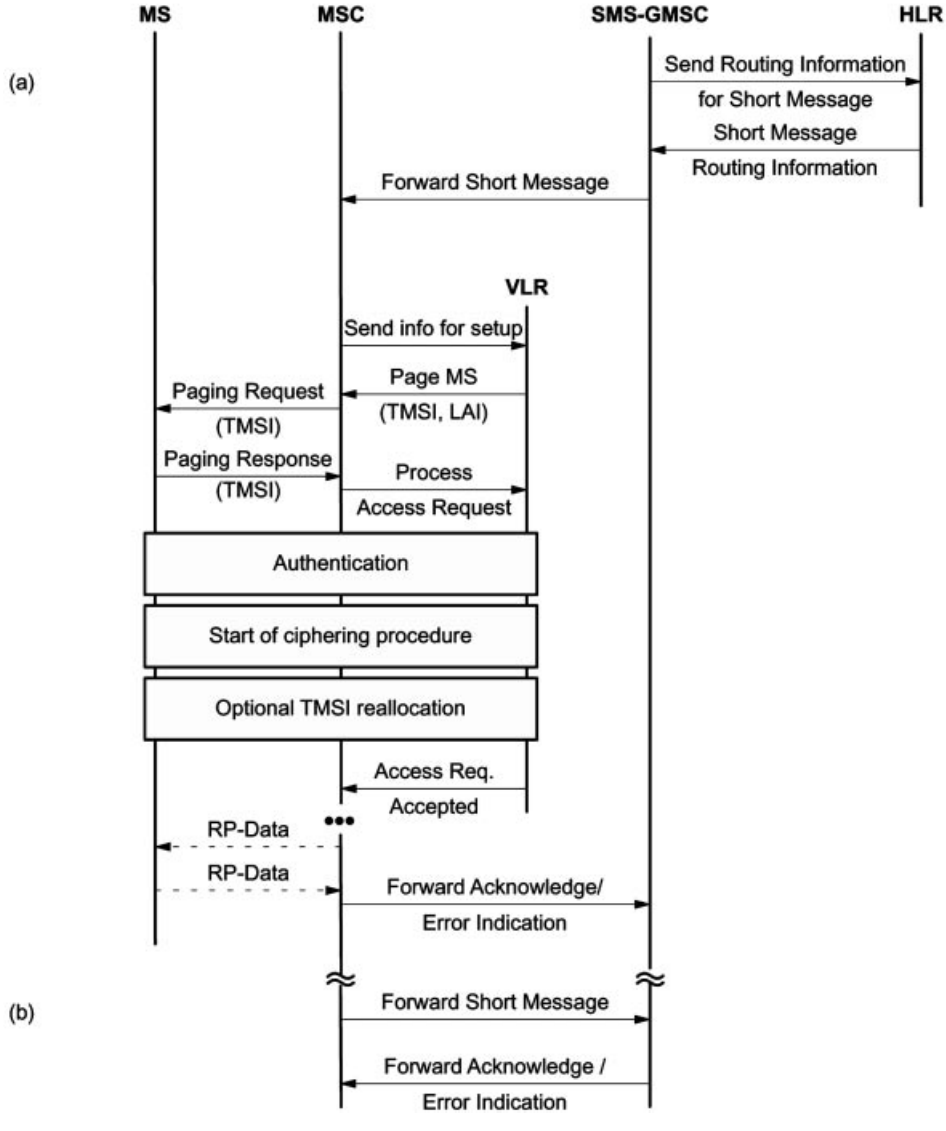

Figure 8.12: Forwarding short messages in a PLMN

involve an MSC. Since the RR module of the network resides in the BSC (see Figure 7.11), the BSS can perform the handover without participation of the MSC. Such handovers occur between cells which are controlled by the same BSC and are called *internal* handover. They can be performed independently by the BSS; the MSC is only informed about the successful execution of internal handovers. All other handovers require participation of at least one MSC, or their BSSMAP and MAP parts, respectively. These handovers are known as *external* handovers.

Participating MSCs can act in the role of MSC-A or MSC-B. MSC-A is the MSC which

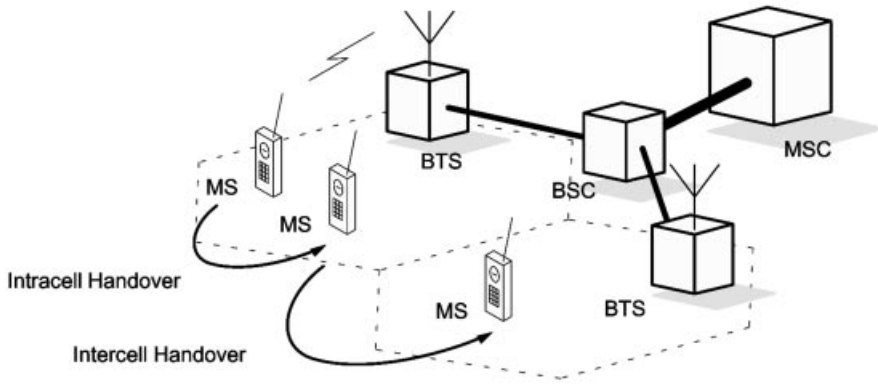

Figure 8.13: Intracell and intercell handover

performed the initial connection setup, and it keeps the MSC-A role and complete control (anchor MSC) for the entire life of the connection. A handover is therefore in general the extension of the connection from the anchor MSC-A to another MSC (MSC-B). In this case, the mobile connection is passed from MSC-A to MSC-B with MSC-A keeping the ultimate control over the connection. An example is presented in Figure 8.14. A mobile station occupies an active connection via BTS1 and moves into the next cell. This cell of BTS2 is controlled by the same BSC so that an internal handover is indicated. The connection is now carried from MSC-A over the BSC and the BTS2 to the mobile station; the connections of BTS1 (radio channel and ISDN channel between BTS and BSC) were taken down. As the mobile station moves on to the cell handled by BTS3, it enters a new BSS which requires an external handover. Besides, this BSS belongs to another MSC, which now has to play the role of MSC-B. Logically, the connection is extended from MSC-A to MSC-B and carried over the BSS to the mobile station. At the next change of the MSC, the connection element between MSC-A and MSC-B is taken down, and a

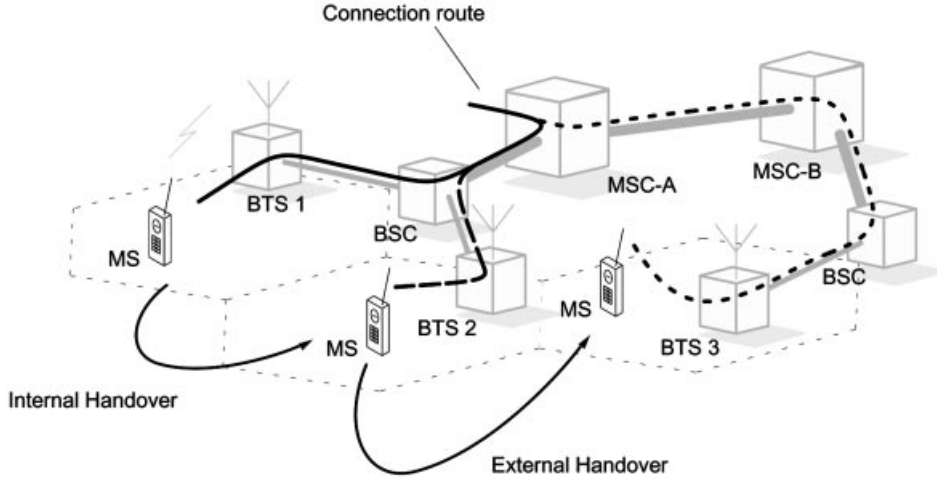

Figure 8.14: Internal and external handover

connection to the new MSC from MSC-A is set up. Then the new MSC takes over the role of MSC-B.

8.4.2 Intra-MSC Handover

The basic structure for an external handover is the handover between two cells of the same MSC (Figure 8.15). The mobile station continually transmits measurement reports with channel monitoring data on its SACCH to the current base station (BSS 1). Based on these measurement results, the BSS decides when to perform a handover and requests this handover from the MSC (message HANDOVER REQUIRED). The respective measurement results can be transmitted in this message to the MSC, to enable its participation in the handover decision. The MSC causes the new BSS to prepare a channel for the handover, and frees the handover to the mobile station (HANDOVER COMMAND), as soon as the reservation is acknowledged by the new BSS. The mobile station now reports to the new BSS (HANDOVER ACCESS) and receives information about the physical channel properties. This includes synchronization data like the new timing advance value and also the new transmitter power level. Once the mobile station is able to occupy the channel successfully, it acknowledges this fact with a message HANDOVER COMPLETE. The resources of the old BSS can then be released.

8.4.3 Decision Algorithm for Handover Timing

The basis for processing a successful handover is a decision algorithm which uses measurement results from mobile and base station to identify possible other base stations as targets for handovers and which determines the optimal moment to execute the handover. The objective is to keep the number of handovers per cell change as small as possible. Ideally, there should not be more than one handover per cell change. In reality, this is often not achievable. When a mobile station leaves the radio range of a base station and enters one of a neighboring station, the radio conditions are often not very stable, so that several handovers must be executed before a stable state is reached. Simulation results in [44] and [36] give a mean value of about 1.5–5 handovers per cell change.

Since every handover incurs not only increased traffic load for the signaling and transport system but also reductions in speech quality, the importance of a well-dimensioned handover decision algorithm is obvious, an algorithm which also takes into account the momentary local conditions. This is also a reason for GSM not having standardized a uniform algorithm for the determination of the moment of the handover. For this decision about when to perform a handover, network operators can develop and deploy their own algorithms which are optimally tuned for their networks. This is made possible through standardizing only the signaling interface that defines the processing of the handover and through transferring the handover decision to the BSS. The GSM handover is thus a *network-originated handover* as opposed to a *mobile-originated handover*, where the handover decision is made by the mobile station. An advantage of this handover approach is that the software of the mobile station need not be changed when the handover strategy or the handover decision algorithm is changed in all or parts of the network. Even though the GSM standard does not prescribe a mandatory handover decision algorithm, a simple

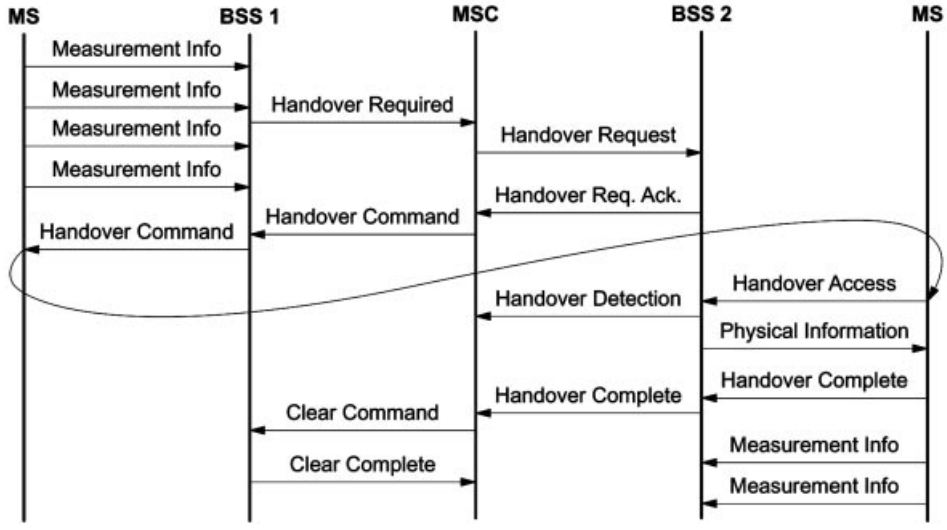

Figure 8.15: Principal signaling sequence for an intra-MSC handover

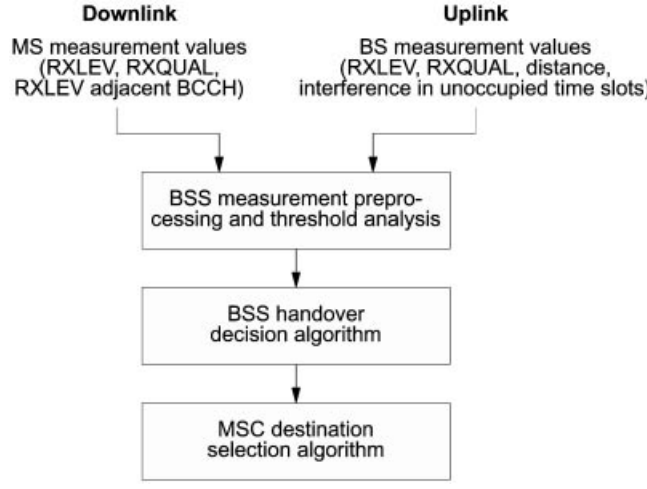

Figure 8.16: Decision steps in a GSM handover

algorithm is proposed, which can be selected by the network operator or replaced by a more complex algorithm.

In principle, a GSM handover always proceeds in three steps (Figure 8.16), which are based on the measurement data provided by the mobile station over the SACCH, and on the measurements performed by the BSS itself. Foremost among these data items are the current channel's received signal level (RXLEV) and the signal quality (RXQUAL), both on the uplink (measured by the BSS) and on the downlink (measured by the MS). In order to be able to identify neighboring cells as potential targets for a handover, the mobile station measures in addition the received signal level RXLEV_CELL(n) of up to 16

neighboring base stations. The RXLEV values of the six base stations which can be received best are reported every 480 ms to the BSS. Further criteria for the handover decision algorithm are the distance between MS and BTS measured via the *Timing Advance* (TA) of the *Adaptive Frame Alignment* (see Section 5.3.2) and measurements of the interference in unused time slots. A new value of each of these measurements is available every 480 ms.

Measurement preprocessing calculates average values from these measurements, whereby at least the last 32 values of RXLEV and RXQUAL must be averaged. The resulting mean values are continuously compared with thresholds (see Table 8.1) after every SACCH interval.

These threshold values can be configured individually for each BSS through management interfaces of the OMSS (see Section 3.3.4). The principle used for the comparison of measurements with the threshold is to conduct a so-called Bernoulli experiment: if out of the last N_i mean values of a criterion i more than P_i go under (RXLEV) or over (RXQUAL, MS_RANGE) the threshold, then a handover may be a necessary. The values of N_i and P_i can also be configured through network management. Their allowed range is defined as the interval [0; 31].

In addition to these mean values, a BSS can calculate the current power budget PBGT(n), which represents a measure for the respective path loss between mobile station and current base station or a neighboring base station n. Using this criterion, a handover can always be caused to occur to the base station with the least path loss for the signals from or to the mobile station. The PBGT takes into consideration not only the RXLEV_DL of the current downlink and the RXLEV_NCELL(n) of the neighboring BCCH but also the maximal transmitter power P (see Table 5.8) of a mobile station, the maximal power MS_TXPWR_MAX allowed to a mobile station in the current cell, and the maximal power MS_TXPWR_MAX(n) allowed to mobiles in the neighboring cells. In addition, the calculation uses the value PWR_C_D, which is the difference between maximal transmitter power on the downlink and current transmitter power of the BTS in the downlink, a measure for the available power control reserve.

Thus the power budget for a neighboring base station n is calculated as follows:

$$\begin{aligned}\text{PBGT}(n) = &(\text{Minimum}\ (\text{MS\_TXPWR\_MAX}, P) - \text{RXLEV\_DL} - \text{PWR\_C\_D}) \\ &- (\text{Minimum}\ (\text{MS\_TXPWR\_MAX}(n), P) - \text{RXLEV\_NCELL}(n))\end{aligned}$$

A handover to a neighboring base station can be requested, if the power budget is PBGT(n) > 0 and greater than the threshold HO_MARGIN(n). The causes for handover which are possible using these criteria are summarized in Table 8.2. As can be seen, the signal criteria of the uplink and downlink as well as the distance from the base station and power budget can lead to a handover.

The BSS makes a handover decision by first determining the necessity of a handover using the threshold values of Table 8.1. In principle, one can distinguish three categories:

- Handover because of more favorable path loss conditions
- Mandatory intercell handover
- Mandatory intracell handover

Table 8.1 Threshold values for the GSM handover

Threshold value	Typical value	Meaning
L_RXLEV_UL_H	−103 to −73 dBm	Upper handover threshold of received signal level in uplink
L_RXLEV_DL_H	−103 to −73 dBm	Upper handover threshold of received signal level in downlink
L_RXLEV_UL_IH	−85 to −40 dBm	Lower(!) received signal level threshold in uplink for internal handover
L_RXLEV_DL_IH	−85 to −40 dBm	Lower(!) received signal level threshold in downlink for internal handover
RXLEV_MIN(n)	approx. −85 dBm	Minimum required RXLEV of BCCH of cell n to perform a handover to this cell
L_RXQUAL_UL_H	–	Lower handover threshold of bit error ratio in uplink
L_RXQUAL_DL_H	–	Lower handover threshold of bit error ratio in downlink
MS_RANGE_MAX	2 to 35 km	Maximum distance between mobile and base station
HO_MARGIN(n)	0 to 24 dB	Hysteresis to avoid multiple handovers between two cells

Table 8.2 Handover causes

Handover cause	Meaning
UL_RXLEV	Uplink received signal level too low
DL_RXLEV	Downlink received signal level too low
UL_RXQUAL	Uplink bit error ratio too high
DL_RXQUAL	Downlink bit error ratio too high
PWR_CTRL_FAIL	Power control range exceeded
DISTANCE	MS to BTS distance too high
PBGT(n)	Lower value of path loss to BTS n

Situations where a neighboring base station shows more favorable propagation conditions and therefore lower path loss, do not necessarily force a handover. Such potential handover situations to a neighboring cell are discovered through the PBGT(n) calculations. To make a handover necessary, the power budget of the neighboring cell must be greater than the threshold HO_MARGIN(n).

The recognition of a mandatory handover situation (Figure 8.17) within the framework of

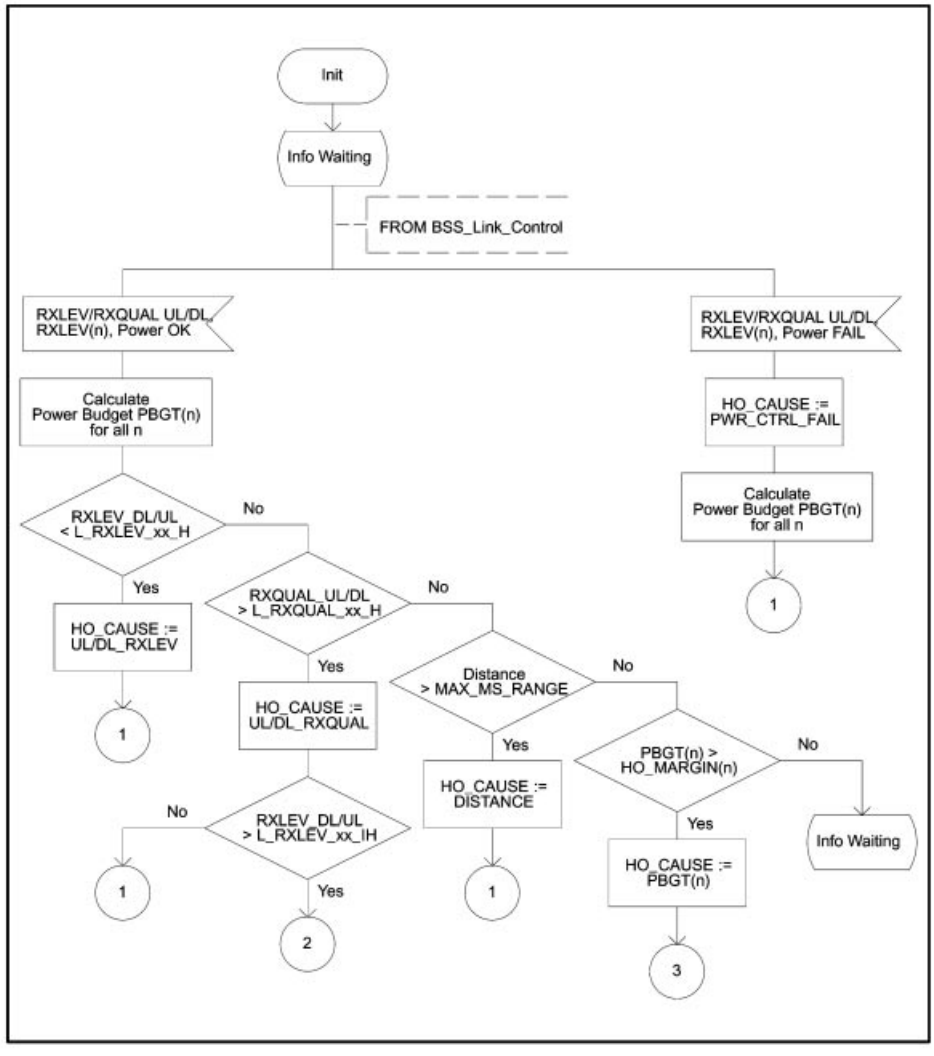

Figure 8.17: Detection of mandatory handover (abbreviated)

the *Radio Subsystem Link Control* (see also Section 5.5 and Figure 5.19) is based on the received signal level and signal quality in uplink and downlink as well as on the distance between MS and BTS. Going over or under the respective thresholds always necessitates a handover. Here are the typical situations for a mandatory handover:

- The received signal level in the uplink or downlink (RXLEV_UL/RXLEV_DL) drops below the respective handover threshold value (L_RXLEV_UL_H/L_RXLEV_DL_H) and the power control range has been exhausted, i.e. the MS and/or the BSS have reached their maximal transmitter power (see Section 5.5.2).
- The bit error ratio as a measure of signal quality in uplink and/or downlink (RXQUAL_UL/RXQUAL_DL) exceeds the respective handover threshold value (L_RXQUAL_

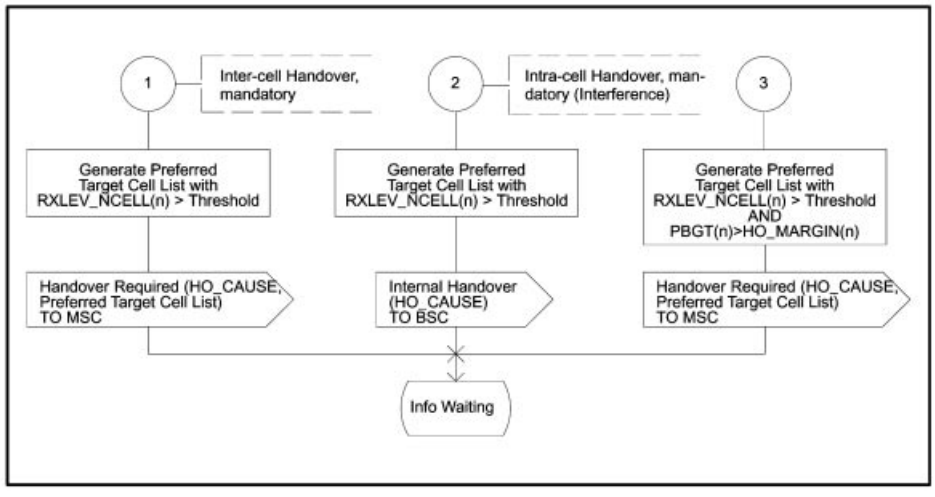

Figure 8.18: Completion of handover decision in the BSS

UL_H/L_RXQUAL_DL_H), while at the same time the received signal level drops into the neighborhood of the threshold value.

- The maximum distance to the base station (MAX_MS_RANGE) has been reached.

A handover can also become mandatory, even if the handover thresholds are not exceeded, if the lower thresholds of the transmitter power control are exceeded (L_RXLEV_xx_P/ L_RXQUAL_xx_P, see Table 5.9), even though the maximum transmitter power has been reached already. The cause of handover indicated is the failure of the transmitter power control (PWR_CTR_FAIL, see Table 8.2).

A special handover situation exists, if the bit error ratio RXQUAL as a measurement for signal quality in uplink and/or downlink exceeds its threshold and at the same time the received signal level is greater than the thresholds L_RXLEC_UL_IH/L_RXLEC_DL_IH. This strongly hints at an existing severe cochannel interference. This problem can be solved with an (internal) intracell handover, which the BSS can perform on its own without support from the MSC. It is also considered as a mandatory handover.

If the BSS has detected a handover situation, a list of candidates as possible handover targets is assembled using the BSS decision algorithm. For this purpose, one first determines which BCCH of the neighboring cell n is received with sufficient signal level:

$$\text{RXLEV\_NCELL}(n) > (\text{RXLEV\_MIN}(n) + \text{Maximum}\ (0, (\text{MS\_TXPWR\_MAX}(n) - P)))$$

The potential handover targets are then assembled in an ordered list of preferred cells according to their path loss compared to the current cell (Figure 8.18). For this purpose, the power budget of the neighboring cells in question is again evaluated:

$$\text{PBGT}(n) - \text{HO\_MARGIN}(n) > 0$$

All cells n which are potential targets for a handover due to RXLEV_NCELL(n) and lower

path loss than the current channel are then reported to the MSC with the message HANDOVER REQUIRED (Figure 8.18) as possible handover targets. This list is sorted by priority according to the difference (PBGT(n) − HO_MARGIN(n)). The same message HANDOVER REQUIRED is also generated if the MSC has sent a message HANDOVER CANDIDATE ENQUIRY to the BSS.

The conditions at a cell boundary in the case of exhausted transmitter power control (PWR_C_D = 0) are shown in Figure 8.19 with a mobile station moving from the current cell to a cell B. The threshold RXLEV_MIN(B) is reached very early; however, the handover is somewhat moved in the direction of cell B because of the positive HO_MARGIN(B) for the power budget. When moving in the opposite direction, the handover would be delayed in the other direction due to HO_MARGIN(A) of cell A. This has the effect of a hysteresis which reduces repeated handovers between both cells due to fading (ping-pong handover). Besides varying radio conditions (fading due to multipath propagation, shadowing, etc.) there are many other sources of error with this kind of handover. Recognize, on one hand, that there are substantial delays between measurement and reaction due to the averaging process. This leads to executing the handover too late on a few occasions. It is more important, however, that the current channel is compared with the BCCH of the neighboring cells rather than the traffic channel to be used after the handover decision, which could suffer from different propagation conditions (frequency-selective fading etc.). Finally, the MSC decides about the target cell of the handover. This decision takes into consideration the following criteria in decreasing order of priority: handover due to signal quality (RXQUAL), received signal level (RXLEV), distance, and path loss (PBGT). This prioritization is especially effective when there are not enough traffic channels available and handover requests are competing for the available channels.

The standard explicitly points out that all measurement results must be sent with the

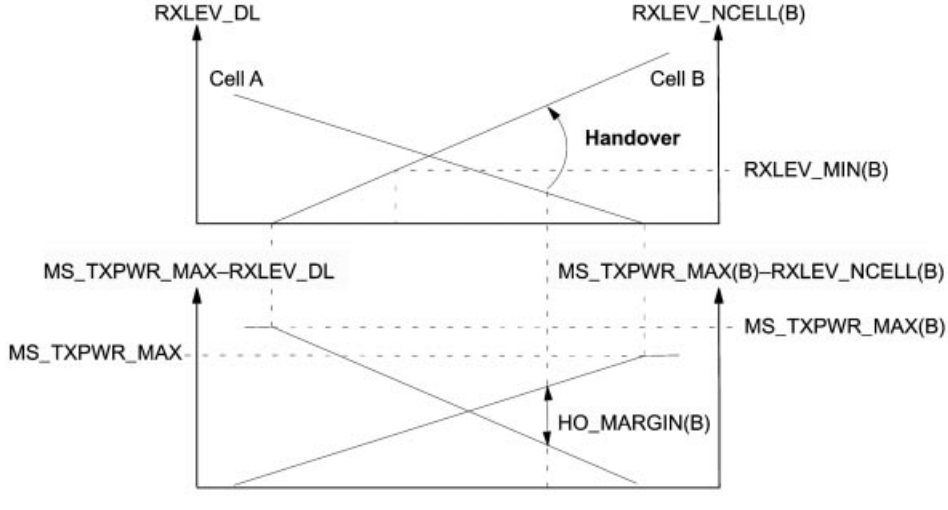

Figure 8.19: Handover criteria for exhausted transmitter power control

message HANDOVER REQUIRED to the MSC, so that in the end the option remains open to implement the complete handover decision algorithm in the MSC.

8.4.4 MAP and Inter-MSC Handover

The most general form of handover is the inter-MSC handover. The mobile station moves over a cell boundary and enters the area of responsibility of a new MSC. The handover caused by this move requires communication between the involved MSCs. This occurs through the SS#7 using transactions of the *Mobile Application Part* (MAP).

8.4.4.1 Basic Handover between two MSCs

The principal sequence of operations for a basic handover between two MSCs is shown in Figure 8.20. The MS has indicated the conditions for the handover, and the BSS requests the handover from MSC-A (HANDOVER REQUIRED). MSC-A decides positively for a handover and sends a message PERFORM HANDOVER to MSC-B. This message contains the necessary data to enable MSC-B to reserve a radio channel for the MS. Above all, it identifies the BSS which is to receive the connection. MSC-B assigns a handover number and tries to allocate a channel for the MS. If a channel is available, the response RADIO

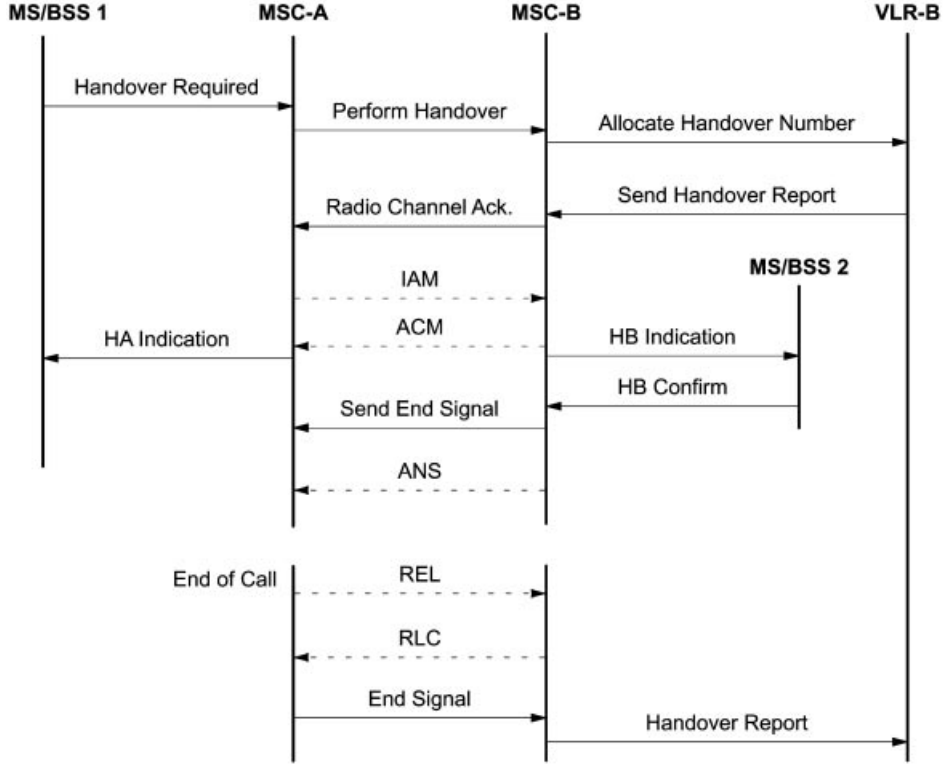

Figure 8.20: Principal operation of a basic handover

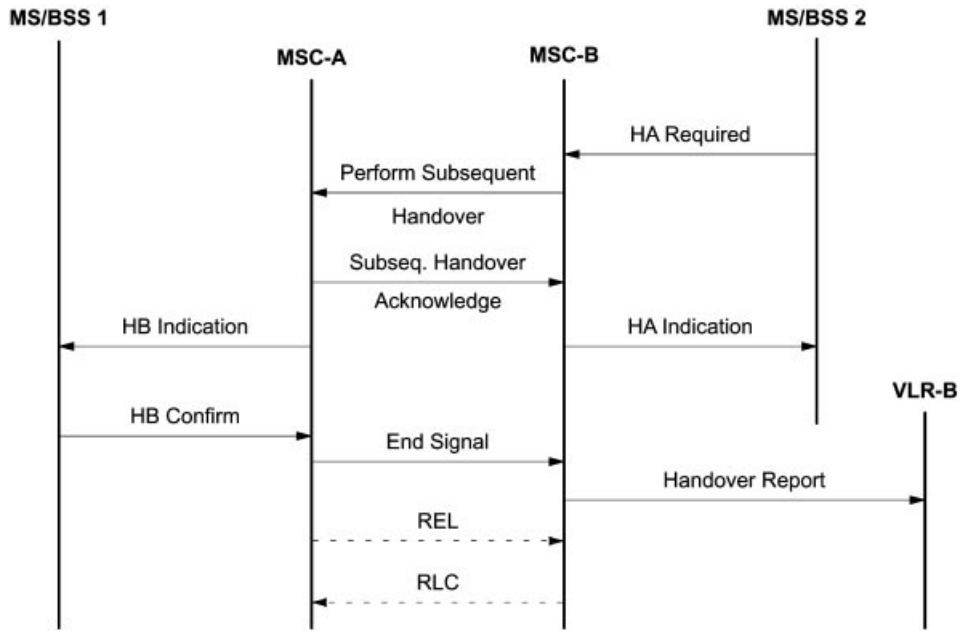

Figure 8.21: Principle of subsequent handover from MSC-B to MSC-A (handback)

CHANNEL ACKNOWLEDGE contains the new MSRN to the MS and the designation of the new channel. If no channel is available, this is also reported to MSC-A which then terminates the handover procedure.

When a RADIO CHANNEL ACKNOWLEDGE is successful, an ISDN channel is switched through between the two MSCs (ISUP messages IAM and ACM), and both MSCs send an acknowledgement to the MS (HA INDICATION, HB INDICATION). The MS then resumes the connection on the new channel after a short interruption (HB CONFIRM). MSC-B then sends a message SEND END SIGNAL to MSC-A and thus causes the release of the old radio connection. After the end of the connection (ISUP messages REL, RLC), MSC-A generates a message END SIGNAL for MSC-B which then sends a HANDOVER REPORT to its VLR.

8.4.4.2 Subsequent Handover

After a first basic handover of a connection from MSC-A to MSC-B, a mobile station can move on freely. Further intra-MSC handovers can occur (Figure 8.15), which are processed by MSC-B.

If, however, the mobile station leaves the area of MSC-B during this connection, a *Subsequent Handover* becomes necessary. Two cases are distinguished: in the first case, the mobile station returns to the area of MSC-A, whereas in the second case it enters the area of a new MSC, now called MSC-B′. In both cases, the connection is newly routed from MSC-A. The connection between MSC-A and MSC-B is taken down after a successful subsequent handover.

A subsequent handover from MSC-B back to MSC-A is also called *handback* (Figure

8.21). In this case, MSC-A, as the controlling entity, does not need to assign a handover number and can search directly for a new radio channel for the mobile station. If a radio channel can be allocated in time, both MSCs start their handover procedures at the air interface (HA/HB INDICATION) and complete the handover. After completion, MSC-A terminates the connection to MSC-B. The message END SIGNAL terminates the MAP

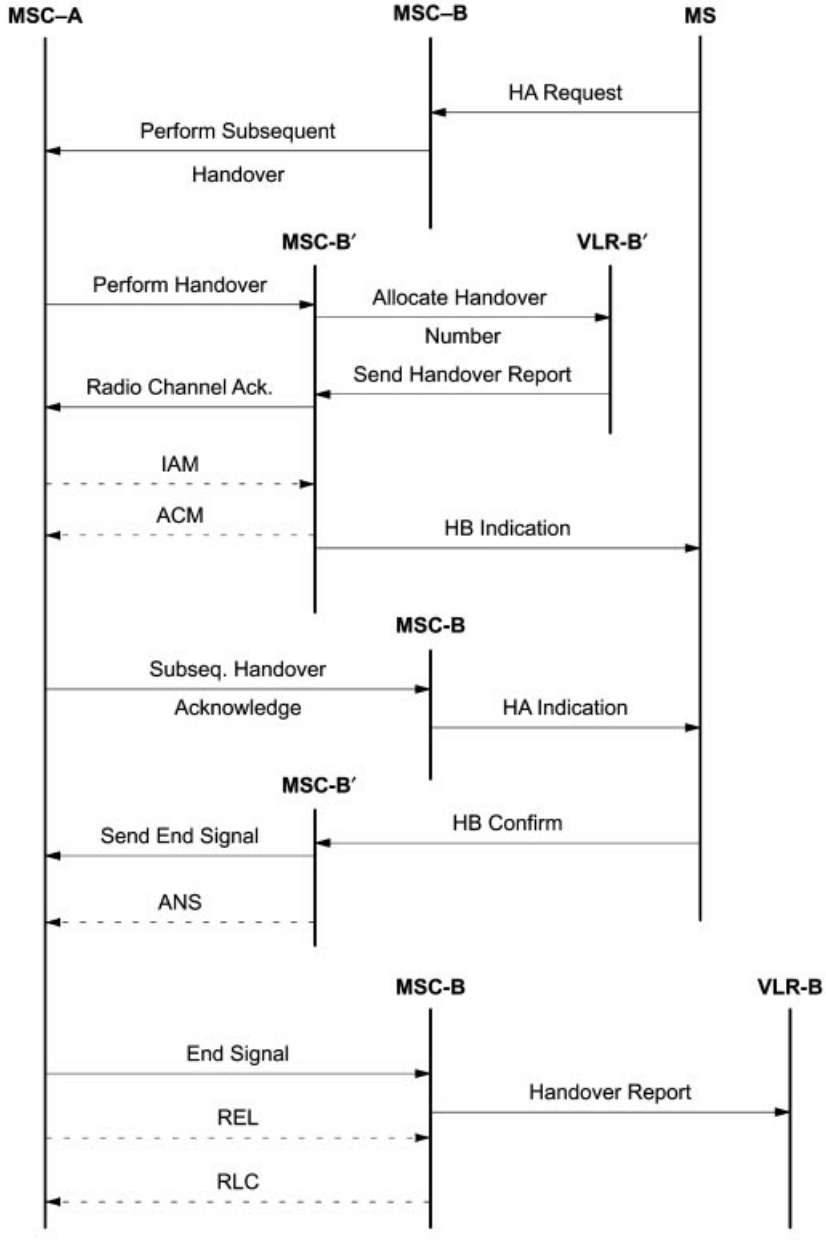

Figure 8.22: Principle of subsequent handover from MSC-B to MSC-B′

process in MSC-B and causes a HANDOVER REPORT to be sent to the VLR of MSC-B; the ISUP message RELEASE releases the ISDN connection.

The procedure for a subsequent handover from MSC-B to MSC-B is more complicated. It consists of two parts:

- A *Subsequent Handover* from MSC-B to MSC-A
- A *Basic Handover* between MSC-A and MSC-B′

The principal operation of this handover is illustrated in Figure 8.22.

In this case, MSC-A recognizes from the message PERFORM SUBSEQUENT HANDOVER, sent by MSC-B, that it is a case of handover to an MSC-B′, and it initiates a *Basic Handover* to MSC-B′. MSC-A informs MSC-B after receiving the ISUP message ACM from MSC-B about the start of handover at MSC-B′ and thereby frees the handover procedure at the radio interface from MSC-B. Once MSC-A receives the MAP message SEND END SIGNAL from MSC-B, it considers the handover as complete, sends the message END SIGNAL to MSC-B to terminate the MAP procedure and cancels the ISDN connection.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

9 Data Communication and Networking

9.1 Reference Configuration

GSM was conceived in accordance with the guidelines of ISDN. Therefore, a reference configuration is also defined for GSM systems, similar to the one used in ISDN systems. Using the reference configuration, one gets an impression of the range of services and the kinds of interfaces to be provided by mobile stations. Furthermore, the reference configuration indicates at which interface which protocols or functions terminate and where adaptation functions may have to be provided.

The GSM reference configuration comprises the functional blocks of a mobile station (Figure 9.1) at the user–network interface Um. The mobile equipment is subdivided into a *Mobile Termination* (MT) and various combinations of *Terminal Adapter* (TA) and *Terminal Equipment* (TE), depending on the kind of service access and interfaces offered to the subscriber.

At the interface to the mobile network, the air interface Um, MT units are defined. An integrated mobile speech or data terminal is represented only by an MT0. The MT1 unit goes one step further and offers an interface for standard-conforming equipment at the ISDN S reference point, which can be connected directly as end equipment. Likewise, normal data terminal equipment with a standard interface (e.g. V.24) can be connected via a TA and this way use the mobile transmission services. Finally, the TA functionality has been integrated into units of type MT2.

At the S or R reference point, the GSM bearer or data services are available (access points 1 and 2 in Figure 9.1), whereas the teleservices are offered at the user interfaces of the TE (access point 3, Figure 9.1). Among the bearer services besides the transmission of digitized speech, there are circuit-switched and packet-switched data transmission. Typical teleservices besides telephony are, for example, *Short Message Service* (SMS), Group 3 fax service, or emergency calls from anywhere.

9.2 Overview of Data Communication

Voice service needs only a switched-through physical connection, which changes its bit rate in the BSS due to the speech transcoding in the TRAU. From the MSC on, the speech

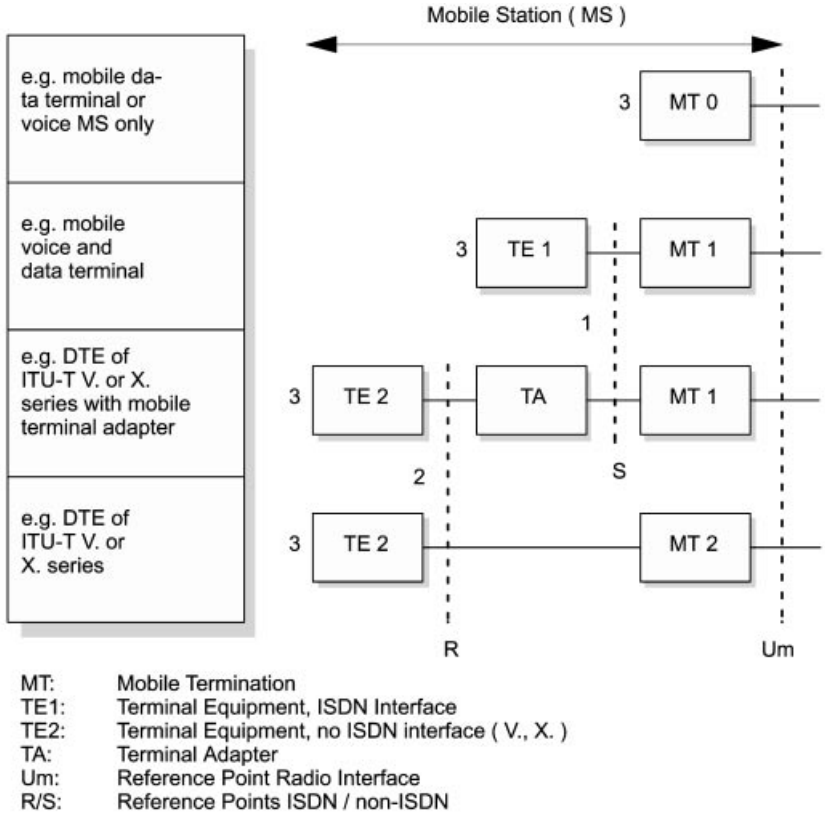

Figure 9.1: GSM reference configuration

signals in GSM networks are transported in standard ISDN format with a bit rate of 64 kbit/s. In comparison, realizing data services and the other teleservices like Group 3 fax is considerably more complicated. Because of the psychoacoustic compression procedures of the GSM speech codec, data cannot be simply transmitted as a voiceband signal as in the analog network – a complete reconstruction of the data signal would not be possible. Therefore, a solution to digitize the voiceband signal similar to ISDN is not possible. Rather the available digital data must be transmitted in unchanged digital form by avoiding speech codecs in the PLMN, as is possible in the ISDN. Here we have to distinguish two areas where special measures have to be taken: first, the realization of data and teleservices at the air interface or within the mobile network, and second, at the transition between mobile and fixed network with the associated mapping of service features. These two areas are illustrated schematically in Figure 9.2.

A PLMN offers transparent and nontransparent services. These bearer services carry data between the MT of the mobile station and the *Interworking Function* (IWF) of the MSC. For the realization of bearer services, the individual units of the GSM network define several functions:

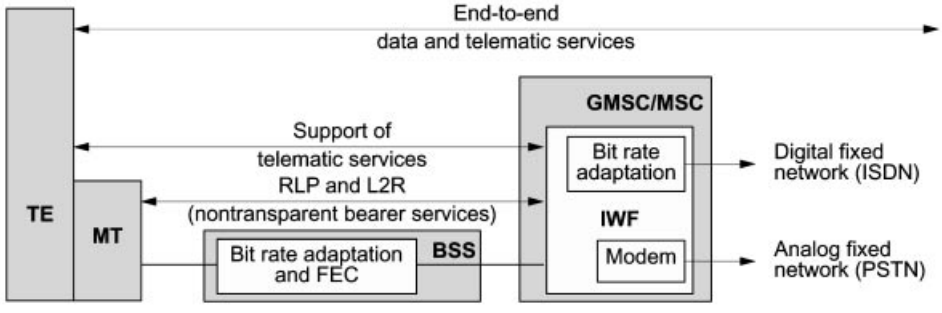

Figure 9.2: Bearer services, interworking, and teleservices

- *Bit Rate Adaptation* (RA)
- *Forward Error Correction* (FEC)
- ARQ error correction with the *Radio Link Protocol* (RLP)
- Adaptation protocol *Layer 2 Relay* (L2R)

For the transmission of transparent and nontransparent data, several rate adaptation stages are required to adapt the bit rates of the bearer services to the channel data rates of the radio interface (traffic channels with 3.6 kbit/s, 6 kbit/s, and 12 kbit/s) and to the transmission rate of the fixed connections. A bearer service for data transmission can be realized in the following two ways: 9.6 kbit/s data service requires a full-rate traffic channel, all other data services can either be realized on a full-rate or half-rate channel. A mobile station must support both types of data traffic channels, independent of what is used for speech transmission. The data signals are transcoded first from the user data rate (9.6 kbit/s, 4.8 kbit/s, 2.4 kbit/s, etc.) to the channel data rate of the traffic channel, then further to the data rate of the fixed connection between BSS and MSC (64 kbit/s) and finally back to the user data rate. This bit rate adaptation (RA) in GSM corresponds in essence to the bit rate adaptation in the ITU-T standard V.110, which specifies the support of data terminals with an interface according to the V. series on an ISDN network [34].

On the radio channel, data is protected through the forward error correction procedures (FEC) of the GSM PLMN; and for nontransparent data services, data is additionally protected by the ARQ procedure of RLP on the whole network path between MT and MSC. Thus RLP is terminated in the MT and MSC. The protocol adaptation to RLP of Layers 1 and 2 at the user interface is done by the *Layer 2 Relay* (L2R) protocol.

Finally, the data is passed on from MSC or GMSC over an *Interworking Function* (IWF) to the respective data connection. The bearer services of the PLMN are transformed to the bearer services of the ISDN or another PLMN in the IWF, which is usually activated in an MSC near the MS, but could also reside in the GMSC of the network transition. In the case of ISDN this transition is relatively simple, since it may just require a potential bit rate adaptation. In the case of an analog PSTN, the available digital data must be transformed by a modem into a voiceband signal, which can then be transmitted on an analog voiceband of 3.1 kHz.

The bearer services realized in this way can offer the protocols that may be required for the support of teleservices between TE and IWF. An example is the fax adaptation protocol.

The fax adapter is a special TE which maps the Group 3 fax protocols with their analog physical interface upon the digital bearer services of a GSM PLMN. Thus, after another adaptation into an analog fax signal in the IWF of the MSC, it enables the end-to-end transfer of fax messages according to the ITU-T Standard T.30.

A possible interworking scenario for transparent data services of GSM with transition to a PSTN is shown in Figure 9.3. The analog circuit-switched connection of the PSTN represents a transparent channel which can be used to transport arbitrary digital data signals in the voiceband. In the analog network, a subscriber selects telephone or modem depending on whether he or she wants to transmit speech or data. In the PLMN, however, the channel coding has to be changed for different services (error protection for different bearer services, see Section 6.2). The bit rate adaptation has to be activated and the speech coding deactivated. In the IWF of the MSC, besides the bit rate adaptation, a modem needs to be added for data communication with the partner in the fixed network. In the GSM network, voice signals therefore take a different path than data signals; in the case shown in Figure 9.3, the data signals are directed from the IWF to the modem, where they are digitized, then passed on after bit rate adaptation to transmission on the radio channel. In the opposite direction, the IWF passes the PCM-coded information on an ISDN channel (64 kbit/s) to the GMSC. From there it is transformed into an analog signal in a network transition switching unit and carried as a voiceband signal in the PSTN to the analog terminal.

After these introductory remarks, the GSM data and teleservices and their realization are discussed in more detail in the following sections.

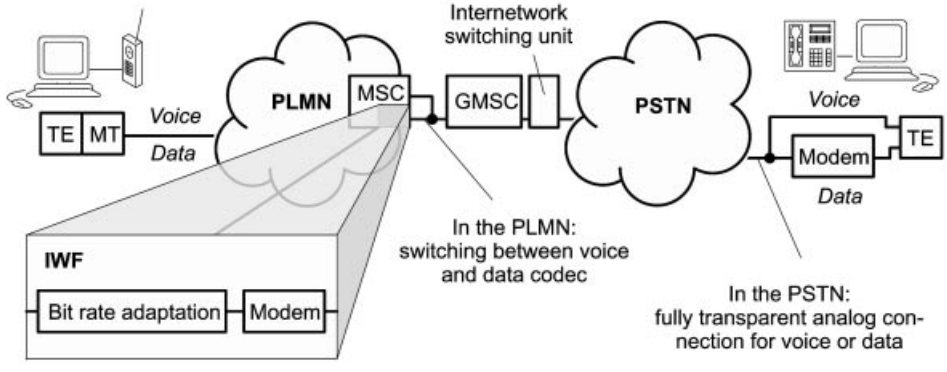

Figure 9.3: Interworking scenario PLMN-PSTN for transparent data services

9.3 Service Selection at Transitions between Networks

A specific interworking problem arises for data services between PLMN and ISDN/PSTN networks. Mobile-terminated calls require that the calling subscriber (ISDN or PSTN subscriber) tells the GMSC which service (speech, data, fax, etc.) he or she wants to use. In ISDN, a *Bearer Capability* (BC) information element would have to be included in the SETUP message. This BC information element could then be passed on by the

network transition switching unit to the GMSC and from there to the local MSC, which could thus activate the required resources. In the course of call processing (CC, see Section 7.4.5), the mobile station would also be informed about the kind of service requested by the calling subscriber and could activate the needed functions. The calling subscriber, however, if there is no ISDN signaling as in analog networks, is not able to do this kind of BC signaling. The service selection therefore has to use another mechanism. The GSM standard proposes two possible solutions, which are always to be used for service selection independent of the type of originating network (ISDN or PSTN).

- *Multinumbering*: the home network with this option assigns to each mobile subscriber several MSISDN numbers, each with a specific *Bearer Capability* (BC), which can be obtained at each call from the HLR. This way the service that an incoming call wants is always uniquely determined. The BC information element is given to the mobile station when the call is being set up, so the MS can decide based on its technical features whether it wants to accept the call.
- *Single numbering*: only a single MSISDN is assigned to the mobile subscriber, and there is no BC information element transmitted with an incoming call. The MS recognizes then that a specific BC is needed when a call is accepted and requests the BC from the MSC. If the network is able to offer the requested service, the call is switched through.

Usually, the multinumbering solution is favored, since one can already verify at call arrival time in the MSC whether the requested resources are available, and the MSC side can decide about accepting the call. There is no negotiation about the BC between MS and MSC, so no radio resources are occupied unnecessarily, and the call set-up phase is not extended.

9.4 Bit Rate Adaptation

Five basic traffic channels are available in GSM for the realization of bearer services: TCH/H2.4, TCH/H4.8, TCH/F2.4, TCH/F4.8, TCH/F9.6 (see Tables 5.2 and 6.2) with bit rates of 3.6 kbit/s, 6 kbit/s, and 12 kbit/s. In recent standardization efforts, a TCH/F14.4 has also been defined. The bearer services (Table 4.2) with bit rates from 300 bit/s up to 9.6 kbit/s must be realized on these traffic channels. Furthermore, on the fixed connections of the GSM network, the data signals are transmitted with a data rate of 64 kbit/s.

The terminals connected at reference point R have the conventional asynchronous and synchronous interfaces. The data services at these interfaces work at bit rates as realized by GSM bearer services. Therefore, the data terminals at the R reference point have to be bit rate adapted to the radio interface. This bit rate adaptation is derived from the V.110 standard used in ISDN in which the bit rates of the synchronous data streams are going through a two-step procedure; first, frames are formed at an intermediate rate which is a multiple of 8 kbit/s; this stream is converted to the channel bit rate of 64 kbit/s [7]. The asynchronous services are preprocessed by a stuffing procedure using stop bits to form a synchronous data stream.

A V.110 procedure modified according to the requirements of the air interface is also used in GSM. In essence, GSM performs a transformation of the data signals from the user data rate (e.g. 2.4 kbit/s or 9.6 kbit/s) at the R reference point to the intermediate data rate

(8 kbit/s or 16 kbit/s) and finally to the ISDN bit rate of 64 kbit/s. The adaptation function from user to intermediate rate is called RA1; the adaptation function from intermediate rate to ISDN is called RA2. A GSM-specific bit rate adaptation step is added between the intermediate rate and the channel data rate (3.6 kbit/s, 6 kbit/s, or 12 kbit/s) of the traffic channel at the reference point Um of the air interface. This adaptation function from intermediate to channel bit rate is designated as RA1/RA1′. An adaptation function RA1′ performs the direct adaptation from user to channel data rate without going through the intermediate data rate. Table 9.1 gives an overview of the bit rates at the reference points and the intermediate data rates between the RA modules.

Table 9.1: Data rates for GSM bit rate adaptation

Interface	Data rate (kbit/s)		Interface (kbit/s)	
	User	Intermediate	Radio	S
Reference point	R	–	Um	S
RA1	≤2.4	8		
RA1	4.8	8		
RA1	9.6	16		
RA2		8		64
RA2		16		64
RA1/RA1′		8	3.6	
RA1/RA1′		8	6	
RA1/RA1′		16	12	

Adaptation frames are defined for the individual bit rate adaptation steps. These frames contain signaling and synchronization data besides the user data. They are defined based on V.110 frames, and one distinguishes three types of GSM adaptation frames according to their length (36 bits, 60 bits, and 80 bits) as shown in Figures 9.4 and 9.5.

The conversion of data signals from user to intermediate rate in the RA1 stage uses the regular 80-bit frame of the V.110 standard. In this adaptation step, groups of 48 user data bits are supplemented with 17 fill bits and 15 signaling bits to form an 80-bit V.100 frame. Because of the ratio 0.6 of user data to total frame length, this adaptation step converts user data rates of 4.8 kbit/s into 8 kbit/s and from 9.6 kbit/s to 16 kbit/s. All user data frames of less then 4.8 kbit/s are "inflated" to a data signal of 4.8 kbit/s by repeating the individual data bits; for example, a 2.4 kbit/s signal all bits are doubled, or with a 600 bit/s signal the bits are written eight times into an RA1 frame.

At the conversion of the intermediate data rate to the channel data rate in the RA1/RA1 stage, the 17 fill bits and 3 of the signaling bits are removed from the RA1 frame, since they are only used for synchronization and not needed for transmission across the air interface. This yields a modified V.100 frame of length 60 bits (Figure 9.5), and the data rate is adapted from 16 kbit/s to 12 kbit/s or from 8 kbit/s to 6 kbit/s, respectively.

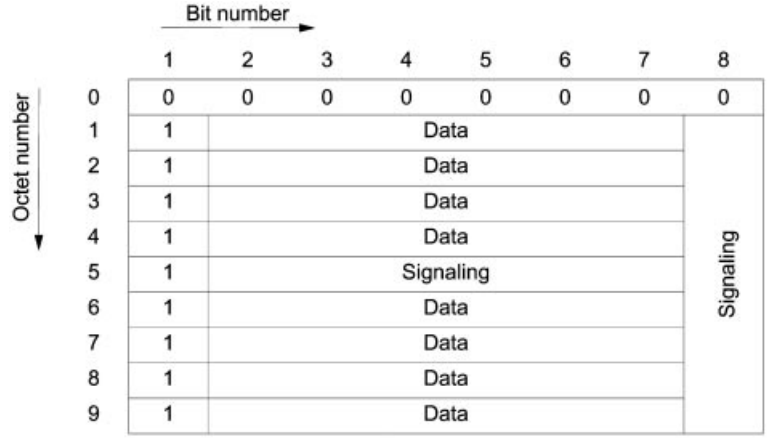

Figure 9.4: V.110 80-bit adaptation frame for the RA1 stage

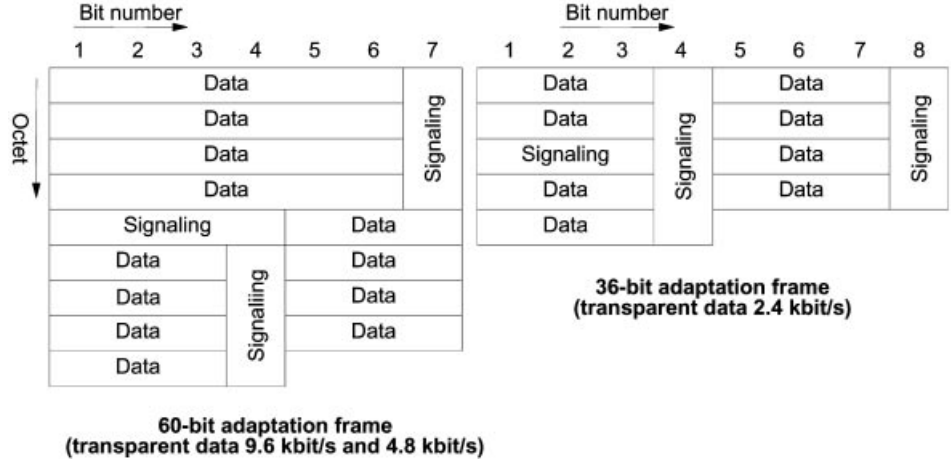

Figure 9.5: Modified V.110 adaptation frame for the RA1′ stage

In the case of user data rates of 4.8 kbit/s or 9.6 kbit/s, adaptation to the channel data rate is already complete. Only for user data rates of less than 4.8 kbit/s do additional parts of the multiple user bits need to be removed, which results in a modified V.110 frame of 36 bits. Thus the user data rates of less than 4.8 kbit/s are adapted to a channel data rate of 3.6 kbit/s. The user data bits of a 2.4 kbit/s signal are then not transmitted twice anymore, or the 600 bit/s user data signals are only written four times into the frames of the RA1′ stage. This is, however, only true for the transparent bearer services. For the nontransparent bearer services, the modified 60-bit V.110 frame is used completely for the transmission of the 60 data bits of an RLP PDU. The required signaling bits are multiplexed with user data into the RLP frame through the Layer 2 Relay protocol L2R.

The modem used for communication over the PSTN resides in the IWF of the MSC, since data is transmitted from here on in digital form within the PLMN. For congestion and flow

control and other functions at the modem interface, the interface signals must therefore be carried from the modem through the PLMN to the mobile station. For this purpose, signaling bits are reserved in the frames of the bit rate adaptation function, which represent these signals and thus give the MS direct modem control. The connection of such a bearer service is therefore transparent not only for user data, but also for out-of-band signaling of the (serial) modem interface in the IWF.

9.5 Asynchronous Data Services

Asynchronous data transmission based on the V. and X. series interfaces is widespread in fixed networks. In order to support such "non-GSM" interfaces, the mobile station can include a *Terminal Adapter* (TA) over which standard terminals with a V. or X. interface (e.g. V.24) can be connected. Such an adaptation unit can also be integrated into the mobile station (MT2, Figure 9.1).

Flow control between TA and IWF can be supported in different ways, just as in ISDN:

- *No Flow Control*: It is handled end-to-end in higher protocol layers (e.g. transport layer)
- *Inband Flow Control* with X-ON/X-OFF protocol
- *Out-of-Band Flow Control* according to V.110 through interface leads 105 and 106.

9.5.1 Transparent Transmission in the Mobile Network

In the case of transparent transmission, data is transmitted with pure Layer 1 functionality. Besides error protection at the air interface, only bit rate adaptations are performed.

User data is adapted to the traffic channel at the air interface according to the data rate and protected with forward error-correcting codes (FECs) against transmission errors. As an example, Figure 9.6 shows the protocol model for transparent asynchronous data transmission over an MT1 with an S interface. Data is first converted in the TE1 or TA into a synchronous data stream by bit rate adaptation (stage RA0). In further stages, data rates are adapted with an MT1 to the standard ISDN (RA1, RA2), and then converted in MT1 over RA2, RA1, and RA1′ to the channel bit rate at the air interface. Provided with an FEC, the data is transmitted and then converted again in the BSS by the inverse operations of bit rate adaptation to 64 kbit/s at the MSC interface. But much more frequently than an MT1 with an (internal) S interface, mobile stations realize a pure R interface without internal conversion to the full ISDN rate in the RA2 stage. This avoids the bit rate adaptation step RA2 and thus the conversion to the intermediate data rate in the RA1 stage. The signal is converted immediately after the asynchronous–synchronous conversion in the RA0 stage from the user data rate to the channel data rate (stage RA1′).

A variation without terminal adapter is shown schematically in Figure 9.7. Here the complete interface functionality, *Interface Circuit* (I/Fcct), for a serial V. interface is integrated with the required adaptation units. The data signals D are converted into a synchronous signal in MT2 (RA0), packed into a modified V.100 frame together with signaling information S from the V.-interface, and adapted to the channel data rate (RA1′). After FEC, the data signals are transmitted over the air interface and finally

converted for further transmission to the data rate of an ISDN B channel after decoding and potential error correction in the BSS (RA2).

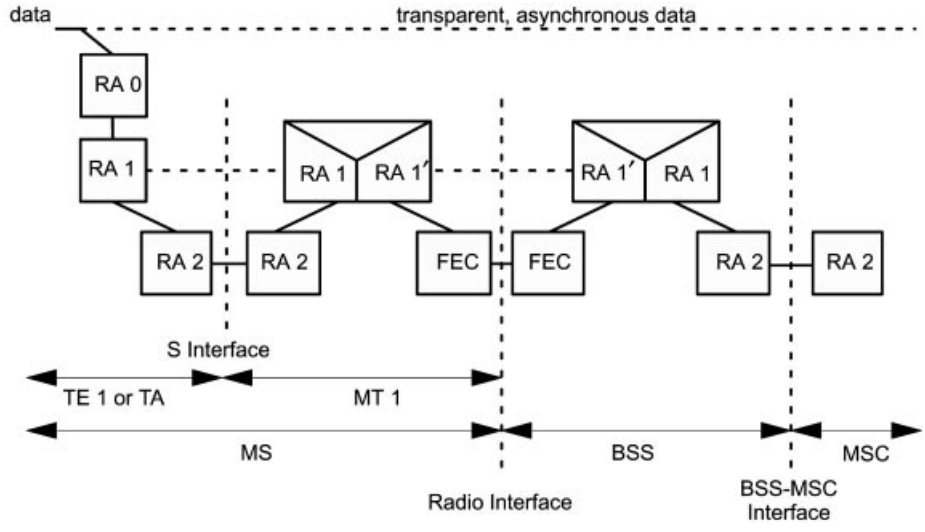

Figure 9.6: Transparent transmission of asynchronous data in GSM

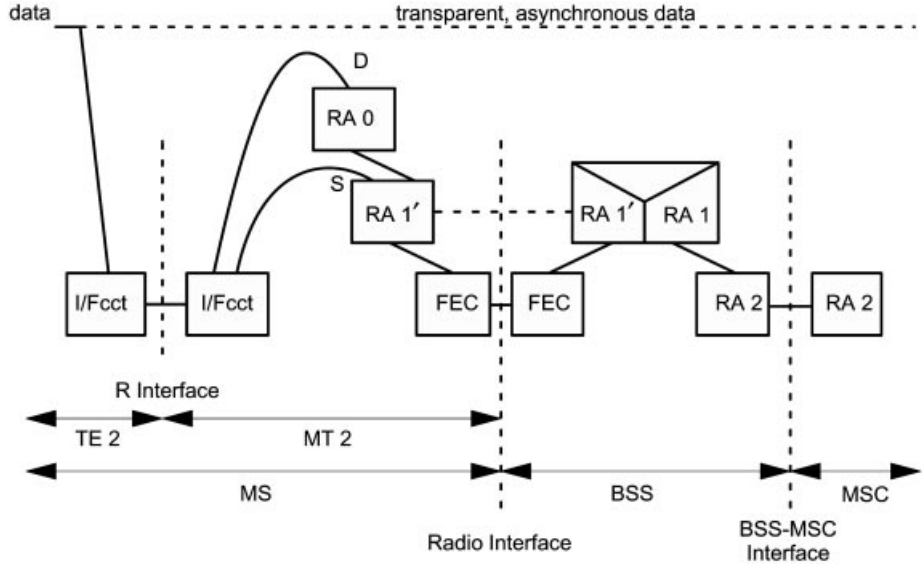

Figure 9.7: Transparent transmission of asynchronous data across the R interface

Figure 9.8 shows a complete scenario with all appropriate network transitions for a transparent bearer service with modem in the interworking function for the conversion of the digital data signals into an analog voiceband signal. A mobile data terminal uses the transparent bearer service of a GSM PLMN over an R interface (S interface is also

possible). The data is circuit-switched to the IWF in the MSC. To communicate with a modem in the fixed network, the IWF activates an appropriate modem function and converts the digital data signals into an analog voiceband signal. The IWF digitizes this voiceband signal again and passes the data on in PCM-coded format through the GMSC. After the network transition, the data signal is finally transmitted to the modem of the communication partner. This modem can be within a terminal in the PSTN or belong to an ISDN terminal. Before being transmitted in the PSTN, the PCM-coded signal is again converted into an analog voiceband signal. In ISDN the signal is transmitted as a PCM-coded signal of category *3.1 kHz audio*; a repeated conversion is not necessary. An ISDN subscriber needs an adaptation unit TA′ for the conversion of the digital voiceband signal into an analog signal, which can then be processed further with a modem and passed on to the data terminal.

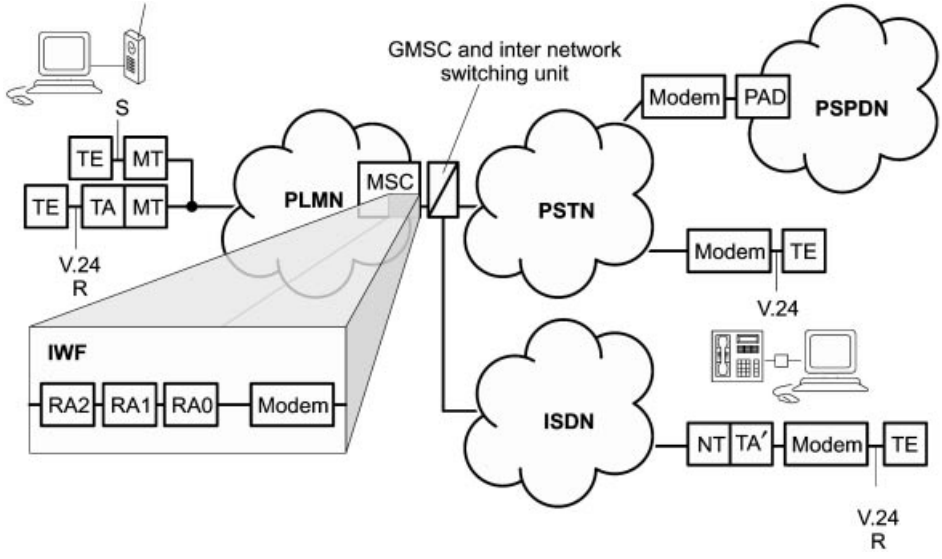

Figure 9.8: Principle of transparent asynchronous data transfer (variant with modem)

Another variant consists of a circuit-switched modem connection to a packet-switched network access node, as is possible from fixed connection ports. In these access nodes, the asynchronous modem signals are combined into packets in a *Packet-Assembler/Disassembler* (PAD) module and then transmitted through the packet-switched network. This variant of packet network access has the disadvantage that one has to switch through to the PAD over a long path, especially in the case of international roaming, since usually the nearest PAD is not the one allowed to the subscriber for access to packet networks.

It is also possible to connect to standard ISDN terminals without an analog modem based on the digital data transmission capability of ISDN. For this purpose, the transmission mode *Unrestricted Digital* has been defined. In this case, there is only a bit rate adaptation according to V.110 (Figure 9.9). The data arrives at the MSC from the BSS in V.110 frames on an ISDN channel with 64 kbit/s and transparently is passed on to the ISDN using a B channel again in V.110 frames. The otherwise necessary modems are entirely un-

necessary in the case of *Unrestricted Digital* connections. However, an ISDN subscriber can connect a terminal through an analog modem by using an adaptation unit TA′ which converts the unrestricted digital signal into a voiceband signal according to one of the V. standards.

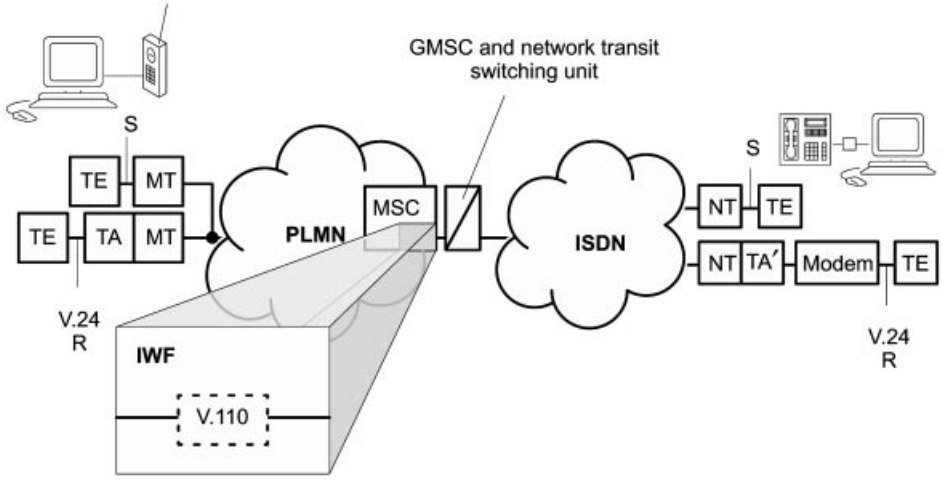

Figure 9.9: Transparent data transfer to an ISDN (unrestricted digital)

The quality of transparent data services in GSM varies with the radio field conditions. This is illustrated by examples of comparative field measurements of the transparent data service BS26 with 9.6 kbit/s data rate, comparing a moving and a standing mobile station. Figure 9.10 shows the weighted distribution of the bit errors of these two cases for a block length of 1024 bits. The weighted distribution indicates the frequency with which m bit errors occur in a block of length n bits (here $n = 1024$):

$$P(m, n) = \sum_{i=m}^{\infty} P(i \text{ errors in } n\text{-bit-block})$$

The distribution shown in Figure 9.10 represents measurements of the error statistics of BS26 which were performed in 1994 in a suburban area for moving and standing mobile stations [58]. Notice that the error frequency for the standing mobile station (“stationaer”) is clearly lower than for the moving mobile station (“mobil”). This result is obtained by averaging measurements over several locations and measurement tours. The resulting mean shows, for the moving station, a sometimes heavily varying channel due to the unavoidable fading phenomena, which frequently cause bursts with high bit error ratios, again resulting in an aggregate higher mean bit error ratio and thus also a higher packet error ratio $P(1, 1024)$.

9.5.2 Nontransparent Data Transmission

In contrast to the transparent transmission mode, the nontransparent mode in GSM data

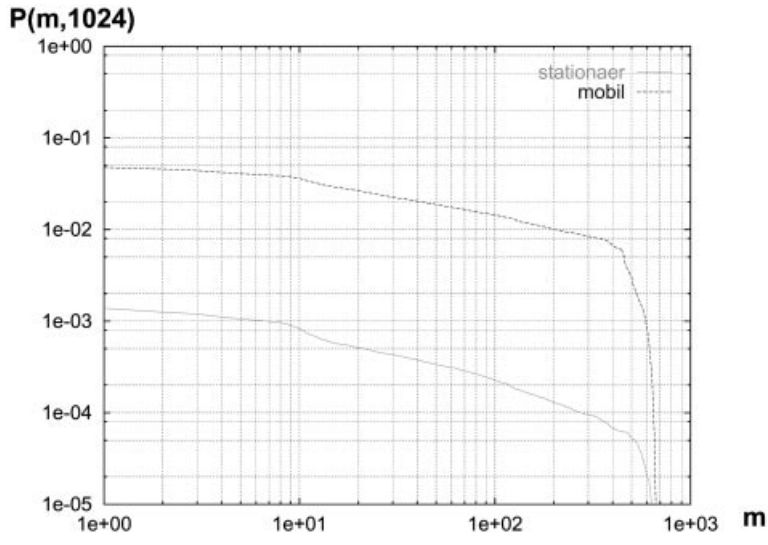

Figure 9.10: Weighted distribution of a transparent GSM bearer service (BS26)

services protects the user data within the PLMN through a Layer 2 protocol, the *Radio Link Protocol* (RLP), in addition to the FEC procedures (convolutional coding, interleaving). This protection protocol further reduces the residual bit error ratio of data transmission in the mobile network. However, the automatic repeat requests (ARQ) of the RLP introduces additional transmission delays on the data path, and the effective user data throughput is reduced (protocol overhead).

User data between MT and MSC/IWF is protected on Layer 2 by the RLP. Two kinds of transmission errors are corrected this way: first, those caused by radio interference and remaining uncorrected by FEC, and second, those caused through the interruptions of handover. For signaling on the FACCH, time slots are "stolen" from the data traffic channel, which can cause data losses. The RLP protects user data against such losses, too.

For nontransparent data transfer with RLP, an additional sublayer in Layer 2 is required, the *Layer 2 Relay* (L2R) protocol. This relay protocol maps user data and status information of the IWF user–modem interface onto the information frames of the RLP. Depending on the kind of user data (character- or bit-oriented), one of two variations of the L2R is used: *Layer 2 Relay Bit Oriented Protocol* (L2RBOP), or *Layer 2 Relay Character Oriented Protocol* (L2RCOP). An L2R PDU is handed to RLP as a service data unit (SDU) and inserted into the RLP frame as a data field. The first octet of an L2R PDU always contains control information, like the status of the signaling lines of the serial interface. Beyond that, the L2R PDU can contain an arbitrary number of such status octets. They are always inserted into the user data stream when the state of the interface changes (e.g. hardware flow control). Thus, of the 200 user data bits (Figure 7.10), only a maximum of 192 can be used for payload. However, since the signaling information is already contained in the L2R PDU, these bits must not be considered for bit rate adaptation. Thus the modified 60-bit V.110 frame (Figure 9.5) can be completely occupied with data bits, and the full channel rate of maximum 12 kbit/s can be used for the transmission

of RLP data. Next come a set of four modified V.110 frames which carry a complete RLP frame. Considering the protocol overhead of RLP (16.7%) and the minimum overhead of an L2R PDU (0.5%), one obtains a usable subscriber data rate of up to 9.95 kbit/s.

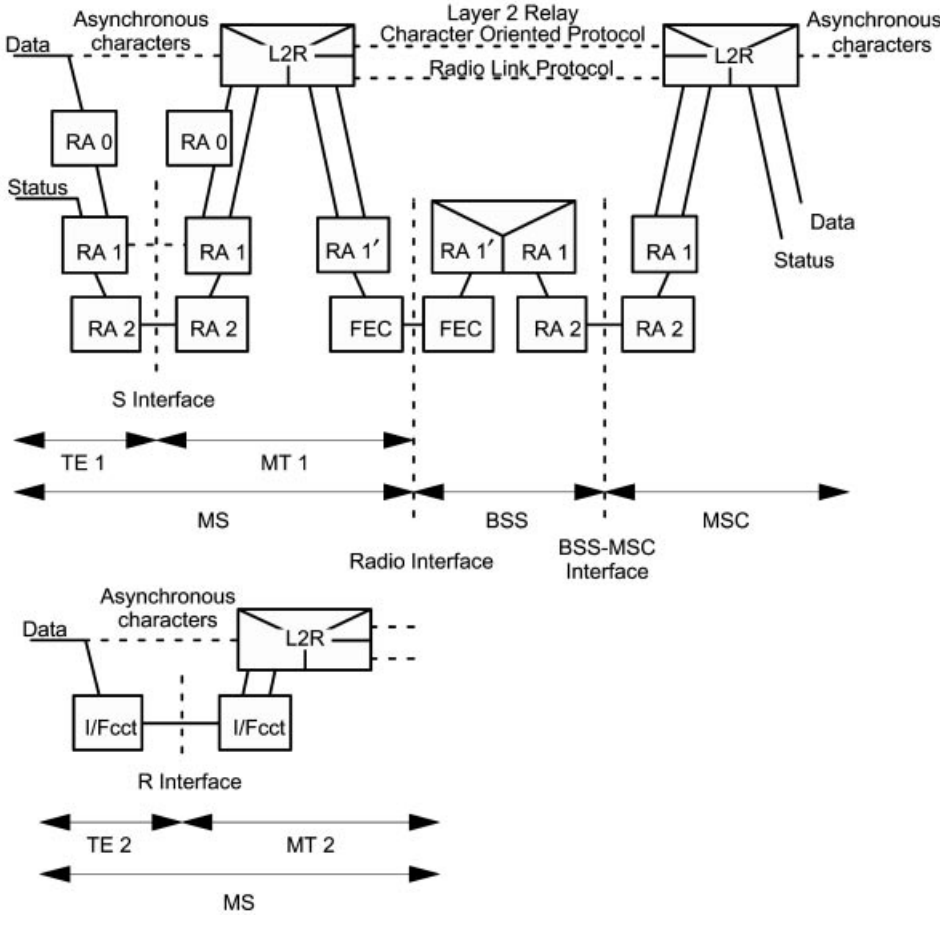

Figure 9.11: Nontransparent data transmission in GSM

The protocol model for asynchronous nontransparent character-oriented data transmission over the S interface in GSM is now presented as an example (Figure 9.11). Data is transmitted by L2R protocol and *L2R Character Oriented Protocol* (L2RCOP) and RLP from the MT1 termination to the MSC. In between, as in the transparent case, are FEC, RA1, RA1′, and RA2. The RLP frames are transported in a synchronous mode. Only the user data stream at the user interface is asynchronous, and in the case of the model in Figure 9.11 the user data stream must be converted for the S interface into a synchronous data stream (RA0). In the case of a terminal with V. interface and an MT2 termination (reference point R), this bit rate adaptation at the S interface would be avoided. The asynchronous data is then directly accepted at the serial interface by the L2R (I/Fcct); potential start/stop bits are removed, and data are combined into L2R PDUs.

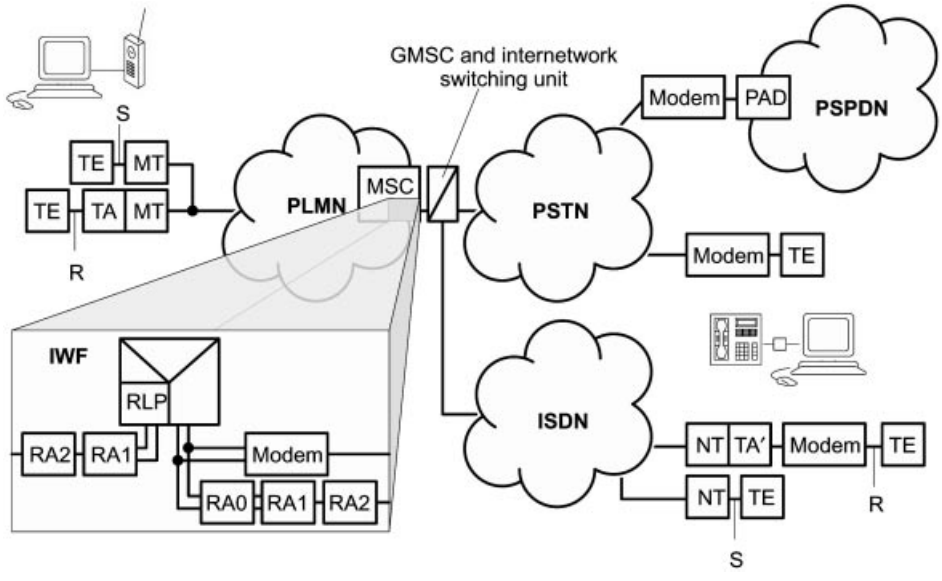

Figure 9.12: Principle of nontransparent data transfer

A complete scenario for network transition with nontransparent asynchronous GSM data services is shown in Figure 9.12. The RLP is terminated in the MSC/IWF, and user data are converted again into an asynchronous data stream by the associated L2R. For the nontransparent case too, the IWF offers both variants for the network transition: first, using modems and, second, *Unrestricted Digital*. In the case of a network transition to the PSTN or to a *3.1 kHz audio* connection into ISDN, the respective modem function is inserted and passes PCM-coded data to the GMSC (not shown in Figure 9.12), which directs them into the fixed networks. Using several bit rate adaptation steps (RA0 – RA1 – RA2, Figure 9.12), the user data can also be converted in the IWF into a synchronous *Unrestricted Digital* signal, which is then carried transparently over an ISDN B channel with 64 kbit/s.

9.5.3 PAD Access to Public Packet-Switched Data Networks

9.5.3.1 Asynchronous Connection to PSPDN PADs

As shown in Figures 9.8 and 9.12, access to *Packet Switched Public Data Networks* (PSPDNs), e.g. Accunet in the USA or Datex-P in Germany, is already possible using the asynchronous services of GSM. This requires a *Packet Assembler/Disassembler* (PAD) in the PSPDN, which packages the asynchronous data on the modem path into X.25 packets and also performs the reverse operation of unpacking. PAD access uses the protocols X.3, X.28, and X.29 (Triple X Profile). Just as from the fixed network, the mobile subscriber dials the extension of a PAD for access to the service of the PSPDN, provided packet network access is allowed. In this way, the subscriber has the same kind of access to the packet network as a subscriber from the fixed network, aside from the longer transmis-

sion delays and the higher bit error ratios. It is therefore recommended to transmit data for PAD access across the air interface in the PLMN in nontransparent mode with RLP [21].

9.5.3.2 Dedicated PAD Access in GSM

However, direct access to packet data networks through the asynchronous GSM data services has disadvantages:

- One needs another subscription to a packet data network operator besides to GSM.
- Independent of the current mobile subscriber's location, a circuit-switched connection to a PAD of a packet service provider is needed. Sometimes the packet network access is only allowed to specific PADs. This is a particular disadvantage if the mobile subscriber is currently in a foreign GSM network and incurs fees for international lines.

Therefore, GSM has defined another PSPDN access without these disadvantages: *Dedicated PAD Access* (Figure 9.13). The services are defined as *Bearer Services* BS41 through BS46 (Table 4.2). With this kind of PSPDN access from a PLMN, each PLMN has at least one PAD that is responsible for the packaging/unpackaging of the X.25 packets of the respective mobile subscriber.

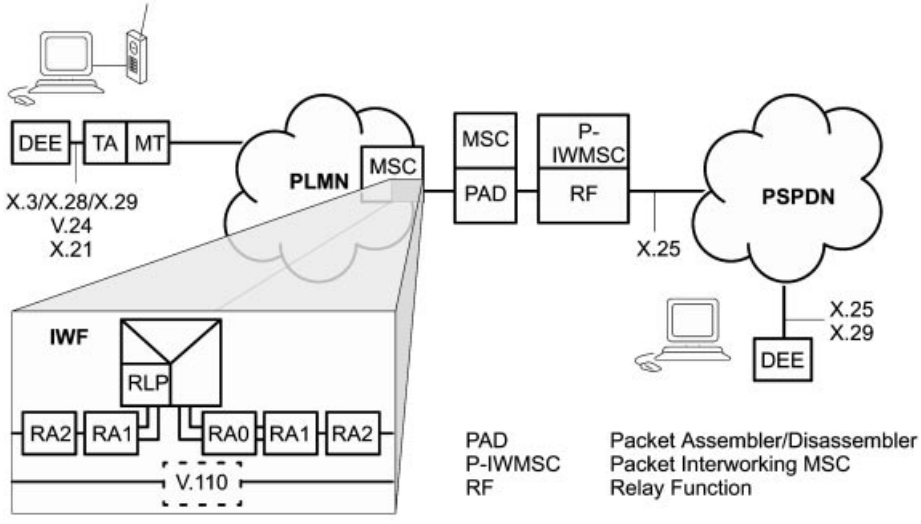

Figure 9.13: Dedicated PAD access through asynchronous GSM data services

For this purpose, a PAD is activated as an additional resource in the IWF or in an especially reserved MSC (Figure 9.13). This PAD can be reached in asynchronous mode again over transparent or nontransparent PLMN connections. However, with this solution the connection to the PAD is as short as possible, since a PAD is already reached in the nearest IWF, and international lines are never occupied for PAD access. Packetization of user data is already performed within the mobile network rather than in the remote PAD of a packet network operator, hence there is no need for a separate subscriber agreement with this

packet network operator. The dedicated PAD of the current PLMN is now responsible for the packaging/unpackaging of the asynchronous data into/from X.25 packets, which are then passed on through a specific interworking MSC (P-IWMSC) to a public PSPDN (e.g. Transpac in France or Sprint's Telenet in the USA) [21].

The dedicated PAD has a uniform profile in all GSM networks; it is reached in each network with the same access procedure. Even in foreign PLMNs, a mobile station therefore gets the earliest and lowest cost access to the packet data network. Charging occurs to the account of the GSM extension (MSISDN) of the mobile subscriber; a separate *Network User Identification* (NUI) for the PSPDN is not necessary. However, only outgoing packet connections are possible.

9.6 Synchronous Data Services

9.6.1 Overview

Synchronous data services allow access to synchronous modems in the PSTN or ISDN as well as to circuit-switched data networks. Such access is not very significant; however, synchronous data services are defined in GSM. The essential differences to the asynchronous data transmission procedures are in bit rate adaptation and modems. For a synchronous data service, no RA0 bit rate adaptation is needed (conversion from asynchronous to synchronous), since data is already in synchronous format. Instead, special synchronous modems are required in the IWF. Synchronous data services can only be offered in transparent mode, with the exception of access to X.25 packet networks, which are a significant application of synchronous data service in GSM.

9.6.2 Synchronous X.25 Packet Data Network Access

The protocol model shown in Figure 9.14 is the model for synchronous data transmission in nontransparent mode with the packet data access protocol according to the ITU-T standard X.25. Because of the nontransparent transmission procedure, the X.25 *Link Access Procedure B* (LAPB) must be terminated in the MT as well as in the IWF. Since LAPB of the X.25 protocol stack operates in a bit-oriented mode, the *Layer 2 Relay Bit Oriented Protocol* (L2RBOP) is required.

9.6.2.1 Basic Packet Mode

Two variants of PSPDN access can be realized with the protocol model in Figure 9.14:

- PSPDN access according to ITU-T X.32
- Access to PSPDN packet handlers according to ITU-T X.31 Case A (basic packet mode).

PSPDN access according to X.32 has not met much acceptance with the applications; the X.31 procedure is used more often.

PSPDN access according to X.32 is the simpler variant. The X.25 packets can be trans-

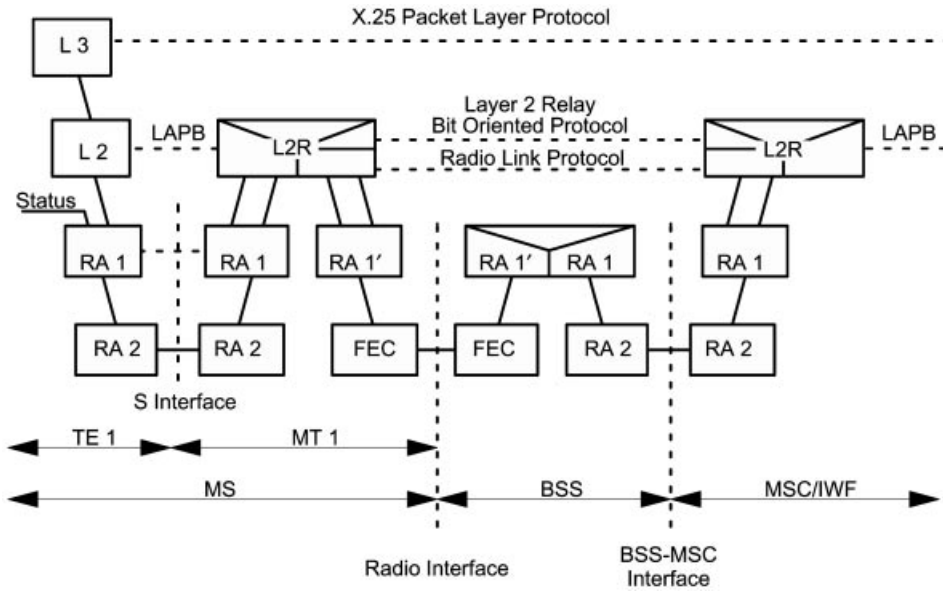

Figure 9.14: X.25 access at the ISDN S interface

ferred directly over a synchronous modem in the IWF to the PSPDN. This does not necessarily require nontransparent transmission in GSM, but it helps because of the lower bit error ratio. In case of the nontransparent transmission, the LAPB protocol has to be terminated in MT and IWF (see above). The subscriber needs a *Network User Identification* (NUI). Incoming and outgoing packet connections are possible, but again there is the problem of needing circuit-switched connections to the home PLMN, just like in case of international roaming.

The access procedure according to ITU-T standard X.31 Case A (basic packet mode) is the more favored variant of this group of services. The X.25 packets of the mobile subscriber are passed from the IWF to the packet handler of the ISDN. Since speed adaptation with the X.31 procedure is performed in the ISDN B channel by flag stuffing, the protocol has to be terminated in the IWF, which means that the nontransparent mode of GSM can be used. In this case too, there are connections possible only through the packet handler of the home network.

9.6.2.2 Dedicated Packet Mode

As in the case of asynchronous PAD access to packet data networks (see Section 9.5.3), the synchronous case of the X.25 access protocol also offers an alternative, which allows the most immediate transition to the PSPDN, even if the mobile station is in a foreign network (international roaming). For this purpose, a dedicated mode is also defined with each PLMN having its own packet handler.

Figure 9.15 shows the principle of this *Dedicated Packet Mode*. It essentially includes the functions of the basic packet mode, with the difference that the packet handler is integrated

into the GSM network. Data is transmitted in the PLMN in nontransparent and synchronous mode. Access to the packet handler is the same in all PLMNs, and is also available to foreign mobile subscribers. Since packet data networks do not know roaming, only outgoing data calls are possible.

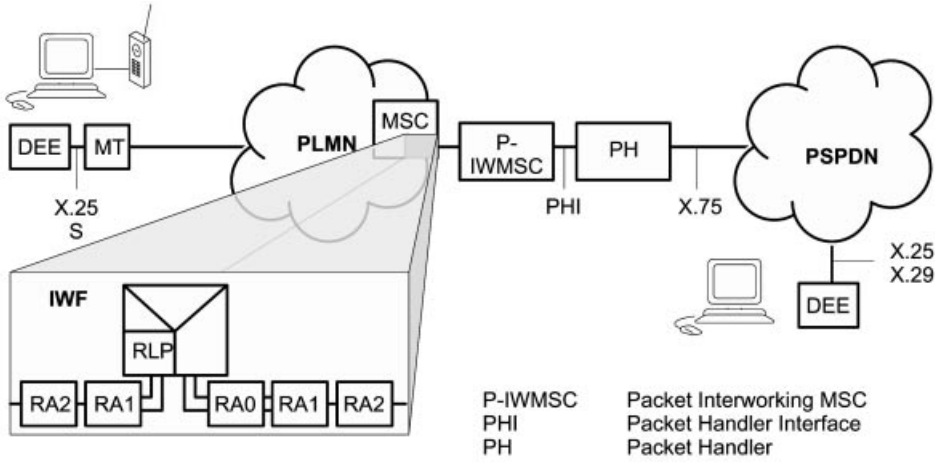

Figure 9.15: Dedicated packet mode with packet handler in GSM

9.7 Teleservices: Fax

In the following, we briefly explain the realization of the GSM fax service.

The GSM standard considers the connection of a regular Group 3 fax terminal with its two-wire interface to an appropriately equipped mobile station as the standard configuration of a mobile fax application. The GSM fax service is supposed to enable this configuration to conduct fax transmissions with standard Group 3 fax terminals over mobile connections. This requires mapping the fax protocol of the analog two-wire interface onto the digital GSM transmission, because the fax procedure defines a complete protocol stack with its own modulation, coding, user data compression, inband signaling, etc. Therefore the *Fax Adapter* (FA) has been defined for this mapping. The principle is summarized in Figure 9.16. The fax adapter of the mobile station converts the fax protocol of a standard Group 3 fax terminal on an analog two-wire line into a GSM internal fax adapter protocol. The PDUs of the adapter protocol are transmitted over the MT with the GSM data services to the fax adapter in the IWF of the MSC, and there they are again converted into the T.30 protocol on the analog line, or transmitted to the ISDN in PCM-coded format, where again a terminal adapter allows connection of a Group 3 fax terminal (not shown in Figure 9.16).

A more compact version of a mobile fax terminal is possible, if the fax adapter and the Group 3 fax terminal are integrated into a compact terminal (GSM fax, Figure 9.16). Such equipment can be connected at reference point R to an MT2. It delivers a digital signal directly. With this integration, the analog two-wire line interface becomes superfluous.

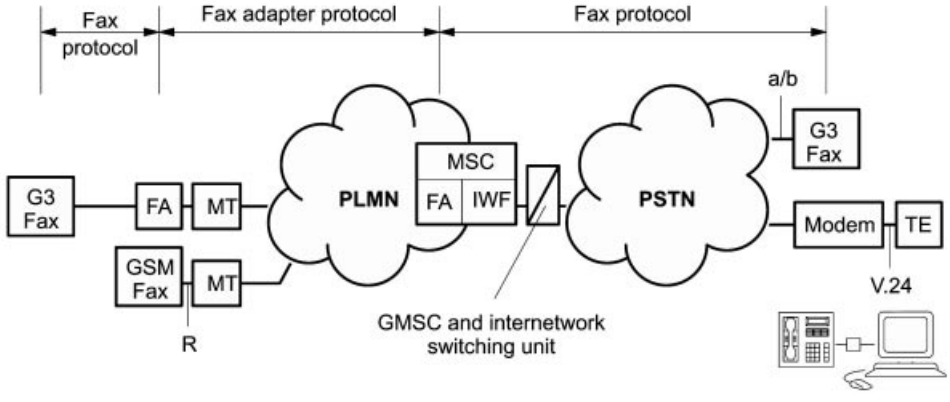

Figure 9.16: Fax adapter in GSM

Hence all the analog functions can be omitted such as demodulation and digitalization of the Group 3 fax signals, i.e. the GSM fax needs no analog components such as a modem building block. Therefore, however, the integrated GSM fax must implement the fax adapter protocol and terminate it at reference point R in order to guarantee correct control of the analog fax components in the fax adapter of the IWF.

A complete fax scenario with the required analog components in the fax adapter is illustrated in Figure 9.17. The FA needs several function blocks in the MS as well as in the IWF for the conversion of the fax protocol of the analog a/b interface to the digital transmission procedure in the PLMN. Group 3 fax equipment according to ITU-T standard

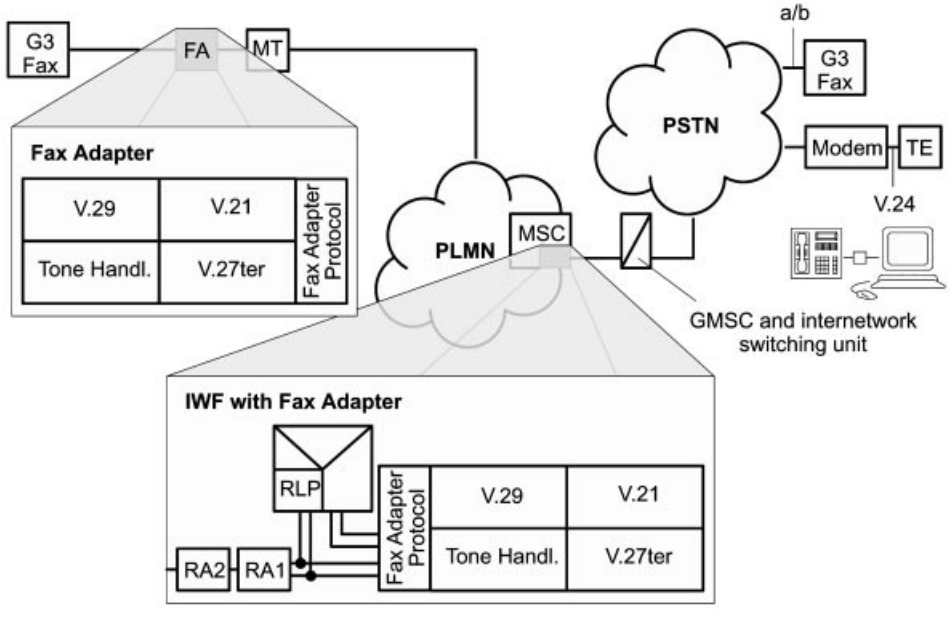

Figure 9.17: Overview of GSM procedure for fax service

T.30 employs three modem building blocks, all three operating in half-duplex mode. A V.21 modem (300 bit/s) is used for the signaling phase to set up a fax connection, whereas the information transfer phase uses a V.27ter modem (4.8 kbit/s or 2.4 kbit/s) or a V.29 modem (9.6 kbit/s, 4.8 kbit/s, and 2.4 kbit/s). For conversion from analog signaling tones into messages of the fax adapter protocol, the fax adapter needs an additional *Tone Handler*. It is used to transfer the (re-)digitized fax signals over either a transparent or a nontransparent GSM bearer service to the IWF in the MSC. The fax adapter protocol provides a complete mapping of the T.30 protocol, such that the IWF is able to handle the complete fax protocol with the partner entity. From the view of the partner entity, the complete GSM connection consisting of fax adapter, mobile station, and IWF represents a physical connection with a mobile Group 3 fax terminal.

Fax service poses special demands on the service quality of the data channel which the GSM network provides for this teleservice. In particular, propagation delays must remain below a maximum threshold, since timers of the T.30 protocol expire otherwise. This is especially critical if RLP is used to reduce transmission errors, which introduces additional delays into the data path through its ARQ procedure.

Two fax services are specified: the transparent procedure and the nontransparent procedure, depending on the kind of bearer service used. The resulting protocol models are shown in Figures 9.18 and 9.19, respectively. It is evident that in both cases the fax protocol is superimposed onto the respective bearer service. The transparent fax procedure

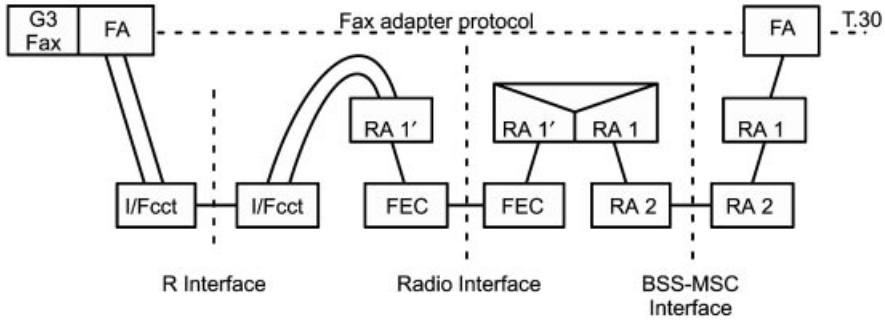

Figure 9.18: Transparent fax procedure in GSM Radio Interface BSS-MSC

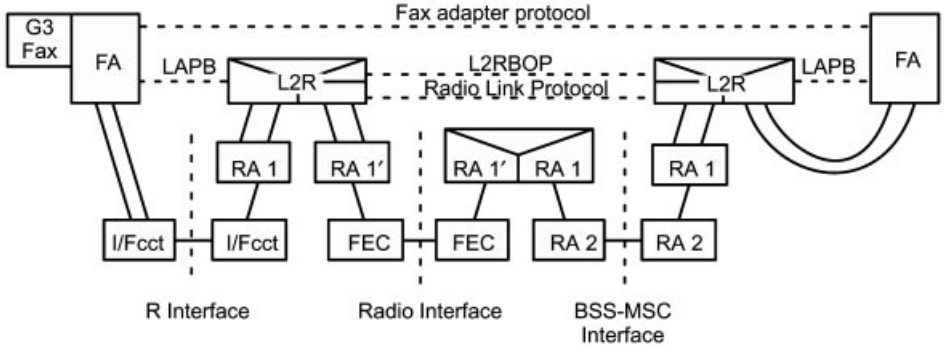

Figure 9.19: Nontransparent fax procedure in GSM

is easier to realize, although with its transparent bearer service, it incurs a correspondingly variable fax quality.

In contrast, the nontransparent fax procedure based on RLP is very well protected against transmission errors, and it delivers very acceptable quality of the transmitted documents over a wide range of distances. However, due to the varying transmission conditions, there are also variable delays of the RLP which lead to intermediate buffering of fax signals in the FA and which, in the worst case, can cause the breakdown of the fax transfer [16].

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

10 Aspects of Network Operation

For the efficient and successful operation of a modern communication network such as a GSM PLMN, a comprehensive *Network Management* (NM) is mandatory. Network management encompasses all functions and activities which control, monitor, and record usage and resource performance of a telecommunication network, with the objective of offering the subscribers telecommunication services of a certain objective level of quality. Various aspects of quality are either defined and prescribed in standards or laid down in operator-specific definitions. Special attention has to be paid to the gap between (mostly simple) measurable technical performance data of the network and the quality of service experienced (subjectively) by the subscriber. Modern network management systems should therefore also include (automated) capabilities to accept reports and complaints from subscribers and convert them into measures to be taken by network management (e.g. trouble ticketing systems).

10.1 Objectives of GSM Network Management

Along with the communication network which realizes the services with its functional units (MS, BSS, MSC, HLR, VLR), one needs to operate a corresponding network management system for support and administration. This NM system is responsible for operation and maintenance of the functional PLMN units and the collection of operational data. The operational data comprise all the measurement data which characterize performance, load, reliability, and usage of the network elements, including times of usage by individual subscribers, which are the basis for calculation of connection fees (billing). Furthermore, in GSM systems in particular, the techniques supporting security must have counterparts in security management functions of network management. This security management is based on two registers: the *Authentication Centre* (AUC) provides key management for authentication and encryption and the *Equipment Identity Register* (EIR) provides barring of service access for individual equipment, "blacklisting." To summarize, for all functions of the telecommunication network and its individual functional units (network elements), there are corresponding NM functions.

The GSM standard has defined the following overall objectives of network management:

- International operation of network management

- Cost limitation of GSM systems with regard to short-term as well as long-term aspects
- Achievement of service quality which at least matches the competing analog mobile radio systems

The international operation of a GSM system includes among others the interoperability with other GSM networks (including different countries) and with ISDN networks, as well as the information exchange among network operators (billing, statistical data, subscriber complaints, invalid IMEI etc.). These NM functions are in large part necessary for network operation allowing international roaming of subscribers, and therefore they must be standardized. Their implementation is mandatory.

The costs of a telecommunication system consist of invested capital and operational costs. The investments comprise the cost of the installation of the network and of the network management, as well as development and licensing costs. The periodically incurred costs include operation, maintenance, and administration as well as interest, amortization, and taxes. Lost revenues due to failing equipment or partial or complete network failure must be included in the periodically incurred costs, whereas consequential losses due to cases of failure, e.g. because of lost customers, cannot be estimated and included. Therefore the reliability and maintainability of the network equipment is of course of immense importance and heavily impacts costs. The installation of an NM system on one hand increases the need for investment capital for the infrastructure as well as for spare capacities in the network. On the other hand, these costs for a standardized comprehensive NM system must be compared with the expenses for administration, operation, and maintenance of network elements with manufacturer-proprietary management, or the costs which arise from not recognizing and repairing network failures early enough. Therefore it has to be the objective of a cost-efficient NM system to define and implement uniform vendor-independent network management concepts and protocols for all network elements, and also to guarantee interoperability of network components from different manufacturers through uniform interfaces within the network.

The quality of service to be achieved can be characterized with technical criteria like speech quality, bit error ratio, network capacity, blocking probability, call disconnection rates, supply probability, and availability, and it can also be characterized with nontechnical criteria like ease of operation and comfort of subscriber access or even hot-line and support services.

Considering these objectives, the following functional areas for network management systems can be identified:

- Administrative and business area (subscribers, terminal equipment, charging, billing, statistics)
- Security management
- Operation and performance management
- System version control
- Maintenance

These functions are realized in GSM based on the concept of the *Telecommunication Management Network* (TMN). In general, they are summarized with the acronym

FCAPS – *Fault*, *Configuration*, *Accounting*, *Performance*, and *Security Management* (Figure 10.1).

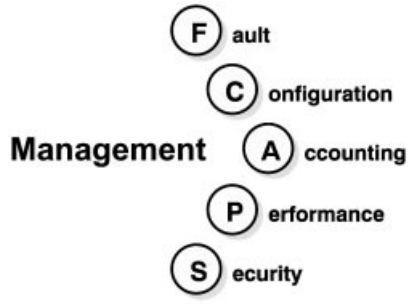

Figure 10.1: Functional areas of TMN systems

Fault management includes functions like failure recognition, failure diagnosis, alarm administration and filtering as well as capabilities for the identification of causes of failures or alarms and keeping of failure logs. Configuration management administers network configurations and handles changes, activates/deactivates equipment, and provides tools for the automatic determination of network topology and connectivity. Accounting management deals with the subscribers and is responsible for the establishment and administration of subscriber accounts and service profiles. Periodic billing for the individual subscribers originates here, based on measured usage times and durations; statistics are calculated, in certain circumstances only for network subareas (billing domains). In performance management, one observes, measures, and monitors performance (throughput, failure rates, response times, etc.), and utilization of network components (hardware and software). The objective is on one hand to ensure a good utilization of resources and on the other hand to recognize trends leading to overload and to be able to start countermeasures early enough. Finally, security management provides for a thorough access control, the authentication of subscribers, and an effective encryption of sensitive data.

10.2 Telecommunication Management Network (TMN)

TMN was standardized within ITU-T/ETSI/CEPT almost simultaneously with the pan-European mobile radio system GSM. The guidelines of the M. series of the ITU-T (M.20, M.30) serve as a framework.

TMN defines an open system with standardized interfaces. This standardization enables a platform-independent multivendor environment for management of all components of a telecommunication network. Essentially it realizes the communication of a management system with network elements it administers, which are considered as managed objects. These objects are abstract information models of the physical resources. A manager can send commands to these administered objects over a standardized interface, can request or change parameters, or be informed by the objects about events that occurred (notification). For this purpose, an agent resides in the managed object, which generates the management messages or evaluates the requests from the manager, and maps them onto corresponding

operations or manipulations of the physical resources. This mapping is system specific as well as implementation dependent and hence not standardized. The generalized architecture of a TMN is illustrated in Figure 10.2.

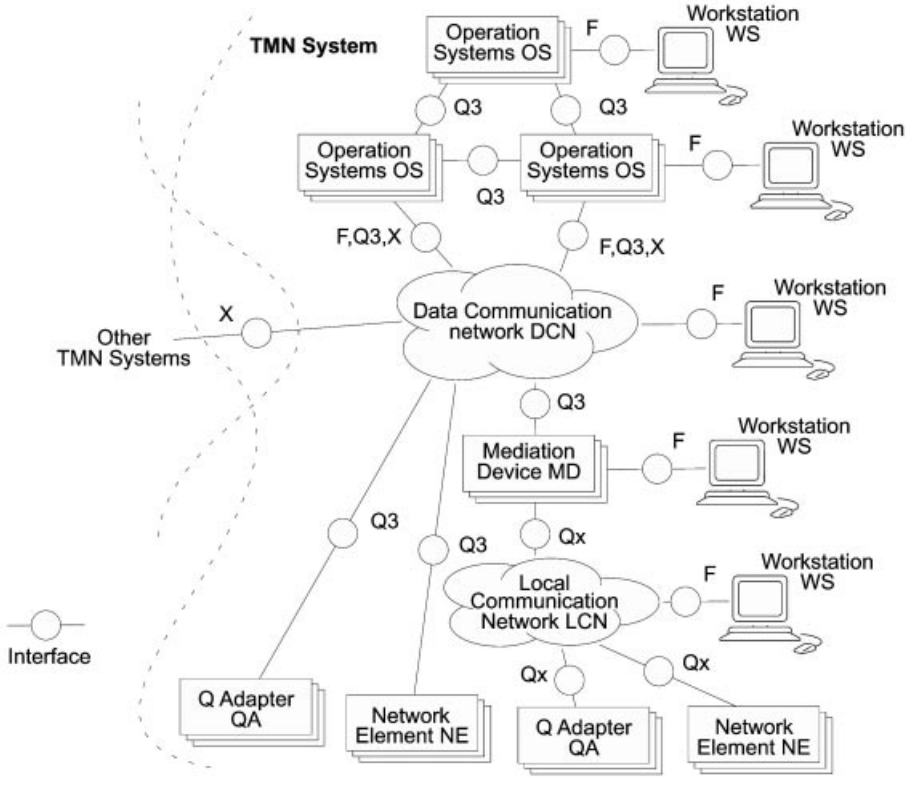

Figure 10.2: TMN architecture (schematically, according to M.3010 [13])

The network management proper is realized in an *Operation System* (OS). The operation systems represent the surveillance and control systems of a TMN system. These systems can communicate with each other directly or form hierarchies. A standardized interface Q3 serves for the communication of the OSs within a TMN, whereas the interconnection of two TMN systems occurs over the X interface (Figure 10.2). The management functionality can also be subdivided into several logical layers according to the OSI hierarchy. For this approach, TMN provides the *Logical Layered Architecture* (LLA) as a framework. The exact numbering and corresponding functionality of each LLA plane were not yet finalized in the standardization process at the time of writing, however, the following planes have been found to be useful (Figure 10.3): *Business Management Layer* (BML), *Service Management Layer* (SML), *Network Management Layer* (NML), and *Element Management Layer* (EML).

The TMN functions of the EML are realized by the network elements NE and contain basic TMN functions such as performance data collection, alarm generation and collection, self diagnosis, address conversion, and protocol conversion. Frequently, the EML is also

known as a *Network EML* (NEML) or a *Subnetwork Management Layer* (SNML) [52]. The NML-TMN functions are normally performed by operation systems and used for the realization of network management applications, which require a network-wide scope. For this purpose, the NML receives aggregate data from the EML and generates a global system view from them. On the SML plane, management activities are performed which concern the subscriber and his or her service profile rather than physical network components. The customer contact is administered in the SML, which includes functions like establishing a subscriber account, initializing supplementary services, and several others. The highest degree of abstraction is reached in the BML, which has the responsibility for the total network operation. The BML supports strategic network planning and the cooperation among network operators [52].

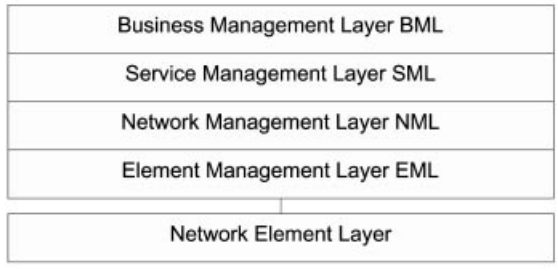

Figure 10.3: Logical layered architecture of a TMN system

For example, an operation system OS could act as a *Basic OS* and be in charge only of a region with a subset of network elements, or it could be a *Network OS* which communicates with all the basic OSs and implements network-wide management functionality. As a *Service OS*, an OS assumes network-wide responsibility for the management of one service, whereas on the BML plane, care is taken of charging, billing, and administration of the whole network and its services.

The individual functional units of the telecommunication network are mapped into *Network Elements* (NEs). These elements are abstract representations of the physical components of the telecommunication network, which is administered by this TMN. The OSs communicate with the network elements over a comprehensive data network, a *Data Communication Network* (DCN). For this purpose, an interface Q3 has been defined, whose protocols comprise all seven layers of the OSI model. However, not every network element must support the full range of Q3 interface capabilities.

For network elements whose TMN interface contains a reduced range of functionality (Qx), a *Mediation Device* (MD) is interposed, which essentially performs the task of protocol conversion between Qx and Q3. A mediator can serve several network elements with incomplete Q3 interfaces, which can be connected to the mediator through a *Local Communication Network* (LCN). The functions of a mediator are difficult to define in general and depend on the respective application, since the range of restrictions of a Qx interface with regard to the Q3 interface is not standardized [23]. Therefore, a mediator could for example realize functions like data storage, filtering, protocol adaptation, or data aggregation and compression.

In spite of ongoing TMN standardization, new network elements and systems without a TMN interface are continuously added and must be integrated. For such cases, the function of the *Q Adapter* (QA) has been defined. In contrast to the mediator MD, which is prefixed to TMN-capable devices with reduced functionality at the Q interface, a QA allows integration of devices which are not TMN capable, and the QA must therefore be tailored for each respective device.

Finally, the operator personnel have access to the TMN system at the F interface through management *Workstations* (WSs) in order to perform management transactions and to check or change parameters. Thus a TMN system gives the network operator at a workstation the capability to supply any network element with configuration data, to receive and analyze failure reports and alarms, or to download locally collected measurement data and usage information. The TMN protocol stack required for this purpose is based on OSI protocols and comprises all seven layers (see also Figure 10.6). The main element of the TMN protocol architecture is the *Common Management Information Service Element* (CMISE) from the OSI system management, which resides in the application layer (OSI Layer 7) [52]. The CMISE consists of a service definition, the *Common Management Information Service* (CMIS), and a protocol definition, the *Common Management Information Protocol* (CMIP). The CMISE defines a uniform message format for requests and notifications between management OS and the managed elements NE or the respective QA.

10.3 TMN Realization in GSM Networks

TMN and GSM were standardized approximately at the same time, so that there was a good opportunity to apply TMN principles and methods in a complete TMN system for network management in GSM from the beginning and from ground up. For this purpose, specific working groups were founded for the five TMN categories (Figure 10.1) as well as for architecture and protocol questions which were supposed to develop as much as possible of the TMN system and its services, while following the top-down methodology [13,14] recommended by the ITU-T. This objective could be pretty much achieved, only that the development methodology was complemented by a bottom-up approach which was rooted in the detailed knowledge about the network components being specified at the same time. The intent was to reach the objective of a complete standard earlier [43,57]. The five TMN categories are essentially realized for all of the GSM system; however, there are some limitations in failure, configuration, and security management. Failure and configuration management are specified only for the BSS; the reasons are that on one hand the databases (HLR, VLR) were assigned to accounting management, and on the other hand standardization efforts were to concentrate on GSM-specific areas. Concentration on GSM-specific areas thereby excluded failure and configuration management for the MSC, which from the management point of view is essentially a standard ISDN switching exchange. For the same reasons, security management is also limited to GSM-specific areas.

The resulting GSM TMN architecture is shown schematically in Figure 10.4. In GSM, the BSC and the MSC have a Q3 interface as network elements to the OS. Besides the BSS

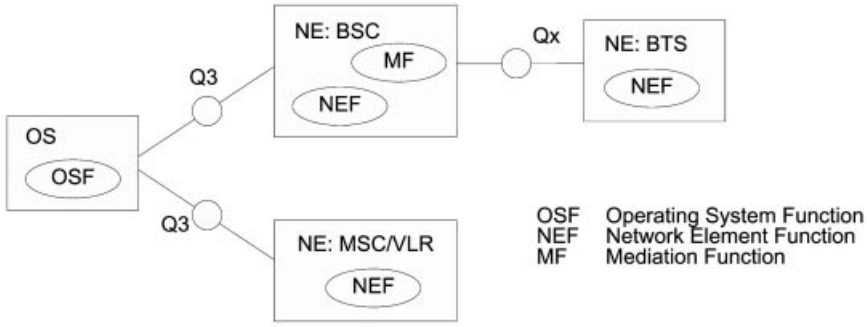

Figure 10.4: A simple TMN architecture of a GSM system (according to [57])

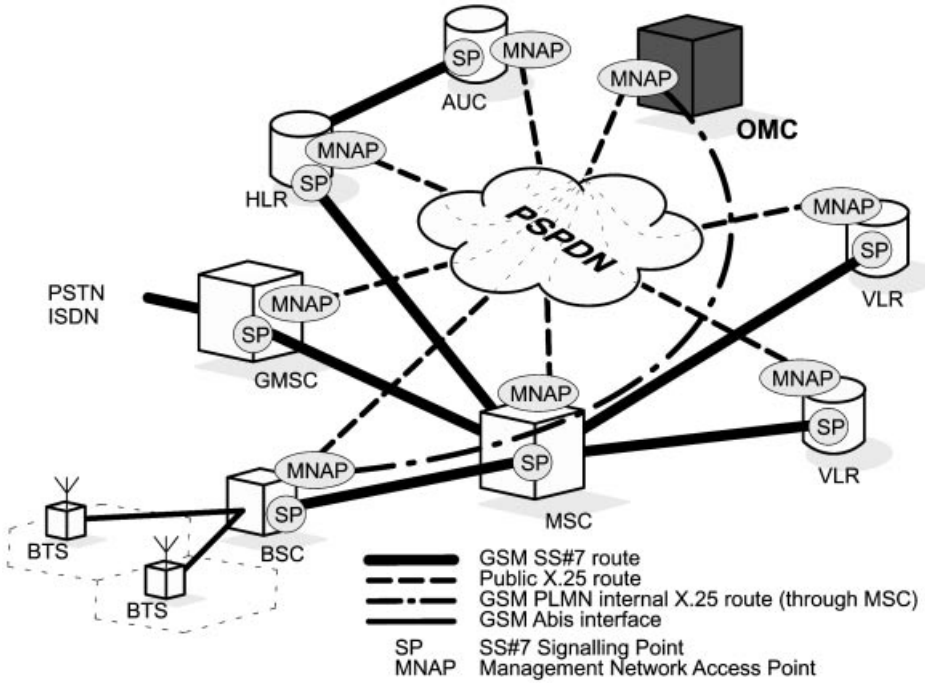

Figure 10.5: Potential signalling interfaces in a GSM TMN

management, the BSC NE always contains a *Mediation Function* (MF) and a Qx interface to the NE supporting the BTS functionality.

An object-oriented information model of the network has been defined for the realization of the GSM TMN services. The model contains more than 100 *Managed Object Classes* (MOCs) with a total of about 500 attributes. This includes the ITU-T standard objects as well as GSM-specific objects, which include the GSM network elements (BSS, HLR, VLR, MSC, AUC, EIR) on one hand, but also represent network and management resources (e.g. for SMS service realization or for file transfer between OS and NE) as

managed objects. These objects usually contain a state space and attributes which can be checked or changed (request) as well as mechanisms for notification, which report the state or attribute changes. In addition, there are commands for creation or deletion of objects, e.g. in the HLR with *create/modify/delete subscriber* or *create/modify/delete MSISDN* or in the EIR with *create/interrogate/delete equipment* [57]. File transfer objects are used especially in the information model dealing with the registers, since it involves movement of large amounts of data.

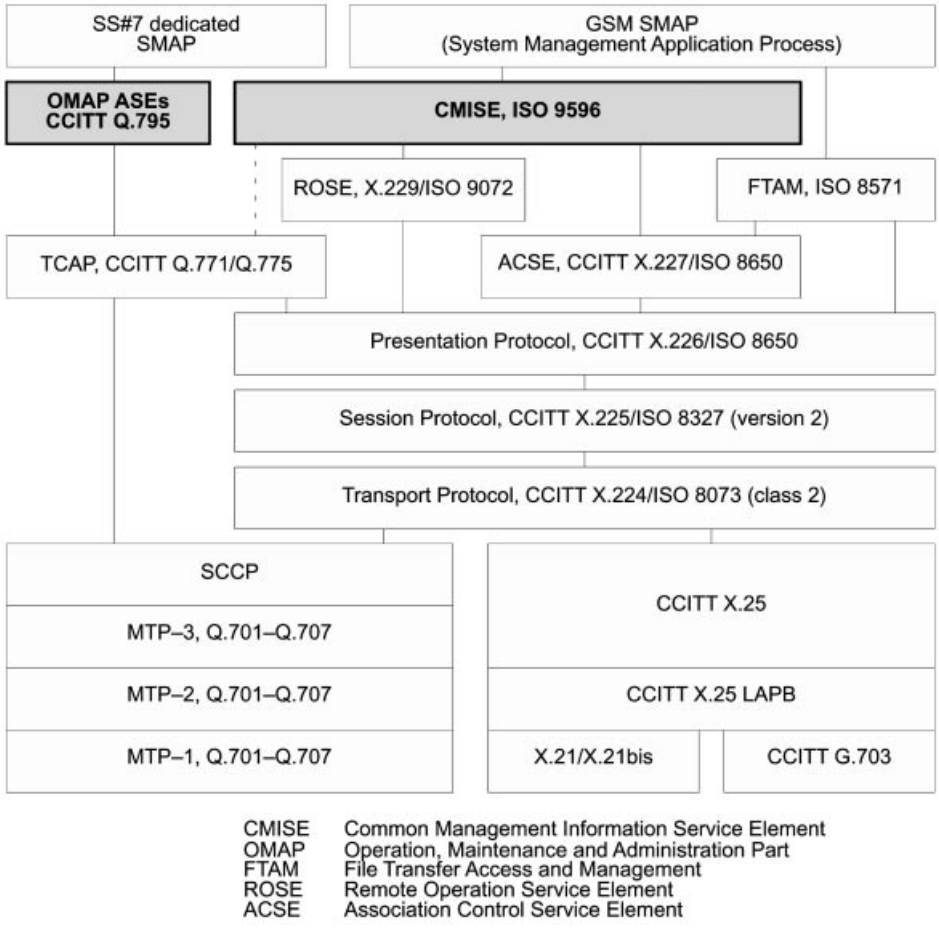

Figure 10.6: GSM network management protocols at the Q3-interface

The TMN communication platform to be used as *Data Communication Network* (DCN) can be either an OSI X.25 packet network or the SS#7 signaling network (MTP and SCCP). Both offer a packet switching service which can be used to transport management messages. Each network element is connected to this management network over a *Management Network Access Point* (MNAP); see Figure 10.5.

If the TMN uses X.25, the DCN can be the public PSPDN or a dedicated packet switching network within the PLMN with the MSC as a packet switching node. In addition, the MSC

can include an interworking function for protocol conversion from an external X.25 link to the SS#7 SCCP, which realizes the connection of the OMC to the PLMN through an external X.25 link. Further transport of management messages is then performed by the SS#7 network internal to the PLMN.

The framework defined for the GSM TMN protocol stack at the Q3 interface is presented in Figure 10.6. The end-to-end transport of messages between OS and NE is realized with the OSI Class 2 transport protocol (TP2), which allows the setup and multiplexing of end-to-end transport connections over an X.25 or SCCP connection. Error detection and data security are not provided in TP2; they are not needed since X.25 as well as SCCP offer a secure message transport service already.

Of course, the OSI protocol stack also needs the protocols for the data link and presentation layers. The *OSI Common Management Information Service Element* (CMISE) plays the central role in GSM network management. Its services are used by a *System Management Application Process* (SMAP) to issue commands, to receive notifications, to check parameters, etc. For file transfer between objects, GSM TMN uses the *OSI File Transfer Access and Management Protocol* (FTAM). It is designed for the efficient transport of large volumes of data.

CMISE needs a few more *Service Elements* (SEs) in the application layer for providing services: the *Association Control Service Element* (ACSE) and the *Remote Operations Service Element* (ROSE). The ACSE is a sublayer of the application layer which allows application elements (here CMISE) to set up and take down connections between each other. The ROSE services are realized with a protocol which enables initiation or execution of operations on remote systems. This way ROSE implements the paradigm also known as *Remote Procedure Call* (RPC).

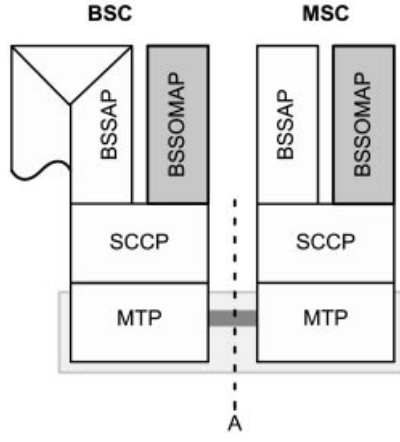

Figure 10.7: Operation and maintenance of the BSS

There is also a management system for the signaling components of a GSM system. This SS#7 SMAP uses the services of the *Operation Maintenance and Administration Part* (OMAP) which allows observation, configuration, and control of the SS#7 network resources. Essentially, the OMAP consists of two *Application Service Elements* (ASEs),

the *MTP Routing Verification Test* (MRVT) and the *SCCP Routing Verification Test* (SRVT) which allow verification of whether the SS#7 network works properly on the MTP or SCCP planes. Another *Management Application Part* is the *Base Station System Operation and Maintenance Application Part* (BSSOMAP) which is used to transport management messages from OMC to BSC through the MSC over the A interface and to execute management activities for the BSS (Figure 10.7, and compare it with Figure 7.11) [53].

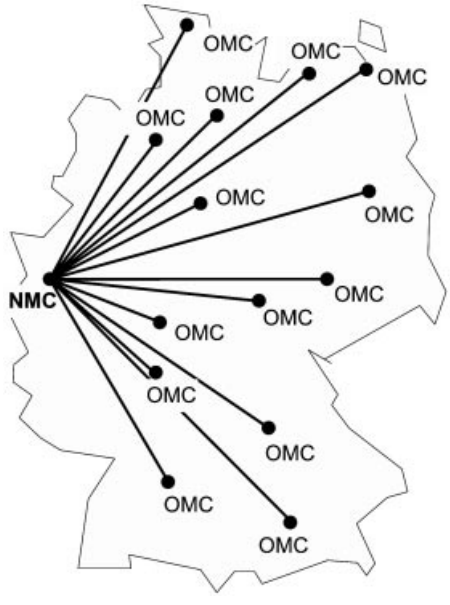

Figure 10.8: Hierarchical organization of network management within Germany

Network management is usually organized in a geographically centralized way. For the remote surveillance and control of network management functions there are usually one or more *Operation and Maintenance Centers* (OMCs). For efficient network management, these OMCs can be operated as regional subcenters according to the LLA hierarchy of the various TMN management planes, and they can be combined under a central *Network Management Centre* (NMC); see Figure 10.8.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

11 General Packet Radio Service (GPRS)

Packet data transmission has already been standardized in GSM phase 2, offering access to the *Packet Switched Public Data Network* (PSPDN); see Sections 9.5.3 and 9.6.2. However, on the air interface such access occupies a complete circuit switched traffic channel for the entire call period. In case of bursty traffic (e.g. Internet traffic), such access leads to a highly inefficient resource utilization. It is obvious that in this case, packet switched bearer services result in a much better utilization of the traffic channels. This is because a packet channel will only be allocated when needed and will be released after the transmission of the packets. With this principle, multiple users can share one physical channel (statistical multiplexing).

In order to address these inefficiencies, the *General Packet Radio Service* (GPRS) has been developed in GSM phase 2+. It offers a genuine packet switched bearer service for GSM also at the air interface. GPRS thus highly improves and simplifies the wireless access to packet data networks. Networks based on the *Internet Protocol* (IP) (e.g. the global Internet or private/corporate intranets) and X.25 networks are supported. In order to introduce GPRS to existing GSM networks, several modifications and enhancements must be made in the network infrastructure as well as in the mobile stations.

Users of GPRS benefit from higher data rates and shorter access times. In conventional GSM, the connection setup takes several seconds and rates for data transmission are restricted to 9.6 kbit/s. GPRS, in practice, offers almost ISDN-like data rates up to approx. 40–50 kbit/s and session establishment times below one second. Furthermore, GPRS supports a more user-friendly billing than that offered by circuit switched data services. In circuit switched services, billing is based on the duration of the connection. This is unsuitable for applications with bursty traffic, since the user must pay for the entire airtime even for idle periods when no packets are sent (e.g. when the user reads a Web page). In contrast to this, with packet switched services, billing can be based on the amount of transmitted data (e.g. Mbyte) and the *Quality of Service* (QoS). The advantage for the user is that he or she can be "online" over a long period of time but will be billed mainly based on the transmitted data volume. The network operators can utilize their radio resources in a more efficient way and simplify the access to external data networks.

The structure of this chapter is as follows:[1] Section 11.1 gives an overview of the GPRS system architecture and explains the fundamental functionality. Next, in Section 11.2, we describe the offered services and the Quality of Service parameters. Section 11.3 explains the session and mobility management and routing. It answers e.g. the questions: How does a GPRS mobile station register with the network? How does the network keep track of the mobile station's location? Section 11.4 gives an overview of the GPRS protocol architecture and briefly introduces the protocols developed for GPRS. Next, an example of a GPRS-Internet interconnection is given (Section 11.5). Section 11.6 discusses the air interface, including the multiple access concept and radio resource management. Moreover, the logical channels and their mapping onto physical channels are explained. Section 11.6.4 considers GPRS channel coding. GPRS security issues are treated in Section 11.7, and, finally, a brief summary of the main features of GPRS is given.

11.1 System Architecture

In order to integrate GPRS into the existing GSM architecture (see e.g. Figure 3.9), a new class of network nodes, called *GPRS Support Nodes* (GSNs), has been introduced. GSNs are responsible for the delivery and routing of data packets between the mobile stations and external *packet data networks* (PDNs). Figure 11.1 illustrates the resulting system architecture.

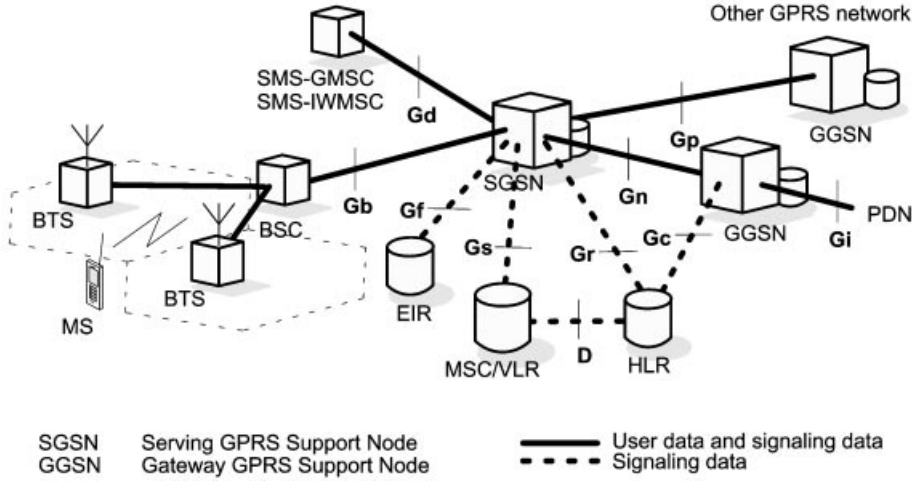

Figure 11.1: GPRS system architecture and interfaces

A *Serving GPRS Support Node* (SGSN) delivers data packets from and to the mobile stations within its service area. Its tasks include packet routing and transfer, functions

[1] Parts of this chapter are based on the authors' publication: Ch. Bettstetter, H.-J. Vögel, J. Eberspächer. GSM Phase 2+ General Packet Radio Service GPRS: Architecture, Protocols, and Air Interface. *IEEE Communications Surveys*, Special Issue on Packet Radio Networks, vol. 2, no. 3, 1999, which can be obtained at http://www.comsoc.org/pubs/surveys. © 1999 IEEE.

for attach/detach of mobile stations and their authentication, and logical link management. The location register of the SGSN stores location information (e.g. current cell, current VLR) and user profiles (e.g. IMSI, address used in the packet data network) of all GPRS users registered with this SGSN.

A *Gateway GPRS Support Node* (GGSN) acts as an interface to external packet data networks (e.g. to the Internet). It converts GPRS packets coming from the SGSN into the appropriate *Packet Data Protocol* (PDP) format (i.e. IP or X.25) and sends them out on the corresponding external network. In the other direction, the PDP address of incoming data packets (e.g. the IP destination address) is converted to the GSM address of the destination user. The readdressed packets are sent to the responsible SGSN. For this purpose, the GGSN stores the current SGSN addresses and profiles of registered users in its location register.

In general, there is a many-to-many relationship between the SGSNs and the GGSNs: A GGSN is the interface to an external network for several SGSNs; an SGSN may route its packets to different GGSNs.

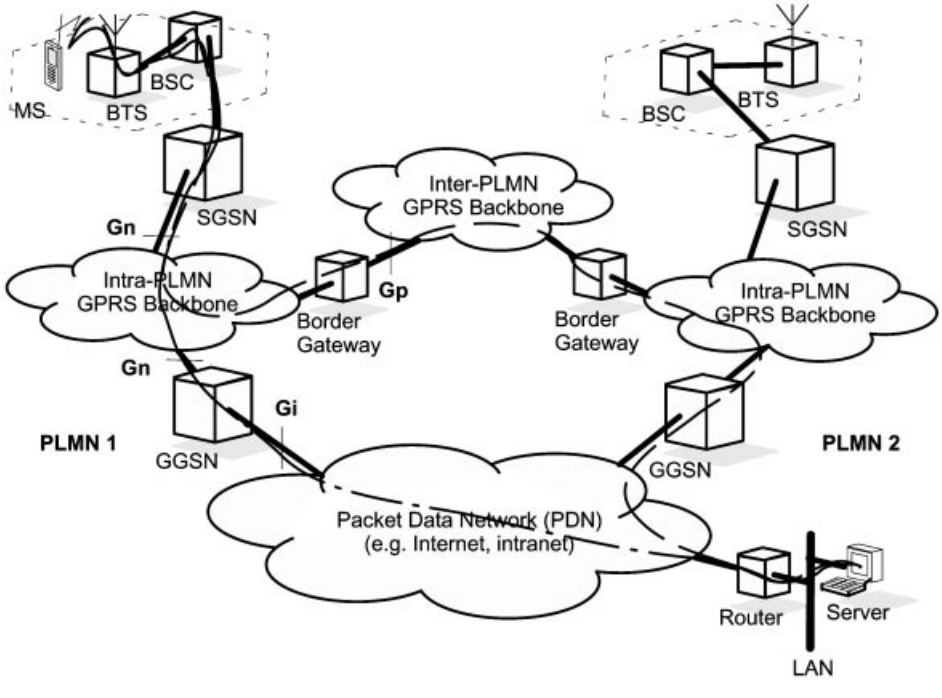

Figure 11.2: GPRS system architecture, interfaces, and routing example

Figure 11.1 also shows the interfaces between the GPRS support nodes and the GSM network. The Gb interface connects the BSC with the SGSN. Via the Gn and the Gp interfaces, user and signaling data are transmitted between the GSNs. The Gn interface is used, if SGSN and GGSN are located in the same PLMN, whereas the Gp interface is used, if they are in different PLMNs.

All GSNs are connected via an IP-based GPRS backbone network. Within this backbone, the GSNs encapsulate the PDN packets and transmit (tunnel) them using the so-called *GPRS Tunneling Protocol* (GTP). In principle, we can distinguish between two kinds of GPRS backbones:

- *Intra-PLMN backbones* are IP-based networks owned by the GPRS network provider connecting the GSNs of the GPRS network.
- *Inter-PLMN backbone networks* connect GSNs of different GPRS networks. They are installed if there is a roaming agreement between two GPRS network providers.

Figure 11.2 shows, how two Intra-PLMN backbone networks of different PLMNs are connected with an Inter-PLMN backbone. The gateways between the PLMNs and the external Inter-PLMN backbone are called *Border Gateways* (BGs). Their main task is to perform security functions in order to protect the private Intra-PLMN backbones against unauthorized users and attacks. The illustrated routing example is explained later.

The Gn and Gp interfaces are also defined between two SGSNs. This allows the SGSNs to exchange user profiles when a mobile station moves from one SGSN area to another.

Across the Gf interface, the SGSN may query and check the IMEI of a mobile station trying to register with the network.

The Gi interface connects the PLMN with external PDNs. In the GPRS standard, interfaces to IP (IPv4 and IPv6) and X.25 networks are supported.

GPRS also adds some more entries to the GSM registers. For mobility management, the user's entry in the HLR is extended with a link to its current SGSN. Moreover, his or her GPRS-specific profile and current PDP address(es) are stored. The Gr interface is used to exchange this information between HLR and SGSN. For example: The SGSN informs the HLR about the current location of the MS. When an MS registers with a new SGSN, the HLR will send the user profile to the new SGSN. In a similar manner, the signaling path between GGSN and HLR (Gc interface) may be used by the GGSN to query the location and profile of a user who is unknown to the GGSN.

In addition, the MSC/VLR may be extended with functions and register entries which allow efficient coordination between packet switched (GPRS) and conventional circuit switched GSM services. Examples for this are combined GPRS and GSM location updates and combined attachment procedures. Moreover, paging requests of circuit switched GSM calls can be performed via the SGSN. For this purpose, the Gs interface connects the registers of SGSN and MSC/VLR.

Finally, it is worth mentioning that it is possible to exchange messages of the *Short Message Service* (SMS) via GPRS. The Gd interface interconnects the *SMS Gateway MSC* (SMS-GMSC) with the SGSN.

11.2 Services

11.2.1 Bearer Services and Supplementary Services

The bearer services of GPRS offer end-to-end packet switched data transfer to mobile

subscribers. Currently, a *Point-to-Point* (PTP) service is specified, which comes in two variants: a connectionless mode (*PTP Connectionless Network Service* (PTP-CLNS), e.g. for IP) and a connection-oriented mode (*PTP Connection Oriented Network Service* (PTP-CONS), e.g. for X.25).

For future releases it is planned to implement a *Point-to-Multipoint* (PTM) service. It will offer transfer of data packets from one user to a group of users/stations. Three kinds of PTM services are possible (also see the comparison in Table 11.1):

- The *Multicast Service* (PTM-M) broadcasts data packets to all users in a certain geographical area. A group identifier indicates, whether the packets are intended for all users in this area or only for a particular group of users.
- Using the *Group Call Service* (PTM-G), data packets are addressed to a particular group of users (PTM group). A geographical area is not taken into account in this case. Users that intend to receive messages must actively become member of this PTM group. PTM-G packets are only sent out in those areas where members of the destination group are currently located.
- Furthermore, it is possible to use IP multicast routing protocols (see e.g. [51]) over GPRS. Packets addressed to an IP multicast group will then be routed to all group members.

Table 11.1: Point-to-Multipoint (PTM) services

Characteristics	PTM-M	PTM-G	IP multicast
Addressed to:	Geographical area	Particular user group	Particular user group
Secondary addressing	Particular user group	–	–
Are the receivers known?	No, anonymous	Yes	Yes
Acknowledged transmission	No	Optional	Yes
Ciphering	No	Yes	Yes

Furthermore, SMS messages can be sent and received over GPRS. It is planned to additionally implement some supplementary services, such as *Closed User Group* (CUG) and *Barring* services.

Based on these standardized services, GPRS providers may offer additional non-standardized services. Examples are access to information databases, messaging services (via store-and-forward mailboxes), and transaction services (e.g. credit card validations and electronic monitoring/surveillance systems). The most important application scenario, however, is the wireless access to the *World Wide Web* (WWW) and to corporate intranets as well as e-mail communication.

11.2.2 Quality of Service

The *Quality of Service* (QoS) requirements for the variety of mobile data applications, in which GPRS is used as transmission technology, are very diverse (for example, compare

the requirements of real-time video conferencing with those of e-mail transfer with respect to packet delay and error-free transmission). Support of different QoS classes is therefore an important feature to support a broad variety of applications but still preserve radio and network resources in an efficient way. Moreover, QoS classes enable providers to offer different billing options. The billing can be based on the amount of transmitted data, the service type itself, and the QoS profile. At the moment, four QoS parameters are defined in GPRS: service precedence, reliability, delay, and throughput. Using these parameters, QoS profiles can be negotiated between the mobile user and the network for each session, depending on the QoS demand and the currently available resources.

The *service precedence* is the priority of a service (in relation to other services). There exist three levels of priority: high, normal, and low. In case of heavy traffic load, for example, packets of low priority will be discarded first.

The *reliability* indicates the transmission characteristics required by an application. Three reliability classes are defined (see Table 11.2), which guarantee certain maximum values for the probability of packet loss, packet duplication, mis-sequencing, and packet corruption (i.e. undetected error in a packet).

Table 11.2: Reliability classes

Class	Probability for			
	Lost packet	Duplicated packet	Out of sequence packet	Corrupted packet
1	10^{-9}	10^{-9}	10^{-9}	10^{-9}
2	10^{-4}	10^{-5}	10^{-5}	10^{-6}
3	10^{-2}	10^{-5}	10^{-5}	10^{-2}

Table 11.3: Delay classes

Class	128 byte packet		1024 byte packet	
	Mean delay (s)	95% delay (s)	Mean delay (s)	95% delay (s)
1	< 0.5	< 1.5	< 2	< 7
2	< 5	< 25	< 15	< 75
3	< 50	< 250	< 75	< 375
4	Best effort	Best effort	Best effort	Best effort

The *delay* parameters define maximum values for the *mean delay* and the *95-percentile delay* (see Table 11.3). The latter is the maximum delay guaranteed in 95% of all transfers. Here, ''delay'' is defined as the end-to-end transfer time between two communicating mobile stations or between a mobile station and the Gi interface to an external network,

respectively. This includes all delays within the GPRS network, e.g., the delay for request and assignment of radio resources, transmission over the air interface, and the transit delay in the GPRS backbone network. Delays outside the GPRS network, e.g., in external transit networks, are not taken into account. Table 11.3 lists the four defined delay classes and their parameters for a 128 byte and 1024 byte packet, respectively.

Finally, the *throughput* parameter specifies the *maximum/peak bit rate* and the *mean bit rate*.

11.2.3 Simultaneous Usage of Packet Switched and Circuit Switched Services

In a GSM/GPRS network, conventional circuit switched services (GSM speech, data, and SMS) and GPRS services can be used in parallel. The GPRS standard defines three classes of mobile stations: Mobile stations of class A fully support simultaneous operation of GPRS and conventional GSM services. Class B mobile stations are able to register with the network for both GPRS and conventional GSM services simultaneously and listen to both types of signaling messages, but can only use one of the service types at a given time. Finally, class C mobile stations can attach for either GPRS or conventional GSM services at a given time. Simultaneous registration (and usage) is not possible, except for SMS messages, which can be received and sent at any time.

11.3 Session Management, Mobility Management, and Routing

In this section we describe how a mobile station registers with the GPRS network and becomes known to an external packet data network. We show how packets are routed to or from mobile stations, and how the network keeps track of the user's current location.

11.3.1 Attachment and Detachment Procedure

Before a mobile station can use GPRS services, it must attach to the network (similar to the *IMSI Attach* used for circuit switched GSM services). The mobile station's ATTACH REQUEST message is sent to the SGSN. The network then checks if the user is authorized, copies the user profile from the HLR to the SGSN, and assigns a *Packet Temporary Mobile Subscriber Identity* (P-TMSI) to the user. This procedure is called *GPRS Attach*. For mobile stations using both circuit switched and packet switched services, it is possible to perform combined GPRS/IMSI attach procedures. The disconnection from the GPRS network is called *GPRS Detach*. It can be initiated by the mobile station or by the network.

11.3.2 Session Management and PDP Context

To exchange data packets with external PDNs after a successful GPRS attach, a mobile station must apply for an address used in the PDN. In general, this address is called *PDP*

address (*Packet Data Protocol address*). In case the PDN is an IP network, this will be an IP address.

For each session, a so-called *PDP context* is created, which describes the characteristics of the session. It contains the PDP type (e.g. IPv4), the PDP address assigned to the mobile station (e.g. an IP address), the requested QoS class, and the address of a GGSN that serves as the access point to the external network. This context is stored in the MS, the SGSN, and the GGSN. Once a mobile station has an active PDP context, it is ''visible'' for the external network and can send and receive data packets. The mapping between the two addresses (PDP ↔ GSM address) makes the transfer of data packets between MS and GGSN possible.

The allocation of a PDP address can be static or dynamic. In the first case, the mobile station permanently owns a PDP address, which has been assigned by the network operator of the user's home-PLMN. Using a dynamic addressing concept, a PDP address is assigned upon activation of a PDP context; i.e., each time a mobile station attaches to the network it will in general get a new PDP address, and after its GPRS detach this PDP address will be again available to other MSs. The PDP address can be assigned by the user's home-PLMN operator (*Dynamic Home-PLMN PDP Address*) or by the operator of the visited network (*Dynamic Visited-PLMN PDP Address*). The GGSN is responsible for the allocation and deactivation of the addresses.

Figure 11.3 shows the PDP context activation procedure initialized by the MS. Using the message ACTIVATE PDP CONTEXT REQUEST, the MS informs the SGSN about the requested PDP context. If a dynamic address is requested, the parameter PDP ADDRESS will be left empty. Afterward, the usual GSM security functions (e.g. authentication of the user) are performed. If access is granted, the SGSN will send a CREATE PDP CONTEXT REQUEST to the affected GGSN. The GGSN creates a new entry in its PDP context table, which enables the GGSN to route data packets between the SGSN and the external PDN. It confirms this to the SGSN with a message CREATE PDP CONTEXT RESPONSE, which also contains the dynamic PDP address (if needed). Finally, the SGSN updates its PDP context table and

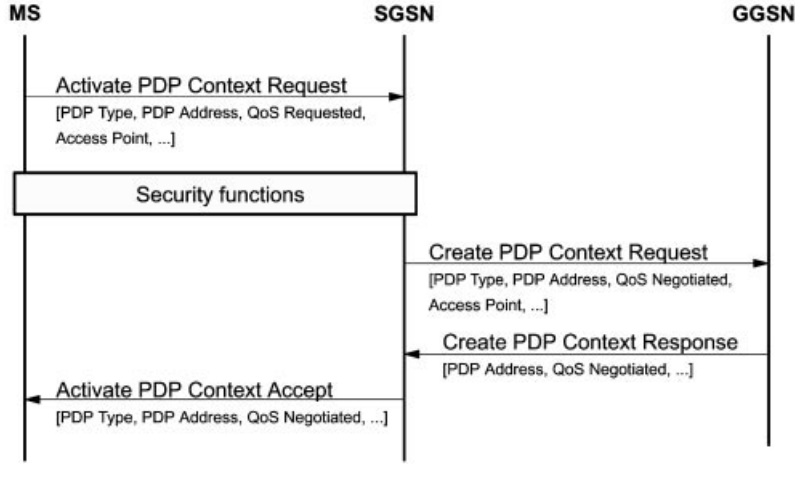

Figure 11.3: PDP context activation

confirms the activation of the new PDP context to the MS (ACTIVATE PDP CONTEXT ACCEPT).

It is also worth mentioning that the GPRS standard supports anonymous PDP context activation, which is useful for special applications such as pre-paid services. In such a session, the user (i.e. the IMSI) using the PDP context remains unknown to the network. Security functions as shown in Figure 11.3 are skipped. Only dynamic address allocation is possible in this case.

11.3.3 Routing

In Figure 11.2 we give an example of how packets can be routed in GPRS. We assume that the packet data network is an IP network.

A GPRS mobile station located in PLMN1 sends IP packets to a Web server connected to the Internet. The SGSN which the mobile station is registered with encapsulates the IP packets coming from the mobile station, examines the PDP context, and routes them through the GPRS backbone to the appropriate GGSN. The GGSN decapsulates the IP-packets and sends them out on the IP network, where IP routing mechanisms transfer the packets to the access router of the destination network. The latter delivers the IP packets to the host.

Let us assume that the mobile station's home-PLMN is PLMN2 and that its IP address has been assigned from the PLMN2 address space – either in a dynamic or static way. When the Web server now addresses IP packets to the MS, they are routed to the GGSN of PLMN2 (the Home-GGSN of the MS). This is because the MS's IP address has the same network prefix as the IP address of its Home-GGSN. The GGSN queries the HLR and obtains the information that the MS is currently located in PLMN1. In the following, it encapsulates the incoming IP packets and tunnels them through the Inter-PLMN GPRS backbone to the appropriate SGSN in PLMN1. The SGSN decapsulates the packets and delivers them to the MS.

11.3.4 Location Management

As in circuit switched GSM, the main task of location management is to keep track of the user's current location, so that incoming packets can be routed to his or her MS. For this purpose, the MS frequently sends *location update* messages to its SGSN.

How often should a mobile station send such a message? If it updates its current location (e.g. its cell) rather seldom, the network must perform a paging process in order to search the MS when packets are coming in. This will result in a significant delivery delay. On the other hand, if location updates happen very often, the MS's location is well known to the network (and thus the packets can be delivered without any additional paging delay), but quite a lot of uplink radio bandwidth and battery power is used for mobility management in this case. Thus, a good location management strategy must be a compromise between these two extreme methods.

For this reason, a state model for GPRS mobile stations has been defined (shown in Figure 11.4). In IDLE state the MS is not reachable. Performing a GPRS attach, it turns into

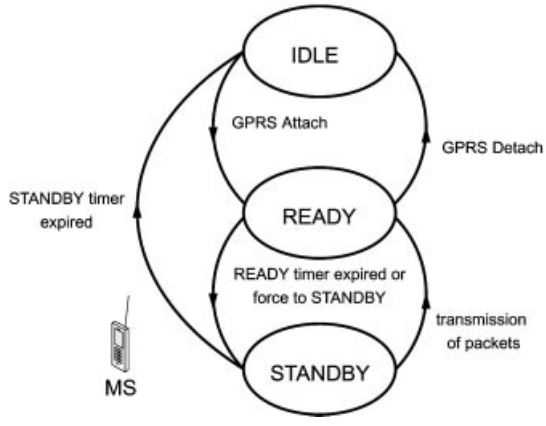

Figure 11.4: State model of a GPRS mobile station

READY state. With a GPRS detach it may deregister from the network and fall back to IDLE state, and all PDP contexts will be deleted. The STANDBY state will be reached when an MS does not send any packets for a long period of time, and therefore the READY timer (which was started at GPRS attach and is reset for each incoming and outgoing transmission) expires. The location update frequency depends on the state in which the MS currently is. In IDLE state, no location updating is performed, i.e., the current location of the MS is unknown. If an MS is in READY state, it will inform its SGSN of every movement to a new cell. For the location management of an MS in STANDBY state, a GSM *Location Area* (LA) is divided into so-called *Routing Areas* (RAs). In general, an RA consists of several cells. The SGSN will only be informed, when an MS moves to a new RA; cell changes will not be indicated.

To find out the current cell of an MS that is in STANDBY state, paging of the MS within a certain RA must be performed (see Figure 11.15). For MSs in READY state, no paging is necessary.

Whenever an MS moves to a new RA, it sends a ROUTING AREA UPDATE REQUEST to its

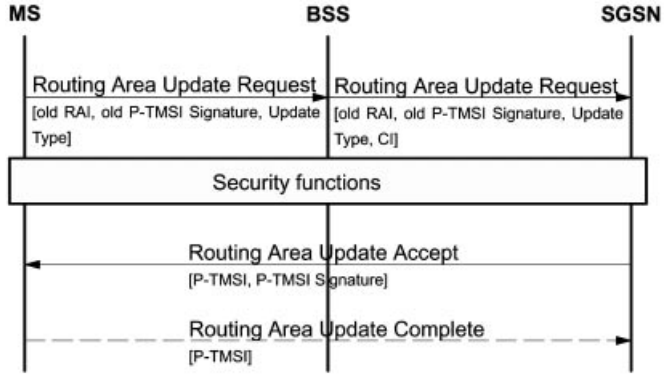

Figure 11.5: Intra-SGSN routing area update

assigned SGSN (see Figure 11.5). The message contains the *Routing Area Identity* (RAI) of its old RA. The BSS adds the *Cell Identifier* (CI) of the new cell to the request, from which the SGSN can derive the new RAI. Two different scenarios are possible:

- Intra-SGSN Routing Area Update (Figure 11.5)
- Inter-SGSN Routing Area Update (Figure 11.6)

In the Intra-SGSN case, the MS has moved to an RA which is assigned to the same SGSN as the old RA. In this case, the SGSN has already stored the necessary user profile and can immediately assign a new P-TMSI (ROUTING AREA UPDATE ACCEPT). Since the routing context does not change, there is no need to inform other network elements, such as GGSN or HLR.

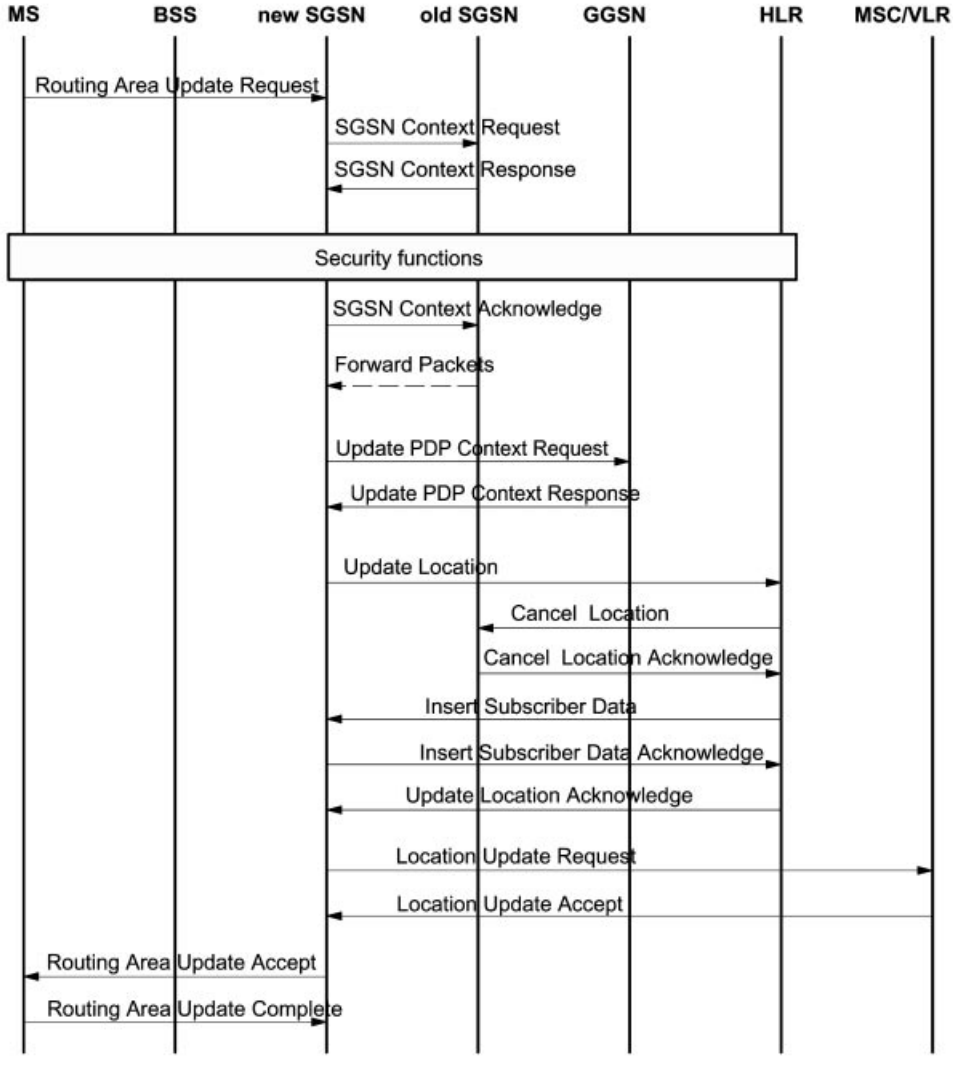

Figure 11.6: Inter-SGSN routing area update

In the Inter-SGSN case, the new RA is administered by a different SGSN than the old RA. The new SGSN realizes that the MS has entered its area and requests the old SGSN to send the PDP contexts of the user (SGSN CONTEXT REQUEST, SGSN CONTEXT RESPONSE, SGSN CONTEXT ACKNOWLEDGE). Afterward, the new SGSN informs the involved GGSNs about the user's new routing context (UPDATE PDP CONTEXT REQUEST, UPDATE PDP CONTEXT RESPONSE). In addition, the HLR and (if needed) the MSC/VLR are informed about the user's new SGSN number (UPDATE LOCATION,..., UPDATE LOCATION ACKNOWLEDGE; LOCATION UPDATE REQUEST, LOCATION UPDATE ACCEPT).

Besides pure RA updates, there also exist combined RA/LA updates. They are performed whenever an MS using GPRS as well as conventional GSM services moves to a new LA. The MS sends a ROUTING AREA UPDATE REQUEST to the SGSN and uses a parameter *update type* to indicate that an LA update is needed. The message is then forwarded from the SGSN to the VLR.

To sum up, we can say that GPRS mobility management consists – as GSM mobility management – of two levels: Micro mobility management tracks the current RA or cell of the user. Macro mobility management keeps track of the user's current SGSN and stores it in the HLR, VLR, and GGSN.

11.4 Protocol Architecture

11.4.1 Transmission Plane

Figure 11.7 illustrates the protocol architecture of the GPRS transmission plane. The protocols offer transmission of user data and its associated signaling (e.g. for flow control, error detection, and error correction). An application running in the GPRS-MS (e.g. a browser) uses IP or X.25, respectively, in the network layer.

11.4.1.1 GPRS Backbone: SGSN–GGSN

As mentioned earlier in this chapter, IP and X.25 packets are transmitted encapsulated within the GPRS backbone network. This is done using the *GPRS Tunneling Protocol* (GTP), i.e., GTP packets carry the user's IP or X.25 packets. GTP is defined both between GSNs within the same PLMN (Gn interface) and between GSNs of different PLMNs (Gp interface).

It contains procedures in the transmission plane as well as in the signaling plane. In the transmission plane, GTP employs a tunnel mechanism to transfer user data packets. In the signaling plane, GTP specifies a tunnel control and management protocol. The signaling is used to create, modify, and delete tunnels. A *Tunnel Identifier* (TID), which is composed of the IMSI of the user and a *Network Layer Service Access Point Identifier* (NSAPI), uniquely indicates a PDP context. Below GTP, the standard protocols TCP or UDP are employed to transport the GTP packets within the backbone network. TCP is used for X.25 (since X.25 expects a reliable end-to-end connection), and UDP is used for access to IP-based networks (which do not expect reliability in the network layer or below). In the network layer, IP is employed to route the packets through the backbone. Ethernet, ISDN,

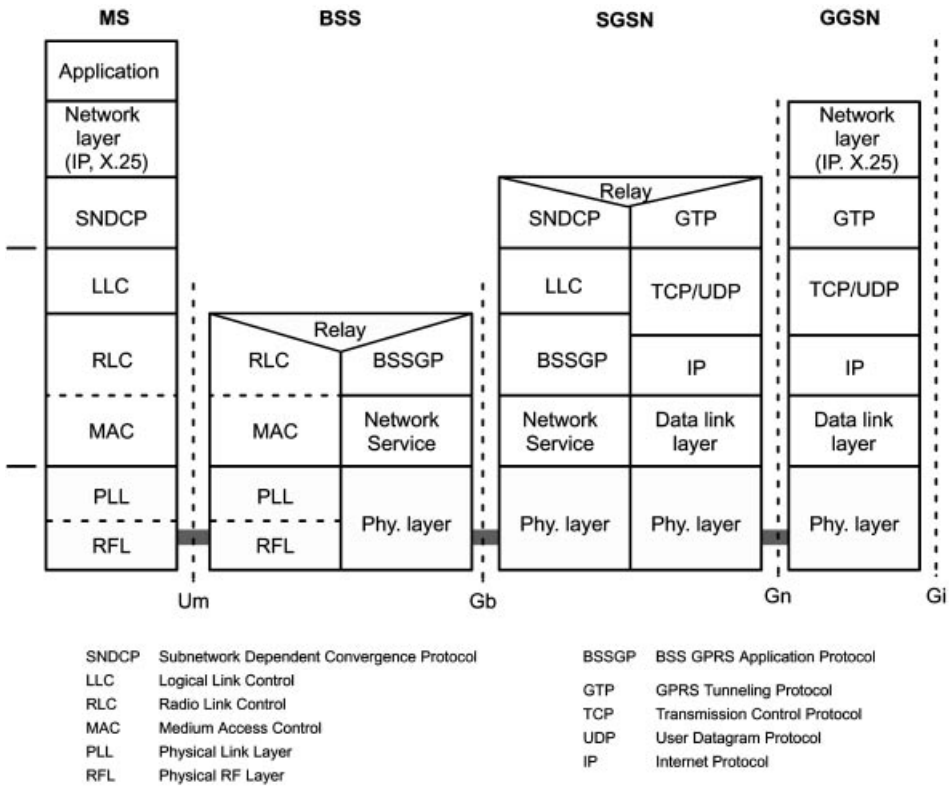

Figure 11.7: Protocol architecture: transmission plane

or ATM-based protocols may be used below IP. To summarize, in the GPRS backbone we have an IP/X.25-over-GTP-over-UDP/TCP-over-IP protocol architecture.

11.4.1.2 Air Interface

In the following we consider the air interface (Um) between MS and BSS or SGSN, respectively.

Subnetwork Dependent Convergence Protocol – The *Subnetwork Dependent Convergence Protocol* (SNDCP) is used to transfer packets of the network layer (IP and X.25 packets) between the MSs and their SGSN. Its functionality includes:

- multiplexing of several PDP contexts of the network layer onto one virtual logical connection of the underlying LLC layer, and
- segmentation of network layer packets onto one frame of the underlying LLC layer and reassembly on the receiver side.

Moreover, SNDCP offers compression and decompression of user data and redundant header information (e.g. TCP/IP header compression).

Data Link Layer – The data link layer is divided into two sublayers:

- *Logical Link Control* (LLC) layer (between MS and SGSN) and
- *Radio Link Control/Medium Access Control* (RLC/MAC) layer (between MS and BSS).

The LLC layer provides a reliable logical link between an MS and its assigned SGSN. Its functionality is based on the LAPDm protocol (which is a protocol similar to HDLC and has been explained in Section 7.3.1). LLC includes in-order delivery, flow control, error detection and retransmission of packets (*Automatic Repeat Request* (ARQ)), and ciphering functions. It supports variable frame lengths and different QoS classes, and besides point-to-point also point-to-multipoint transfer is possible. A logical link is uniquely addressed with a *Temporary Logical Link Identifier* (TLLI). Within one RA the mapping between TLLI and IMSI is unique. However, the user's identity remains confidential, since the TLLI is derived from the P-TMSI of the user.

The RLC/MAC layer has two functions. The purpose of the *Radio Link Control* (RLC) layer is to establish a reliable link between the MS and the BSS. This includes the segmentation and reassembly of LLC frames into RLC data blocks and ARQ of uncorrectable blocks. The *Medium Access Control* (MAC) layer controls the access attempts of mobile stations on the radio channel. It is based on a Slotted-Aloha principle (see Section 5.1). The MAC layer employs algorithms for contention resolution of access attempts, statistical multiplexing of channels, and a scheduling and prioritizing scheme, which takes into account the negotiated QoS. On the one hand, the MAC protocol allows that a single MS simultaneously uses several physical channels (several time slots of the same TDMA frame). On the other hand, it also controls the statistical multiplexing, i.e., it controls how several MSs can access the same physical channel (the same time slot of successive TDMA frames). This is explained in more detail in Section 11.6.

Physical Layer – The physical layer between MS and BSS can be divided into the two sublayers: *Physical Link Layer* (PLL) and *Physical RF Layer* (RFL). The PLL provides a physical channel between the MS and the BSS. Its tasks include channel coding (i.e., detection of transmission errors, *Forward Error Correction* (FEC), and indication of uncorrectable codewords), interleaving, and detection of physical link congestion. The RFL, which operates below the PLL, includes modulation and demodulation.

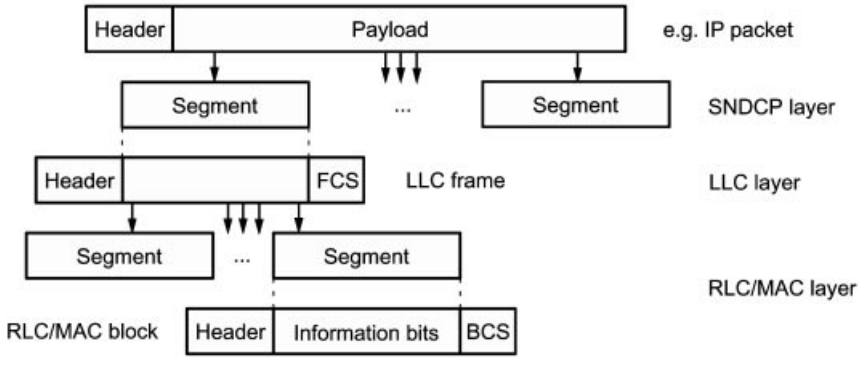

Figure 11.8: Data flow and segmentation between the protocol layers in the MS

To summarize this section, Figure 11.8 illustrates the data flow between the protocol layers in the mobile station. Packets of the network layer (e.g. IP packets) are passed down to the SNDCP layer, where they are segmented to LLC frames. After adding header information and a *Frame Check Sequence* (FCS) for error protection, these frames are segmented into one or several RLC data blocks. Those are then passed down to the MAC layer. One RLC/MAC block contains a MAC and RLC header, the RLC payload (''information bits''), and a *Block Check Sequence* (BCS) at the end. The channel coding of RLC/MAC blocks and the mapping to a burst in the physical layer are explained in Section 11.6.

11.4.1.3 BSS – SGSN Interface

At the Gb interface, the *BSS GPRS Application Protocol* (BSSGP) is defined on Layer 3. It is derived from the *BSS Management Part* (BSSMAP), which has been explained in Section 7.3.1. The BSSGP delivers routing and QoS-related information between BSS and SGSN. The underlying *Network Service* (NS) protocol is based on the *Frame Relay* protocol.

11.4.2 Routing and Conversion of Addresses

Now we explain the routing example of Section 11.3.3 in detail. Figure 11.9 roughly illustrates the transfer of an incoming IP packet. It arrives at the GGSN, is then routed through the GPRS backbone to the responsible SGSN and finally to the MS. Using the PDP context, the GGSN determines from the IP destination address a *Tunnel Identifier* (TID) and the IP address of the relevant SGSN. Between GGSN and the SGSN, the *GPRS Tunneling Protocol* is employed. The SGSN derives the *Temporary Logical Link Identifier* (TLLI) from the TID and finally transfers the IP packet to the MS. The so-called *Network Service Access Point Identifier* (NSAPI) is part of the TID. It maps a given IP

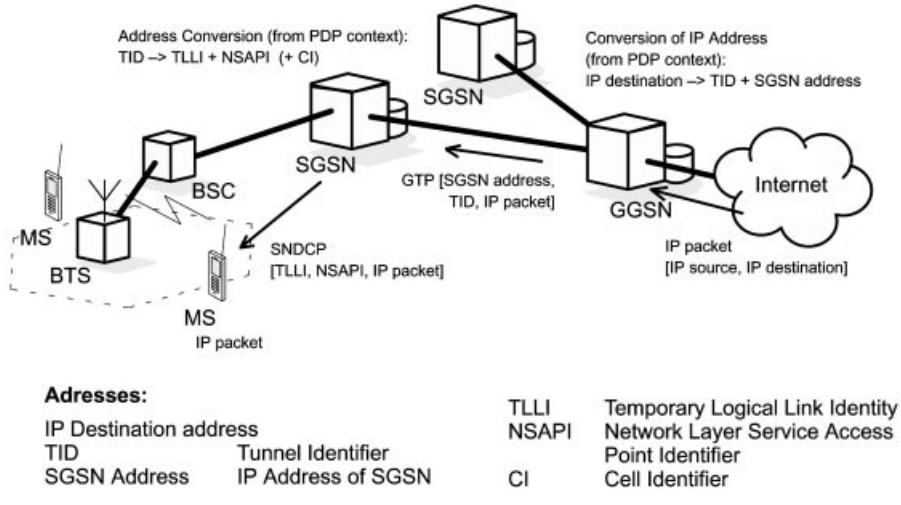

Figure 11.9: Routing and address conversion: Incoming IP packet (mobile terminated data transfer)

address to the corresponding PDP context. An NSAPI/TLLI pair is unique within one RA. Figure 11.10 gives a similar example with an outgoing (mobile originated) IP packet.

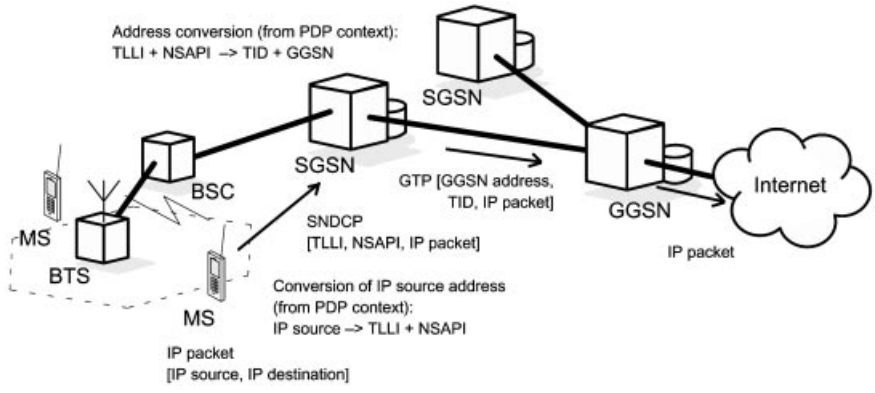

Figure 11.10: Routing and address conversion: Outgoing IP packet (mobile originated data transfer)

11.4.3 Signaling Plane

The protocol architecture of the signaling plane comprises protocols for control and support of the functions of the transmission plane, e.g., for the execution of GPRS attach and detach, PDP context activation, the control of routing paths, and the allocation of network resources.

Between MS and SGSN (Figure 11.11), the *GPRS Mobility Management and Session Management* (GMM/SM) protocol is responsible for mobility and session management. It includes functions for GPRS attach/detach, PDP context activation, routing area updates, and security procedures.

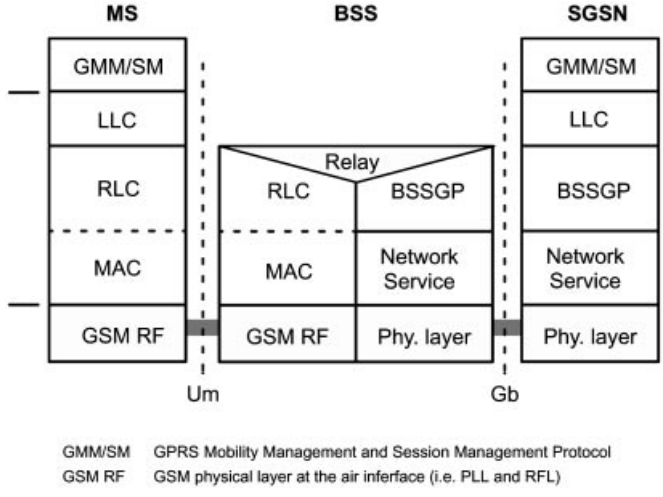

Figure 11.11: Signaling plane: MS-SGSN

The signaling architecture between SGSN and the registers HLR, VLR, and EIR (Figure 11.12) uses protocols known from conventional GSM (Section 7.3) and partly extends them with GPRS-specific functionality. Between SGSN and HLR as well as between SGSN and EIR, an enhanced *Mobile Application Part* (MAP) is employed. The exchange of MAP messages is accomplished over the *Transaction Capabilities Application Part* (TCAP), the *Signaling Connection Control Part* (SCCP), and the *Message Transfer Part* (MTP).

The *BSS Application Part* (BSSAP+) includes functions of GSM's BSSAP. It is applied to transfer signaling information between the SGSN and the VLR (Gs interface). This includes, in particular, signaling of the mobility management when coordination of GPRS and conventional GSM functions is necessary (e.g. for combined GPRS and non-GPRS location update, combined GPRS/IMSI attach, or paging of an MS via GPRS for an incoming GSM call).

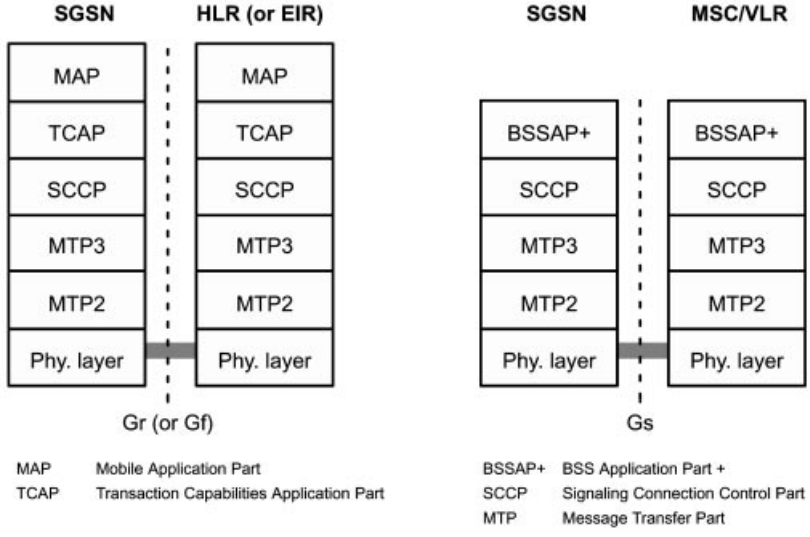

Figure 11.12: Signaling plane: SGSN-HLR, SGSN-EIR, and SGSN-MSC/VLR

11.5 Interworking with IP Networks

Figure 11.13 gives an example of how a GPRS network is interconnected with the Internet. From outside, i.e., from an external IP network's point of view, the GPRS network looks like any other IP subnetwork, and the GGSN looks like a usual IP router.

As explained in Section 11.3.2, each mobile station obtains an IP address after its GPRS attach, which is valid for the duration of the session. The network provider has reserved a certain number of IP addresses, and can dynamically assign these addresses to active mobile stations. To do so, the network provider may install a DHCP server (*Dynamic Host Configuration Protocol*) in its network. This server automatically manages the available address space. The address resolution between IP address and GSM address is performed by the GGSN, using the appropriate PDP context. The routing of IP packets

and the tunneling through the Intra-PLMN backbone (using GTP) has been explained in Sections 11.1, 11.3, and 11.4.2.

Moreover, a *Domain Name Server* (DNS) is used to map between IP addresses and host names. To protect the PLMN from unauthorized access, a firewall is installed between the private GPRS network and the external IP network.

With this configuration, GPRS can be seen as a wireless extension of the Internet all the way to a mobile station. The mobile user has direct connection to the Internet.

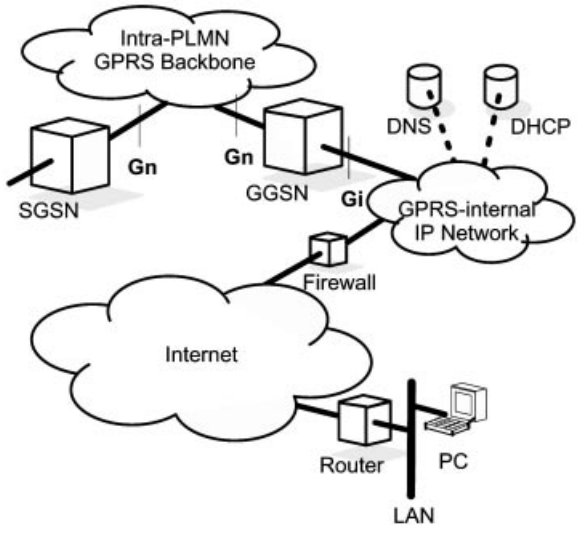

Figure 11.13: GPRS–Internet interconnection

11.6 Air Interface

The enhanced air interface of GPRS offers higher data rates and a packet-oriented transmission. It is therefore considered one of the key aspects in GPRS. In this section, we explain how several mobile stations can share one physical channel (multiple access) and how the assignment of radio resources between circuit-switched GSM services and GPRS services is controlled. Afterward, the logical channels and their mapping onto physical channels (using multiframes) is presented. Finally, GPRS channel coding concludes this chapter.

11.6.1 Multiple Access and Radio Resource Management

On the physical layer, GPRS uses the GSM combination of FDMA and TDMA with 8 time slots per TDMA frame (as explained in Section 5.2.2). However, several new methods are used for channel allocation and multiple access. They have significant impact on the performance of GPRS.

In circuit switched GSM, a physical channel (i.e. one time slot of successive TDMA

frames) is permanently allocated for a particular MS during the entire call period (no matter whether data is transmitted or not). Moreover, it is assigned in the uplink as well as in the downlink.

GPRS enables a far more flexible resource allocation scheme for packet transmission. A GPRS mobile station can transmit on several of the 8 time slots within the same TDMA frame (*multislot operation*). The number of time slots which an MS is able to use is called *multislot class*. In addition, up- and downlink are allocated separately, which saves radio resources, especially for asymmetric traffic (e.g. Web browsing).

A cell supporting GPRS must allocate physical channels for GPRS traffic. In other words, the radio resources of a cell are shared by all mobile stations (GSM and GPRS) located in this cell. The mapping of physical channels to either GPRS or circuit switched GSM services can be performed in a dynamic way. A physical channel which has been allocated for GPRS transmission is denoted as *Packet Data Channel* (PDCH). The number of PDCHs can be adjusted according to the current traffic demand (*Capacity on Demand principle*). For example, physical channels not currently in use by GSM calls can be allocated as PDCHs for GPRS to increase the quality of service for GPRS. When there is a resource demand for GSM calls, PDCHs may be de-allocated.

As already mentioned, physical channels for packet switched transmission (PDCHs) are only allocated for a particular MS when this MS sends or receives data packets, and they are released after the transmission. With this dynamic channel allocation principle, multiple MSs can share one physical channel. For bursty traffic this results in a much more efficient usage of the radio resources.

The channel allocation is controlled by the BSC. To prevent collisions, the network indicates in the downlink, which channels are currently available. An *Uplink State Flag* (USF) in the header of downlink packets shows which MS is allowed to use this channel in the uplink. The allocation of PDCHs to an MS also depends on its multislot class and the QoS of the session.

11.6.2 Logical Channels

Table 11.4 lists the packet data logical channels defined in GPRS. As with logical channels in conventional GSM, they can be divided up into two categories: traffic channels and signaling (control) channels. The signaling channels can further be divided into *Packet Broadcast Control*, *Packet Common Control*, and *Packet Dedicated Control* channels.

The *Packet Data Traffic Channel* (PDTCH) is employed for the transfer of user data. It is assigned to one mobile station (or in case of PTM to multiple mobile stations). One mobile station can use several PDTCHs simultaneously.

The *Packet Broadcast Control Channel* (PBCCH) is a unidirectional point-to-multipoint signaling channel from the BSS to the mobile stations. It is used by the BSS to broadcast information about the organization of the GPRS radio network to all GPRS mobile stations of a cell. Besides system information about GPRS, the PBCCH should also broadcast important system information about circuit switched services, so that a GSM/GPRS mobile station does not need to listen to the *Broadcast Control Channel* (BCCH).

Table 11.4: Logical Channels in GPRS

Group		Channel	Function	Direction
Traffic channels	Packet data traffic channel	PDTCH	Packet data traffic	MS ↔ BSS
Signaling channels	Packet broadcast control channel	PBCCH	Packet broadcast control	MS ← BSS
	Packet common control channel (PCCCH)	PRACH	Packet random access	MS → BSS
		PAGCH	Packet access grant	MS ← BSS
		PPCH	Packet paging	MS ← BSS
		PNCH	Packet notification	MS ← BSS
	Packet dedicated control channels	PACCH	Packet associated control	MS ↔ BSS
		PTCCH	Packet timing advance control	MS ↔ BSS

The *Packet Common Control Channel* (PCCCH) transports signaling information for functions of the network access management, i.e., for allocation of radio channels, medium access control, and paging. Four sub-channels are defined:

- The *Packet Random Access Channel* (PRACH) is used by the mobile stations to request one or more PDTCH.
- The *Packet Access Grant Channel* (PAGCH) is used to allocate one or more PDTCH to a mobile station.
- The *Packet Paging Channel* (PPCH) is used by the BSS to find out the location of a mobile station (paging) prior to downlink packet transmission.
- The *Packet Notification Channel* (PNCH) is used to inform mobile stations of incoming PTM messages.

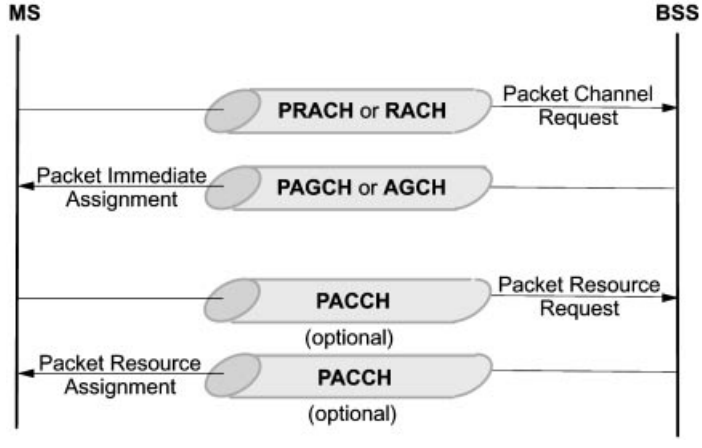

Figure 11.14: Uplink channel allocation (mobile originated packet transfer)

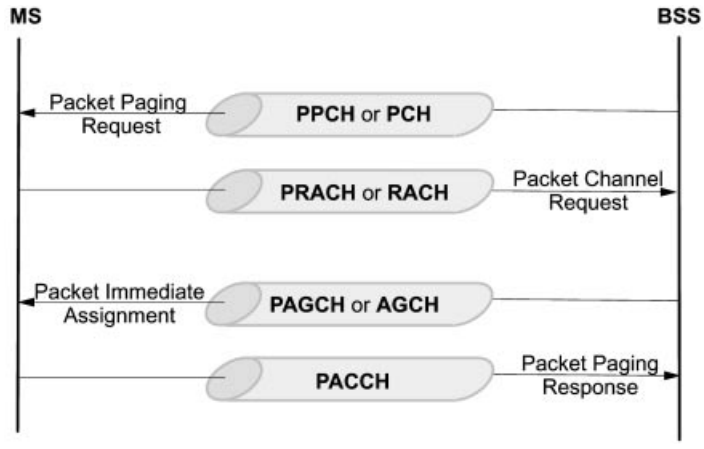

Figure 11.15: Paging (mobile terminated packet transfer)

Figure 11.14 shows the principle of the uplink channel allocation (mobile originated packet transfer). A mobile station requests a channel by sending a PACKET CHANNEL REQUEST on the PRACH or RACH. The BSS answers on the PAGCH or AGCH, respectively. Once the PACKET CHANNEL REQUEST is successful, a so-called *Temporary Block Flow* (TBF) is established. With that, resources (e.g. PDTCH and buffers) are allocated for the mobile station, and data transmission can start. During transfer, the *Uplink State Flag* (USF) in the header of downlink blocks indicates to other MSs that this uplink PDTCH is already in use. On the receiver side, a *Temporary Flow Identifier* (TFI) helps to reassemble the packet. Once all data has been transmitted, the TBF and the resources are released again. Figure 11.15 illustrates the paging procedure of a mobile station (mobile terminated packet transfer).

The *Packet Dedicated Control Channel* is a bidirectional point-to-point signaling channel. It contains the channels PACCH and PTCCH.

- The *Packet Associated Control Channel* (PACCH) is always allocated in combination with one or more PDTCH. It transports signaling information related to one specific mobile station (e.g. power control information).
- The *Packet Timing Advance Control Channel* (PTCCH) is used for adaptive frame synchronization. The MS sends over the uplink part of the PTCCH, the PTCCH/U, access bursts to the BTS. From the delay of these bursts, the correct value for the *Timing Advance* (TA) can be derived; see Section 5.3.2. This value is then transmitted in the downlink part, the PTCCH/D, to inform the MS.

Coordination between circuit switched and packet switched logical channels is also possible here. If the PCCCH is not available in a cell, a GPRS mobile station can use the *Common Control Channel* (CCCH) of circuit switched GSM to initiate the packet transfer. Moreover, if the PBCCH is not available, it can obtain the necessary system information via the *Broadcast Control Channel* (BCCH).

Table 11.5 lists the block lengths and net data throughput of the logical GPRS channels

Table 11.5: Logical channels in GPRS

Channel type	Net data throughput (in kbit/s)	Block length (in bit)	Block distance (in ms)
PDTCH (CS-1)	9.05	181	–
PDTCH (CS-2)	13.4	268	–
PDTCH (CS-3)	15.6	312	–
PDTCH (CS-4)	21.4	428	–
PACCH	Changes dynamically		
PBCCH	s × 181/120	181	120
PAGCH	Changes dynamically	181	
PNCH	Changes dynamically	181	
PPCH	Changes dynamically	181	
PRACH (8 bit Access burst)	Changes dynamically	8	
PRACH (11 bit Access burst)	Changes dynamically	11	

Table 11.6: Combinations of logical GPRS channels

	B10	B11	B12	B13
PDTCH				
PBCCH				
PCCCH				
PACCH				
PTCCH				

Table 11.7: Channel combinations used by the mobile station

	M9	M10
PDTCH		*n+m*
PBCCH		
PCCCH		
PACCH		
PTCCH		

(compare with Table 5.2). Four different coding schemes (CS-1 to CS-4) are defined for data transmission on the PDTCH. They are explained in Section 11.6.4.

As with circuit switched GSM, the GPRS logical channels can be used in certain combinations only. The allowed combinations for multiplexing logical channels onto physical

channels are shown in Table 11.6. Moreover, Table 11.7 shows the channel configurations which a GPRS mobile station can use (dependent on its state). Combination M9 represents a mobile station in IDLE state waiting for incoming packets. Combination M10 is a transmitting mobile station with multislot capabilities. Several PDTCHs are assigned to a single MS, where n denotes the number of PDTCHs which allow bidirectional transmission, and m denotes the number of PDTCHs which allow only unidirectional transmission. We have: $n = 1,\ldots,8$, $m = 0,\ldots,8$, and $n + m = 1,\ldots,8$.

11.6.3 Mapping of packet data logical channels onto physical channels

From Section 5.4 we know that the mapping of logical GSM channels onto physical channels has two components: mapping in frequency and mapping in time. The mapping in frequency is based on the TDMA frame number and the frequencies allocated to the BTS and the mobile station. The mapping in time is based on the definition of complex multiframe structures on top of the TDMA frames.

A multiframe structure for PDCHs consisting of 52 TDMA frames (each with 8 time slots) is shown in Figure 11.16. The corresponding time slots of a PDCH of four consecutive TDMA frames form one *Radio Block* (Blocks B0–B11). Two TDMA frames are reserved for transmission of the PTCCH, and the remaining two frames are IDLE frames. A multiframe has thus a duration of approx. 240 ms (52 × 4.615 ms). A *Radio Block* consists of 456 bits.

The mapping of the logical channels onto the blocks B0–B11 of the multiframe can vary from block to block and is controlled by parameters which are broadcast on the PBCCH. The GPRS recommendations define which time slots may be used by a logical channel.

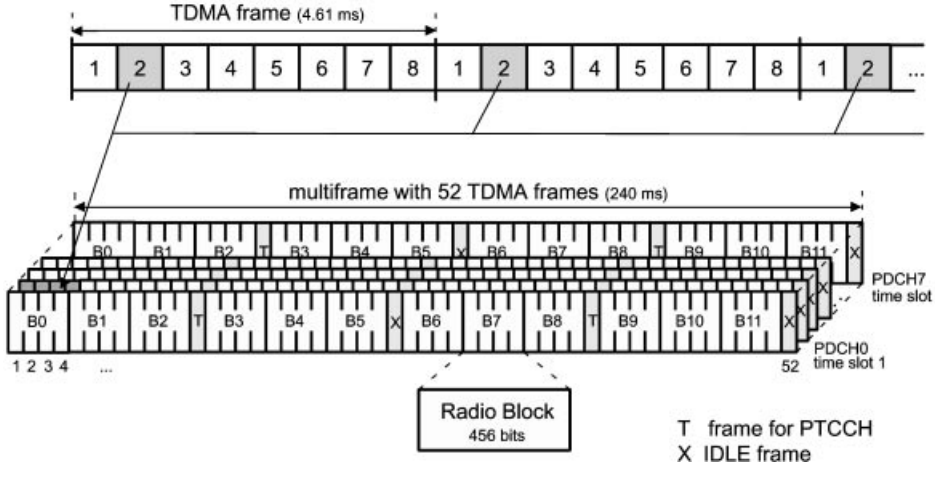

Figure 11.16: Multiframe structure with 52 TDMA frames

Besides the 52-multiframe, which can be used by all logical GPRS channels, also a 51-multiframe structure is defined. It is used for PDCHs carrying only the logical channels PCCCH and PBCCH (channel combination B13 in Table 11.6). In the downlink, it consists

of 10 blocks each 4 frames (B0-B9) and 10 IDLE frames. In the uplink, it has 51 Random Access frames. Its duration is 235.4 ms.

11.6.4 Channel Coding

Figure 11.17 shows how a block of the RLC/MAC layer (compare with Figure 11.8) is encoded and mapped onto four bursts. Channel coding is used to protect the transmitted data packets against errors and perform *Forward Error Correction* (FEC). The channel coding technique in GPRS is quite similar to the one employed in conventional GSM. An outer block coding, an inner convolutional coding, and an interleaving scheme is used (Figure 6.6).

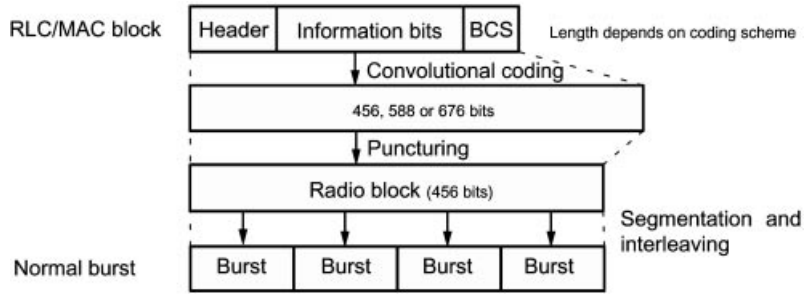

Figure 11.17: Physical layer at the air interface: channel coding, interleaving, and formation of bursts (continued from Figure 11.8)

Four coding schemes with different code rates are defined. Their parameters are listed in Table 11.8. For each scheme, a block of 456 bits results after encoding. Figure 11.18 illustrates the encoding process, which is briefly explained in the following.

Table 11.8: Channel coding schemes for the traffic channels in GPRS

Coding scheme	Pre-encoded USF	Infobits without USF and BCS	Parity bits BCS	Tail bits	Output convolutional encoder	Punctured bits	Code rate	Data rate (kbit/s)
CS-1	3	181	40	4	456	0	1/2	9.05
CS-2	6	268	16	4	588	132	≈2/3	13.4
CS-3	6	312	16	4	676	220	≈3/4	15.6
CS-4	12	428	16	–	456	–	1	21.4

Let us employ coding scheme CS-2. First of all, the 271 information bits of an RLC/MAC block (268 bits plus 3 bits USF, see Table 11.5) are mapped to 287 bits using a systematic block encoder, i.e., 16 parity bits are added. These parity bits are denoted as *Block Check Sequence* (BCS). The USF pre-encoding maps the first 3 bits of the block (i.e. the USF) to 6 bits in a systematic way. Afterward, 4 zero bits (tail bits) are added at the end of the entire block. The tail bits are needed for termination of the subsequent convolutional coding.

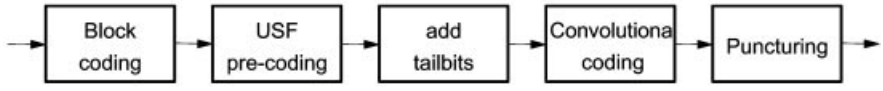

Figure 11.18: Encoding of GPRS data blocks

For the convolutional coding, a non-systematic rate-1/2 encoder with memory 4 is used, which is defined by the generator polynomials

$$G_0(d) = 1 + d^3 + d^4$$

$$G_1(d) = 1 + d + d^3 + d^4$$

This is the same encoder as used in conventional GSM. A possible encoder realization is shown in Figure 6.9. At the output of the convolutional encoder a codeword of length 588 bit results. Afterward, 132 bits are punctured, resulting in a radio block of length 456 bit. Thus, we obtain a code rate of the convolutional encoder (including the puncturing) of

$$r = \frac{6 + 268 + 16 + 4}{456} \approx \frac{2}{3}$$

Coding scheme 1 is equivalent to the coding of the SACCH. A systematic Fire code is used for block coding (see Section 6.2.1.3, first paragraph). There is no pre-coding of the USF bits. The convolutional coding is done with the known rate-1/2 encoder, however, this time the output sequence is not punctured. Using CS-4, the 3 USF bits are mapped to 12 bits, and no convolutional coding is applied.

For the coding of the traffic channels (PDTCH), one of the four coding schemes is chosen, depending on the quality of the signal. The two *stealing flags* in a *normal burst* (Figure 5.6) are used to indicate which coding scheme is used. Under very bad channel conditions, CS-1 yields a data rate of only 9.05 kbit/s per time slot, but a very reliable coding. Under good channel conditions, convolutional coding is skipped (CS-4), and we achieve a data rate of 21.4 kbit/s per time slot. Thus, we obtain a theoretical maximum data rate of 171.2 kbit/s per TDMA frame. In practice, multiple users share the time slots, and, thus, a much lower bit rate is available to the individual user. Moreover, the quality of the radio channel will for sure not always allow us to use CS-4 (or CS-4 is not supported by the mobile terminal or by the network operator). The data rate available to the user depends (among other things) on the current total traffic load in the cell (i.e., the number of users and their traffic characteristics), the used coding scheme, and the multislot class of the MS. Data rates between 10 and 50 kbit/s are realistic values. A simulative study on GPRS performance can be found in [37].

After encoding, the codewords are input into a block interleaver of depth 4. For all coding schemes, the interleaving scheme known from the interleaving of the SACCH (see Section 6.2.3, last paragraph) is employed. On the receiver side, the codewords are deinterleaved. As in GSM, the decoding is performed using the Viterbi Algorithm.

The signaling channels are encoded using CS-1. An exception is the PRACH. It can transmit two very short bursts, one burst with 8 information bits and one burst with 11

information bits. The coding for the 8-bit burst is the one used for the RACH (see Sections 6.2.1.3 and 6.2.2) and the coding for the 11-bit burst is a punctured version of it.

11.7 Authentication and Ciphering

The security principles inside the GPRS network are almost equivalent to those used in conventional GSM (Section 6.3). Security functions in the GPRS network

- protect against unauthorized use of services (by authentication and service request validation),
- provide data confidentiality (using ciphering), and
- provide confidentiality of the subscriber identity.

As in GSM, two keys are used: the *Subscriber Authentication Key* Ki and the *Cipher Key* Kc. The main difference is that not the MSC but the SGSN handles authentication. Moreover, a special GPRS ciphering algorithm (A5) has been defined, which is optimized for encryption of packet data.

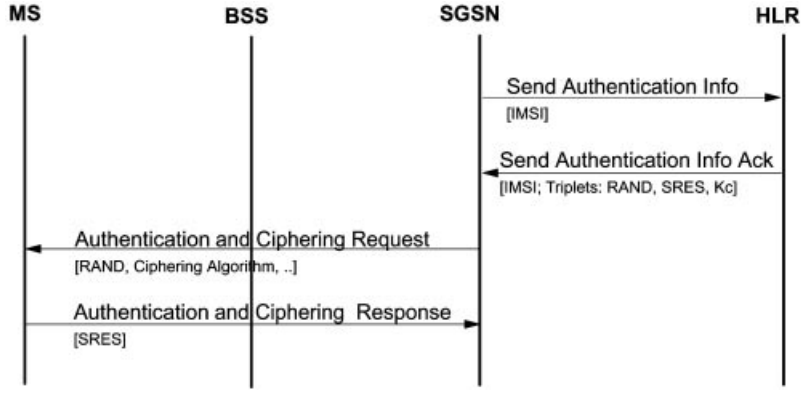

Figure 11.19: Subscriber authentication in GPRS

Figure 11.20: Principle of subscriber authentication in GPRS

11.7.1 User Authentication

Figures 11.19 and 11.20 illustrate the GPRS authentication process. The standard GSM algorithms are used to generate security data. The algorithm A3 calculates the *Signature Response* (SRES) from the *Subscriber Authentication Key* (Ki) and a *Random Number* (RAND).

If the SGSN does not have authentication sets for a user (Kc, RAND, SRES), it requests them from the HLR by sending a message SEND AUTHENTICATION INFO. The HLR responds with a SEND AUTHENTICATION INFO ACK which includes the security data. Now, the SGSN offers a random number RAND to the MS (AUTHENTICATION AND CIPHERING REQUEST). The MS calculates SRES and transmits it back to the SGSN (AUTHENTICATION AND CIPHERING RESPONSE). If the mobile station's SRES equals to the SRES calculated (or maintained) by the SGSN, the user is authenticated and is allowed to use the network.

11.7.2 Ciphering

The ciphering functionality is performed in the LLC layer between MS and SGSN (see Figs. 11.7 and 11.11). Thus, the ciphering scope reaches from the MS all the way to the SGSN (and vice versa), whereas in conventional GSM the scope is only between MS and BTS/BSC.

As in GSM ciphering, the algorithm A8 generates the *Cipher Key* Kc from the key Ki and a random number RAND (see Figure 6.26). Kc is then used by the *GPRS Encryption Algorithm* (GEA) for data encryption (algorithm A5). Note that the key Kc which is handled by the SGSN is independent of the key Kc handled by the MSC for conventional GSM services. An MS may thus have more than one Kc key.

The MS and the SGSN start ciphering after the message AUTHENTICATION AND CIPHERING RESPONSE is sent or received, respectively. Afterward, GPRS user data and signaling during data transfer are transmitted in an encrypted manner.

11.7.3 Subscriber Identity Confidentiality

As in GSM, the identity of the subscriber is held confidential. This is done by using temporary identities on the radio channel. In particular, the user's IMSI is not transmitted unencrypted, but a *Packet Temporary Mobile Subscriber Identity* (P-TMSI) is assigned to each user by the SGSN. This address is temporary and is only valid and unique in the service area of this SGSN. From the P-TMSI, a *Temporary Logical Link Identity* (TLLI) can be derived. The mapping between these temporary identities and the IMSI is stored only in the MS and in the SGSN.

11.8 Summary

The *General Packet Radio Service* (GPRS) is an important step in the evolution of cellular networks toward third-generation and mobile Internet. Its packet-oriented transmission technology enables efficient and simplified wireless access to IP and X.25 networks.

GPRS extends the existing GSM infrastructure in particular with two network nodes, namely the SGSN and GGSN. In Section 11.1 their tasks and the interworking with GSM nodes and registers (HLR, VLR, and EIR) has been explained.

In first implementations, GPRS offers point-to-point bearer services and transport of SMS messages; in future releases also point-to-multipoint services will be offered. An important feature of GPRS is its QoS support. An individual QoS profile (service precedence, reliability, delay, and throughput) can be negotiated for each PDP context. For the simultaneous usage of GPRS and conventional GSM services, three classes of mobile stations are defined in the standard.

Before a GPRS mobile station can use GPRS services it must obtain an address used in the external packet data network (e.g. an IP address) and create a PDP context. This context describes the essential characteristics of the session (PDP type, PDP address, QoS, and GGSN). In order to support a large number of mobile users, it is essential to use dynamic address allocation, e.g., using DHCP for dynamic IP address assignment.

Once an MS has an active PDP context, packets addressed from the external packet data network to the MS will be routed to the responsible GGSN. The GGSN then tunnels them to the current SGSN of the mobile user, which finally forwards the data to the MS. The GPRS location management is based on the definition of an MS state model. Depending on the state of the MS (READY, STANDBY, or IDLE), it performs many or only few location updates. For this purpose, special routing areas are defined, which are sub-areas of the location areas defined in GSM. Although GPRS has its own mobility management, it cooperates with the GSM mobility management. This results, for example, in a more efficient paging mechanism for mobile stations that use circuit- and packet-based services simultaneously.

In Section 11.4 we showed the protocol architecture of the GPRS transmission and signaling plane. GPRS-specific protocols include the *GPRS Tunneling Protocol* (GTP), the *GPRS Mobility Management and Session Management* (GMM/SM) protocol, and the *Subnetwork Dependent Convergence Protocol* (SNDCP). Some GSM protocols, such as the *Mobile Application Part*, have been extended for use with GPRS.

The packet-oriented air interface is one of the key aspects of GPRS. Mobile stations with multislot capability can transmit on several time slots of a TDMA frame, up- and downlink are allocated separately, and physical channels are only assigned for the duration of the transmission, which leads to a statistical multiplex gain. This flexibility in the channel allocation results in a more efficient utilization of the radio resources. On top of the physical channels, a number of logical packet channels have been standardized. The traffic channel PDTCH is used for payload transmission. The GPRS signaling channels are used e.g. for broadcast of system information (PBCCH), access control (PRACH, PAGCH), paging (PPCH), and notification of incoming PTM messages (PNCH). Once more, the coordination between GPRS and GSM channels saves radio resources.

GPRS channel coding defines four different coding schemes, which allow to adjust the tradeoff between the level of error protection and data rate, depending on the current radio channel quality. GPRS security principles include authentication, ciphering, and subscriber identity confidentiality. The SGSN handles authentication, and a special *GPRS Encryption Algorithm* (GEA) has been defined. Moreover, GPRS operators protect their

network with firewalls to external networks and border gateways to other GPRS networks. IP security protocols (IPsec) may be used to communicate over insecure external IP networks.

Typical scenarios for GPRS are the wireless access to the Internet, e-mail communication, *Wireless Application Protocol* (WAP) over GPRS, and applications in the telemetry field. Users can access the Internet without first requiring to dial in to an Internet Service Provider. In particular, mobile e-commerce and location-based services (e.g. tourist guides) will become more important in the future. Main advantages for the users are the higher data rates and volume-based billing. The latter allows them to stay online for a long time.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

12 GSM – The Story Goes On

12.1 Globalisation

GSM is now in more countries than McDonalds.
Mike Short, Chairman MoU Association 1995–1996

GSM was initially designed as a pan-European mobile communication network, but shortly after the successful start of the first commercial networks in Europe, GSM systems were also deployed on other continents (e.g. in Australia, Hong Kong, and New Zealand). In the meantime, 373 networks in 142 countries are in operation (see Section 1.3).

In addition to GSM networks that operate in the 900 MHz frequency band, so-called *Personal Communication Networks* (PCN) and *Personal Communication Systems* (PCS) are in operation. They are using new frequencies around 1800 MHz, and in North America around 1900 MHz. Apart from the peculiarities that result from the different frequency range, PCN/PCS networks are full GSM networks without any restrictions, in particular with respect to services and signaling protocols. International roaming among these networks is possible based on the standardized interface between mobile equipment and the SIM card, which enables personalization of equipment operating in different frequency ranges (*SIM card roaming*). Furthermore, a more general standardization of the SIM concept could allow worldwide roaming across non-GSM networks.

Besides roaming based on the SIM card, the MoU has put increasing emphasis on multiband systems and multiband terminals during the last years (dualband, triband). Multiband systems permit the simultaneous operation of base stations with different frequency ranges. In connection with multiband terminals, this approach leads to a powerful concept. Such terminals can be operated in several frequency bands, and they can adapt automatically to the frequencies used in the network at hand. This enables roaming among networks with different frequency ranges, but also automatic cell selection in multiband networks with different frequencies becomes possible.

12.2 Overview of GSM Services in Phase 2+

GSM is not a closed system that does not undergo any change. The GSM standards are being enhanced; and in the current phase of standardization (Phase 2+) several individual topics are being discussed. Phase 1 of the GSM implementation contained basic teleservices – in the first place voice communication – and a few supplementary services, which had to be offered by all network operators in 1991 when GSM was introduced into the market. The standardization of Phase 2 was completed in 1995 with market introduction following in 1996. Essentially, ETSI added more of the supplementary services, which had been planned already when GSM was initially conceived and which were adopted from the fixed ISDN (see Section 4.3). These new services made it necessary to rework large parts of the GSM standards. For this reason, networks operating according to the revised standard are also called GSM Phase 2 [45]. However, all networks and terminals of Phase 2 preserve the compatibility with the old terminals and network equipment of Phase 1, i.e. all new standard development had to be strictly backward compatible.

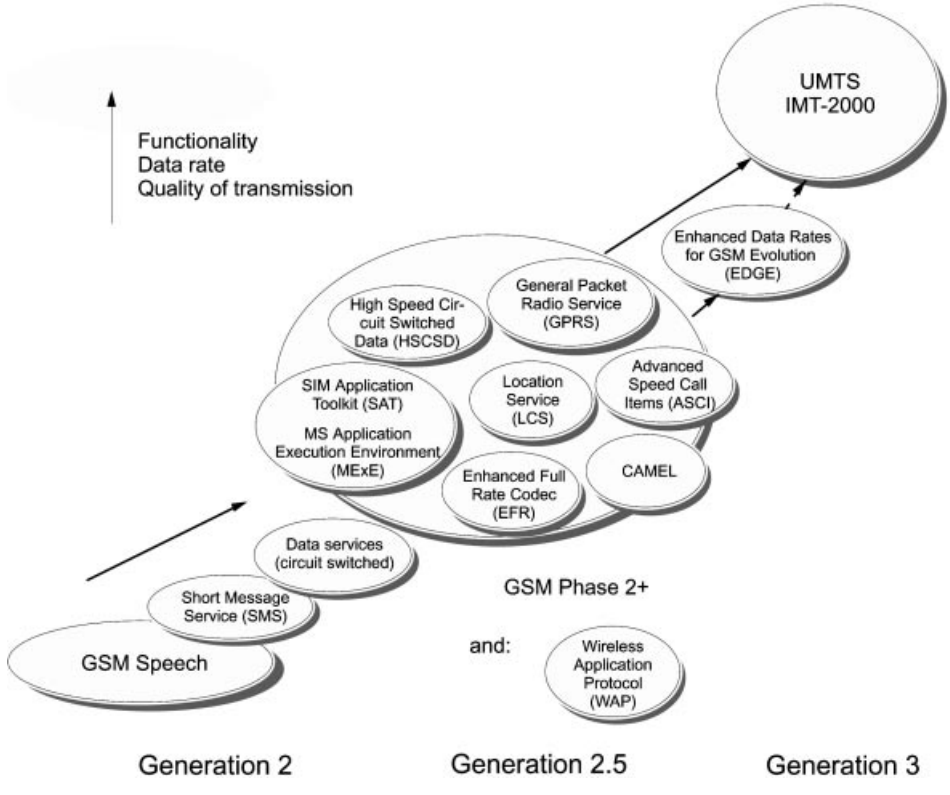

Figure 12.1: Evolution of GSM

The topics of Phase 2+ deal with many aspects ranging from radio transmission to communication and call processing. However, there is no complete revision of the GSM standard; rather single subject areas are treated as separate standardization units, with the intent of allowing them to be implemented and introduced independently from each other.

Thus GSM systems can evolve gradually, and standardization can meet market needs in a flexible way. However, with this approach, a unique identification of a GSM standard version becomes impossible. The designation GSM Phase 2+ is supposed to indicate this openness [45], suggesting an evolutionary process with no endpoint in time or prescribed target dates for the introduction of new services. The GSM standards are now published in so-called releases (e.g. Release 97, 98, 99, and 2000).

A large menu of technical questions is being addressed, only a few of which are presented as examples in the following. Figure 12.1 illustrates the evolution of GSM, from the initial digital speech services toward the 3rd generation of mobile communications (UMTS/IMT-2000). In particular, it shows the services of Phase 2+ that are covered in this book. Most of these services are already offered by GSM network providers today and can be used with enhanced mobile equipment. Some other services are in the planning stage at the time of this writing.

12.3 Bearer and Teleservices of GSM Phase 2+

Whereas GSM Phase 2 defined essentially a set of new supplementary services, Phase 2+ is also addressing new bearer and teleservices. In this section we give an overview of these new speech and data services. They significantly improve the GSM speech quality and make the utilization of available radio resources much more efficient. Furthermore, the new data services are an important step toward wireless Internet access via cellular networks.

12.3.1 Improved Codecs for Speech Services: Half-Rate Codec, EFR Codec, and AMR Codec

One of the most important services in GSM is (of course) voice service. Thus it is obvious, that voice service has to be further improved. In first place is the development of new speech codecs with two competing objectives:

- better utilization of the frequency bands assigned to GSM and
- improvement of speech quality in the direction of the quality offered by ISDN networks, which is primarily requested by professional users.

Half-Rate codec – The reason for improved bandwidth utilization is to increase the network capacity and the spectral efficiency (i.e. traffic carried per cell area and frequency band). Early plans were already in place to introduce a *half-rate speech codec*. Under good channel conditions, this codec achieves, in spite of the half bit rate, almost the same speech quality as the full-rate codec used so far. However, quality loss occurs in particular for mobile-to-mobile communication, since in this case (due to the ISDN architecture) one has to go twice through the GSM speech coding/decoding process. These multiple, or tandem, conversions degrade speech quality. The end-to-end transmission of GSM-coded speech is intended to avoid multiple unnecessary transcoding and the resulting quality loss (Figure 12.2) [45]. This technique has been passed under the name *Tandem Free Operation* (TFO) in GSM Release 98.

Enhanced Full-Rate (EFR) codec – A very important concern is the improvement of

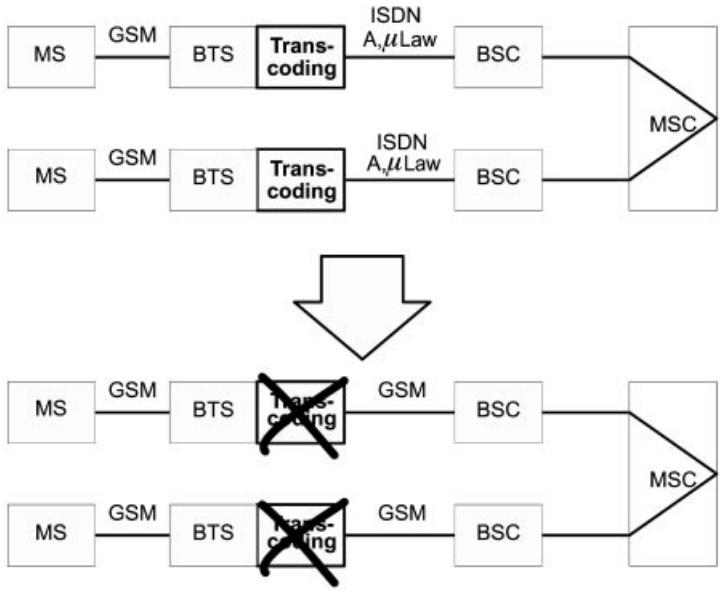

Figure 12.2: Through-transport of GSM-coded speech in Phase 2+ for mobile-to-mobile connections (*tandem free operation*)

speech quality. Speech quality that is close to the one found in fixed networks is especially important for business applications and in cases where GSM systems are intended to replace fixed networks, e.g. for fast installation of telecommunication networks in areas with insufficient or missing telephone infrastructure.

Work on the *Enhanced Full-Rate* (EFR) codec was therefore considered of high priority. This EFR is a full-rate codec (net bit rate 12.2 kbit/s). Nevertheless, it achieves speech quality clearly superior to the previously used full-rate codec. It has been initially standardized and used in North American DCS1900 networks [45] and has been implemented in GSM with very good success. Instead of using the *Regular Pulse Excitation–Long Term Prediction* (RPE-LTP) coding scheme (see Section 6.1), a so-called *Algebraic Code Excitation–Linear Prediction* (ACELP) is employed.

The EFR speech coder delivers data blocks of 244 information bits to the channel encoder (compare with Table 6.2). In addition to grading the bits into important Class I bits and less important Class II bits, EFR further divides into Class Ia bits and Class Ib bits. A special preliminary channel coding is employed for the most significant bits: eight parity bits (generated by a *Cyclic Redundancy Check* (CRC) coding) and eight repetition bits are added to provide additional error-detection. The resulting 260 bits are processed by the block encoder as described in Section 6.2.1.1. For convolutional coding of Class I bits the convolutional encoder defined by the generator polynomials G0 and G1 is employed.

Adaptive Multi-Rate (AMR) codec – The speech codecs mentioned before (full-rate, half-rate, and EFR) all use a fixed source/information bit rate, which has been optimized for typical radio channel conditions. The problem with this approach is its inflexibility: whenever the channel conditions are much worse than usual, very poor speech quality will result, since the channel capacity assigned to the mobile station is too small for error free

transmission. On the other hand, radio resources will be wasted for unneeded error protection if the radio conditions are better than usual.

To overcome these problems, a much more flexible codec has been developed and standardized: the *Adaptive Multi-Rate* (AMR) codec. It can improve speech quality by adaptively switching between different speech coding schemes (with different levels of error protection) according to the current channel quality. To be more precise, AMR has two principles of adaptability [11]: *channel mode adaptation* and *codec mode adaptation.*

Channel mode adaptation dynamically selects the type of traffic channel that a connection should be assigned to: either a full-rate (TCH/F) or a half-rate traffic channel (TCH/H). The basic idea here is to adapt a user's gross bit rate in order to optimize the usage of radio resources. If the traffic load in a cell is high, those connections using a TCH/F (gross bit rate 22.8 kbit/s) and having good channel quality should be switched to a TCH/H (11.4 kbit/s). On the other hand, if the load is low, the speech quality of several TCH/H connections can be improved by switching them to a TCH/F. The signaling information for this type of adaptation is done with existing protocols on GSM signaling channels; the switching between full-rate and half-rate channels is realized by an intracell handover.

The task of codec mode adaptation is to adapt the coding rate (i.e. the trade-off between the level of error protection versus the source bit rate) according to the current channel conditions. When the radio channel is bad, the encoder operates at low source bit rates at its input and uses more bits for forward error protection. When the quality of the channel is good, less error protection is employed.

Table 12.1: AMR codec modes

Source data rate in kbit/s	12.2	10.2	7.95	7.4	6.7	5.9	5.15	4.75
Information bits per block	244	204	159	148	134	118	103	95
– Class Ia bits (CRC-protected)	81	65	75	61	55	55	49	39
– Class Ib bits (not CRC-protected)	163	139	84	87	79	63	54	56
Rate *R* of convolutional encoder	1/2	1/3	1/3	1/3	1/4	1/4	1/5	1/5
Output bits from convolutional encoder	508	642	513	474	576	520	565	535
Punctured bits	60	194	65	26	128	72	117	87

The AMR codec consists of eight different modes with source/information bit rates ranging from 12.2 kbit/s to 4.75 kbit/s (see Table 12.1). All modes are scaled versions of a common ACELP basis codec.

From the results of link quality measures, an adaptation unit selects the most appropriate codec mode. Figure 12.3 illustrates the AMR encoding principle. Channel coding is performed using a punctured recursive systematic convolutional code. Since not all bits of the voice data are equally important for audibility, AMR also employs an *Unequal Error Protection* (UEP) structure. The most important bits (Class Ia; e.g. mode bits and LPC

coefficients) are additionally protected by a *Cyclic Redundancy Check* (CRC) code with 6 parity bits. On the receiver side, the decoder will discard the entire speech frame if the parity check fails. Also the degree of puncturing depends on the importance of the bits. At the end of the encoding process, a block with a fixed number of gross bits results, which is subsequently interleaved to reduce the number of burst errors.

Since the channel conditions can change rapidly, codec mode adaptation requires a fast signaling mechanism. This is achieved by transmitting the information about the used codec mode, link control, and DTX, etc. together with the speech data in the TCH, i.e. a special inband signaling is employed.

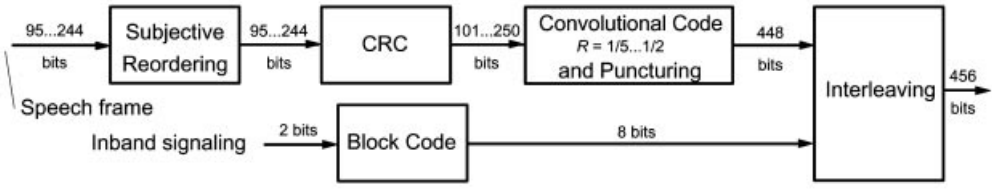

Figure 12.3: AMR channel encoding principle (bit numbers for TCH/F)

We give an example: the 12.2 kbit/s codec for a TCH/F operates with 244 source bits (12.2 kbit/s × 20 ms), which are first rearranged to subjective importance. By adding six CRC bits for Class 1a bits, we obtain 250 bits. The subsequent recursive convolutional encoder, defined by the two generators 1 and $G_1/G_0 = (d^4 + d^3 + d + 1)/(d^4 + d^3 + d)$, with rate $R \approx 1/2$, maps those bits to 508 bits. Next, 60 bits are punctured, which results in an output sequence of 448 bits. Together with the encoded inband signaling (8 bits) this block is interleaved and finally mapped to bursts. The resulting gross bit rate is thus 456 bits/20 ms = 22.8 kbit/s.

12.3.2 Advanced Speech Call Items (ASCI)

GSM systems of Phase 2 offer inadequate features for group communications. For example, group call or "push-to-talk" services with fast connection setup as known from private radio or digital trunked radio systems (e.g. TETRA), are not offered. However, such services are indispensable for most closed user groups (e.g. police, airport staff, railroad or taxi companies). In particular railroad operators had a strong request for such features. In 1992, their international organization, the *Union Internationale des Chemins de Fer* (UIC), selected the GSM system as their standard [45]. This GSM-based uniform international railway communication system should replace a multitude of incompatible radio systems.

In this section we describe the standardized speech teleservices that offer functionality for group communication: the *Voice Broadcast Service* (VBS) and the *Voice Group Call Service* (VGCS). In addition, the *Enhanced Multi-Level Precedence and Pre-emption Service* (eMLPP) is used to assign and control priorities to users and their calls (e.g. for emergency calls). All those services together are referred to as *Advanced Speech Call Items* (ASCI).

12.3.2.1 Voice Broadcast Service (VBS)

The *Voice Broadcast Service* (VBS) allows a user to broadcast a speech message to several other users within a certain geographical area. The user who initiates the call can only send (“speaker”), and all others can only listen (“listeners”).

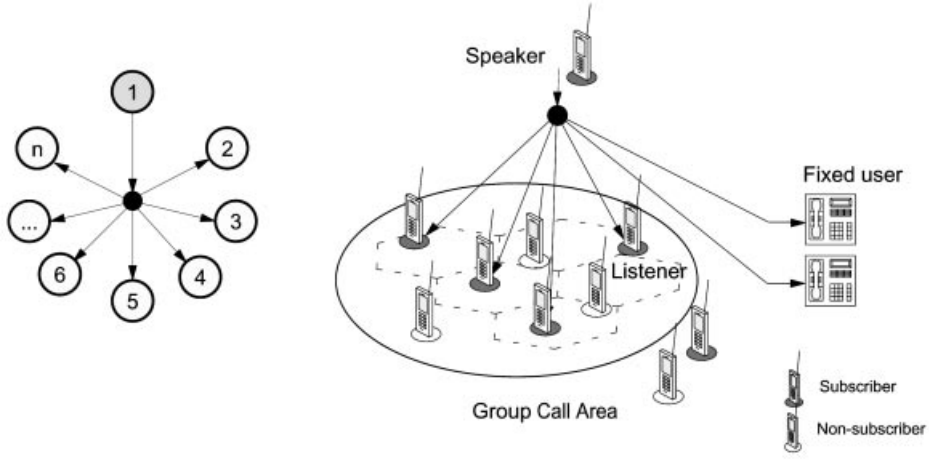

Figure 12.4: VBS scenario (schematic illustration)

Figure 12.4 gives a schematic illustration of a VBS scenario. Mobile users who are interested in a certain VBS group subscribe it and will then receive broadcast calls of this group. A special permission is needed, however, for the right to send broadcast calls, i.e. for the right to act as a speaker. The subscribed VBS groups are stored on the user’s SIM card, and if a subscriber does not want to receive VBS calls for a certain time, he or she can deactivate them. Besides mobile GSM users, also a predefined group of fixed telephone connections can participate in the VBS service (e.g. dispatchers, supervisors, operators, or recording machines).

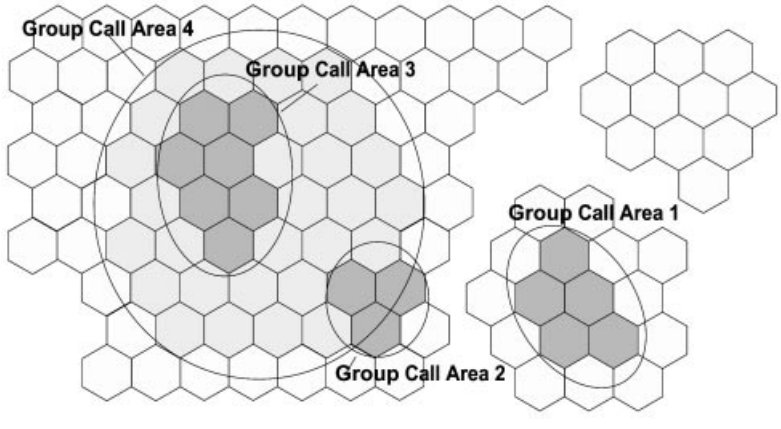

Figure 12.5: Some examples of group call areas

System Concept and Group Call Register – The area in which a speech broadcast call is offered is referred to as *group call area*. As illustrated in Figure 12.5, in general, this area consists of several cells. A group call area may comprise cells of several MSC areas and even of several PLMNs. One MSC is responsible for the handling of the VBS. It is called *Anchor MSC*. In case a voice broadcast should also be transmitted in cells that are not within the service area of this MSC (i.e. if the group call area contains also cells belonging to other MSCs), the MSCs of those cells are also involved. They are then denoted as *Relay MSCs*.

The VBS-specific data is stored in a *Group Call Register* (GCR). Figure 12.6 shows the extended GSM system architecture. The GCR contains the *broadcast call attributes* for each VBS group, which are needed for call forwarding and authentication. For example:

- Which cells belong to the group call area?
- Which MSC is the responsible anchor MSC?
- In which cells are group members currently located, i.e. in which cells is a voice message to be broadcast?
- To which other MSCs is the voice message to be forwarded to reach all group members who are currently located in the group call area?
- To which external fixed telephone connections is the broadcast message addressed?
- Which fixed telephone connections are allowed to act as speakers?

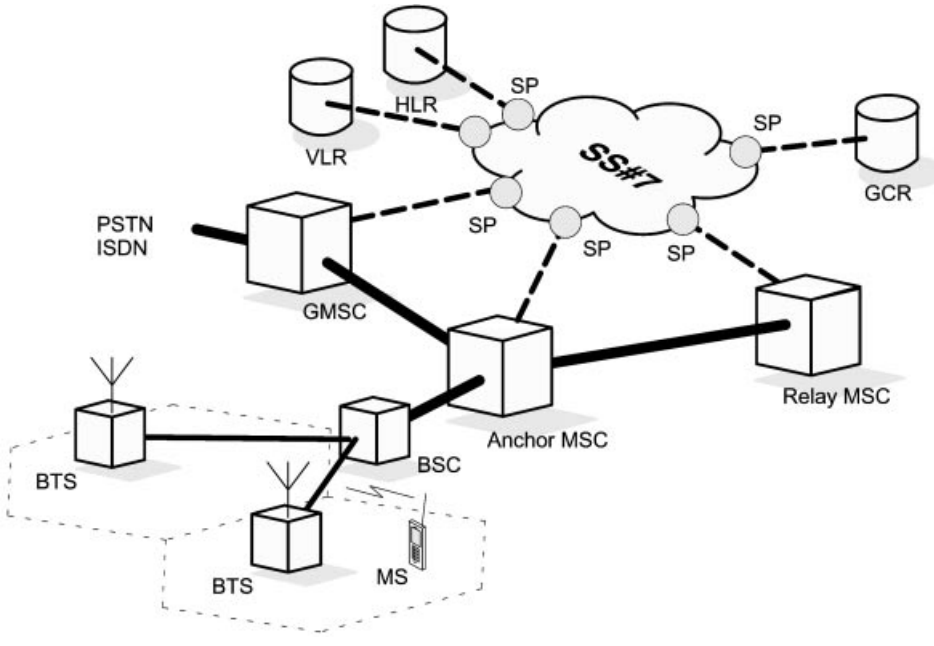

Figure 12.6: Extension of the GSM system architecture with the GCR

Call Establishment and Logical Channels – A mobile station that intends to initiate a voice broadcast call sends a service request to the BSS. The request contains the *Group ID* of the VBS group to be called. Thereupon, the responsible MSC queries the user's profile

from the VLR and checks whether the user is allowed to act as speaker for the stated group. Afterward, some VBS-specific attributes are requested from the GCR. If the broadcast call should also be transmitted in cells that do not belong to the current MSC, an anchor MSC is determined. The anchor MSC then forwards the VBS attributes to all relay MSCs, which then request all affected BSCs to allocate a traffic channel in the respective cells, and to send out notification messages on the NCH (see Section 5.1). When a mobile station receives such a message and it is also subscribing to the respective VBS group, it changes to the given traffic channel and listens to the voice broadcast in the downlink. The speaker is then informed about the successful connection setup and can start talking. The notification message is periodically repeated on the NCH until the speaker terminates the call.

In contrast to the paging procedure in conventional GSM calls, the individual mobile users and their mobile stations are not explicitly addressed by an IMSI or TMSI but with the Group ID of the VBS group. Furthermore, the mobile stations do not acknowledge the reception of VBS calls to the network. To realize the service, traffic channels are not allocated to individual subscribers, but the voice signal of the speaker is broadcast to all listening participants in a cell on one group channel. Thus, in each participating cell, only one full-rate channel is occupied (as in regular voice calls).

12.3.2.2 Voice Group Call Service (VGCS)

Another group communication service is the *Voice Group Call Service* (VGCS). The VGCS defines a closed user group communication service, where the right to talk can now be passed along within the group during a call by using a push-to-talk mechanism as in mobile radio. This principle is illustrated in Figure 12.7: User 1 initializes a group call and speaks, while the other users listen. Afterward, User 1 releases the channel and changes into listener mode. Now, each of the subscribers may apply for the right to become speaker. For example, User 4 requests the channel, and the network assigns it to him/her. He or she talks, releases the channel, and changes back to listener mode. Finally, the group call is terminated by the initiator (in general). Whereas the information flow in the VBS is simplex, the VGCS can be regarded as a half-duplex system (compare Figures 12.4 and 12.7).

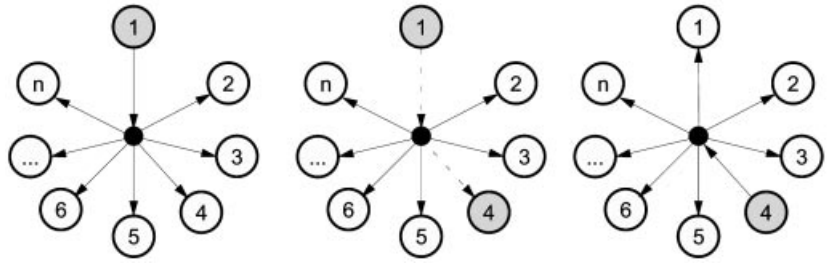

Figure 12.7: Group call scenario (schematic illustration)

The fundamental concepts and entities of the VBS, e.g. the definition of group call areas, group IDs, the GCR, and anchor and relay MSCs are also used in the VGCS.

Logical Channels – A traffic channel is allocated in each cell of the group call area that is involved in the VGCS. All group members listen to this channel in the downlink, and only the speaker uses it in the uplink. Therefore, in addition to the tasks for VBS calls, the network must also control uplink radio resources. The network indicates in the downlink to all mobile stations whether the uplink channel is in use or not. If the channel is free, the group members may send access bursts. Collisions that occur with simultaneous requests are resolved, and the network chooses one user who obtains the channel and thus has the right to talk.

12.3.2.3 Enhanced Multi-Level Precedence and Pre-emption (eMLPP)

Priority services enable a network to process calls with a priority class (*precedence level*). If the network load is high, calls with high priority can then be treated in a preferred manner, and resources for low priority calls can be deallocated. In the extreme case, a call with low priority can be dropped because a call with high priority arrives (*pre-emption*).

Table 12.2: Priority classes in eMLPP

Class	Used by	Connection setup	Call interruption (pre-emption)	Example
A	Operator	Fast (1–2 s)	Yes	Highest priority; VBS/ VGCS emergency calls
B	Operator	Normal (<5 s)	Yes	Calls of operator
0	Subscriber	Normal (<5 s)	Yes	Emergency calls of users
1	Subscriber	Slow (<10 s)	Yes	
2	Subscriber	Slow (<10 s)	No	
3	Subscriber	Slow (<10 s)	No	Standard priority
4	Subscriber	Slow (<10 s)	No	Lowest priority

The control of priorities in GSM is called *Enhanced Multi-Level Precedence and Pre-emption* (eMLPP). It is a supplementary service for point-to-point speech services as well as for VBS and VGCS. The principle of eMLPP is based on the *Multi-Level Precedence and Pre-emption* (MLPP) [33] method used in SS#7. In doing so, MLPP has been enhanced with functions for priority control at the air interface. Table 12.2 lists all priority classes of eMLPP. Besides the five precedence levels that are used in MLPP (Classes 0–4), two additional levels with higher priority are defined (Classes A and B). The table also shows whether a call with higher priority may terminate a call with lower priority. It is important to note that only the operator may use calls of Class A and B, such that for example an emergency call over VBS or VGCS can be initiated in disaster situations. Calls of this class can only be employed within the service area of one MSC. The other five classes can be utilized within the entire PLMN and also in combination with the MLPP of ISDN. The highest priority call that a subscriber is allowed to use is stored on his or her SIM card and in the HLR.

12.3.3 New Data Services and Higher Data Rates: HSCSD, GPRS, and EDGE

Development also continues with data services. The maximal data rate of 9600 bit/s for data services in conventional GSM is rather low compared to fixed networks. The desire for higher data rates in GSM networks is therefore quite obvious. Two prominent trends can be recognized: integration of packet services into GSM networks and high-bit-rate bearer services with data transmission rates up to some 10 kbit/s.

Accordingly, one of the GSM standardization groups specified the *High Speed Circuit Switched Data* (HSCSD) service. By combining several traffic channels, data rates of up to 60 kbit/s are achieved. Whereas this is relatively easy to implement at the base stations, the changes required on the terminal side are substantial. An HSCSD-capable terminal must be able to transmit and receive simultaneously on several time slots (*multislot operation*), and it must also supply the considerable signal processing power for modulation/demodulation and channel coding. This is why a new generation of mobile stations with significantly increased capabilities was required for HSCSD usage. Since 1999 some network operators have been offering HSCSD.

The newly defined packet data service, *General Packet Radio Service* (GPRS), finds great interest among network operators. It offers a genuine packet switched bearer service at the air interface. Its first phase of standardization was completed in Release 97 and is stable. During the year 2000, several operators upgraded their network with GPRS. As for HSCSD, new multislot-capable mobile stations are required (which can use, e.g. 4 time slots in the downlink and 2 time slots in the uplink). The GPRS chapter of this book discusses in detail the system architecture, protocols, air interface, multiple access, and interworking with the Internet. Additional information can be found in [5,10,20,26,37,60]. Release 99 extended the GPRS standard with some new functions, e.g. point-to-multipoint services and prepaid services. Furthermore, existing functionality has been improved.

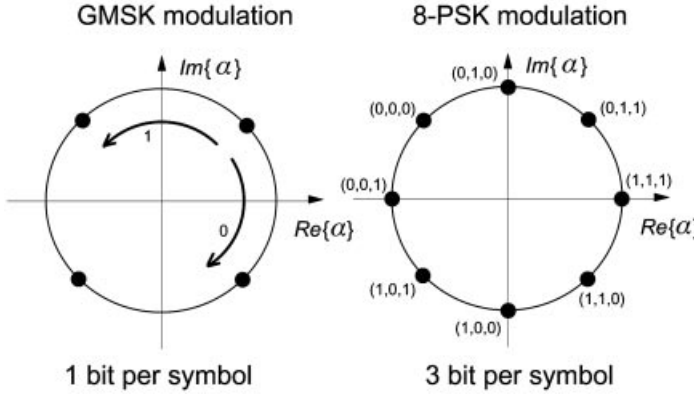

Figure 12.8: Symbol space constellations for GMSK and 8-PSK

While HSCSD and GPRS achieve higher data rates because a mobile station can use several time slots of the same TDMA frame and because new coding schemes are employed, the planned *Enhanced Data Rates for GSM Evolution* (EDGE) system goes

even one step further. EDGE replaces the GMSK modulation scheme used in GSM with an 8-PSK (*8-Phase Shift Keying*) scheme, so that it achieves an approximately three times higher data rate per time slot and a higher spectral efficiency. Using GMSK, one data bit d_i on average is mapped to one symbol a_i (see Section 5.2.1); with 8-PSK three data bits d_i are combined to one symbol a_i and transmitted together. Figure 12.8 shows the symbol constellations in the complex plane and the associated bit sequences. As opposed to GMSK, 8-PSK does not have a constant envelope and therefore puts higher requirements on new transceivers in BTSs and MSs.

Furthermore, a *link adaptation* technique is employed, which dynamically chooses a modulation and coding scheme according to the current radio channel conditions. EDGE exists in two variants for GSM: *Enhanced Circuit Switched Data* (ECSD) for circuit switched services such as HSCSD, and *Enhanced GPRS* (EGPRS). More detailed information can be found in [22,47].

It is interesting to note that both GPRS and EDGE are also being standardized for the North American cellular network TDMA-136 (GPRS-136 and GPRS-136HS EDGE).

12.4 Supplementary Services in GSM Phase 2+

12.4.1 Supplementary Services for Speech

By far the largest part of the supplementary service characteristics known from ISDN have in one way or another already been implemented in GSM (see Section 4.3). The mobility of the users, however, creates the need for new supplementary services. Examples of supplementary services known from ISDN or newly defined are *mobile access hunting*, *short message forwarding*, *multiple subscriber profile*, *call transfer*, or *Completion of Calls to Busy Subscribers* (CCBS).

The example of the CCBS service shows especially clearly how much the role of the HLR is changing from its original function as a database to a more active role as a service control component, similar to the *Service Control Point* (SCP) of the *Intelligent Network* (IN). The supplementary service CCBS basically realizes "call back if busy." If a called subscriber does momentarily not accept a call due to an ongoing connection, the calling subscriber can activate the supplementary service CCBS which causes the network to notify him at the end of the called subscriber's ongoing call and automatically set up the new connection. The subscriber mobility adds more complexity to the implementation of this service. In the fixed network, implementation would require the establishment of queues for call-back requests in the switching center of the calling and called subscribers, respectively. In a mobile network, this may involve additional switches, because after activation of the CCBS service, the calling subscriber may be roaming into another switching center area. If the implementation of the service were only in the MSC, either there could be a centralized solution, or the queuing lists would have to be forwarded to the new MSC – which may even be in another network. The targeted solution is centralized in the HLR, which has to store the subscriber's callback queues (if existing) in addition to the current MSC designation. If the mobile station changes the MSC area, the callback queue is transferred to the new MSC. In this case, therefore, the HLR has to assume an additional

server role and perform call control beyond the originally planned restriction to a pure database function.

12.4.2 Location Service (LCS)

GSM Release 99 introduces a *Location Service* (LCS) making it possible to determine the exact location of a mobile station down to a few meters. One of the motivations for this service has been a law in the USA which demands to locate a person in case of an emergency call.

From GSM mobility management, the network already knows the current cell of the user (*cell identifier*). However, this location accuracy is not sufficient in most cases, and therefore investigations have been made to find a more sophisticated solution. In the so-called *Time of Arrival* (TOA) method, the network listens to handover access bursts of the mobile station and is then able to triangulate its position. In contrast, using the *Enhanced Observed Time Difference* (E-OTD) method, now the mobile stations measure the time difference of received bursts from different base stations. Both methods only work if a mobile station has contact to at least three base stations. The accuracy of E-OTD schemes lies between 50 and 125 m, and the one of TOA is worse [12]. E-OTD schemes require a software update on the mobile equipment as well as modifications in the network, whereas for the TOA method it is mainly sufficient to modify network components. However, the functionality of TOA is provided by synchronization of the cellular network (using *Global Positioning System* (GPS) or precise clocks at each BTS). This capability is currently not provided in asynchronous GSM networks. The most precise way to find out the position of a mobile station is to integrate a GPS receiver into each piece of mobile equipment. The mobile station then receives its current position from GPS satellites. A substantial disadvantage of this approach is that mobile stations cannot always have intervisibility with GPS satellites (e.g. inside buildings). We observe that each of the three methods has its advantages and disadvantages.

In addition to the technical implementation of the location service, two new network nodes have been defined for this type of service: the *Gateway Mobile Location Centre* (GMLC) and the *Serving Mobile Location Centre* (SMLC). The GMLC acts as an interface to applications that use the positioning information of users in a specific way – so-called *location-based* or *location-aware* services. Examples are navigation services (such as "Where is the closest gas station?") or virtual tourist guides ("What is the building on my left side?"). A service provider stores e.g. the locations of gas stations and sightseeing attractions in a database and adds other useful information. At the request of a mobile user, the provider can get the current position of the user from the GMLC and send back the requested information. Other location-based applications include location-based charging ("home zone"), vehicle tracking (e.g. stolen cars), and localized news, weather, and traffic information.

12.5 Service Platforms

The procedures for the development of the GSM standards required close cooperation of

the involved manufacturers and network operators. The international standardization of services and interfaces led to a set of common successful performance characteristics in GSM networks, most prominently the international roaming capability. The more a performance criterion is standardized, the lower are the costs of development and introduction, since all manufacturers and operators contribute to paying the costs. On the other hand, the network operators desire service differentiation to be able to gain competitive advantages. The standardization of services and service performance criteria reduces the possibility for differentiation among competitors. Moreover, the time-to-market is often pushed out because of the prolonged process of standardization.

For these reasons, the service platform concept has been introduced in GSM on both the network and the terminal side. These platforms offer mechanisms, functions, and protocols for definition and control of services and applications. Those services/applications can be operator-specific, such that an international standardization process is not needed in general. The required generic functions can be made available in each mobile station and network node, and they can be used and combined in a flexible way for service execution.

The GSM supplementary services can be regarded as the simplest form of service platform usage. An extended concept are the so-called *service nodes*, such as a voice mail server and an SMS service center. However, both concepts have significant disadvantages: supplementary services are subject to international standardization, and on the other hand these services might not be available to roaming users in foreign networks, since network providers are not obligated to implement all supplementary services. The situation is similar with service nodes, which are often accessible only in the home network. We conclude that these two types of platforms allow the definition of vertical/operator-specific services only in a limited way, and their usage in foreign networks is often not possible or rather complicated.

An extension of the platform concept, which has been taken up by ETSI in the Phase 2+ standardization, attempts to overcome this dilemma. Instead of specifying services and supplementary services directly or completely, only mechanisms are standardized which enable introduction of new services. With this approach it is possible to restrict the implementation of a service to a few switches in the home network of a subscriber, whereas local (visited) switches have to provide only a fixed set of basic functions and the capability to communicate with the home network switch containing service logic.

This group of GSM standards within Phase 2+ is known under the name *Support of Operator-Specific Services* (SOSS), or also as *Customized Applications for Mobile Network Enhanced Logic* (CAMEL). The answer on the terminal side is the *SIM Application Toolkit* (SAT) and *Mobile Station Execution Environment* (MExE). They are explained in the following.

12.5.1 CAMEL – GSM and Intelligent Networks

Essentially, CAMEL represents a convergence of GSM and *Intelligent Network* (IN) technologies. The fundamental concept of IN is to enable flexible implementation, introduction and control of services in public networks and to use the idea of dividing the

switching functionality into basic switching functionality, residing in *Service Switching Points* (SSPs) and centralized service control functionality, residing in *Service Control Points* (SCPs). Both network components communicate with each other over the signaling network using the generic SS#7 protocol extension called *Intelligent Network Application Part* (INAP). This approach enables a centralized, flexible, and rapid introduction of new services [2].

There are already some features in GSM which parallel an intelligent network. Even though GSM standards use neither IN terminology nor IN protocols, i.e. INAP, the GSM network structure follows the IN philosophy [41]. In the GSM architecture, the separation into functional units like MSC and HLR and the consistent use of SS#7 and its MAP extensions are in conformity with the IN architecture, which is split into SSPs and SCPs that communicate using INAP.

The philosophy of CAMEL is to proceed with the implementation of services in GSM in a similar way as in IN. This is reflected in separating a set of basic call processing functions in the MSC or GMSC (which act as SSP), from the intelligent service control functions (SCP) in the home network of the respective subscriber. The HLR in a GSM network already has functions similar to the SCP, especially with regard to supplementary services. Beyond that, the CAMEL approach provides its own dedicated SCPs. Imagine specialized SCPs for the translation of abbreviated numbers in *Virtual Private Networks* (VPN) or for future extended *Short Message Services* (SMSs). With this configuration, the service implementation with its service logic is needed only once, namely in the home network SCP. The network operator offering the service thus has the sole control over the features and performance range of the service. Because of the complete range of generic functions that have to be provided at each SSP (MSC, GMSC, etc.), new services can immediately be provided in each network, and an uninterrupted service availability is guaranteed for roaming subscribers. The sole responsibility and control for the introduction of new services lies in SOSS/CAMEL with the operator of the home network, the contract partner of the subscriber. This opens new competitive possibilities among network operators. Operator-specific services can be introduced rapidly without having to go through the standardization process, and yet they are available worldwide.

Figure 12.9 shows the resulting architecture. The CAMEL specification requires a GSM-specific version of IN. Similar to the IN approach, GSM defines a basic call processing function as *GSM Service Switching Function* (gsmSSF) and a service logic function *GSM Service Control Function* (gsmSCF). In addition to the MAP signaling interfaces already existing in Phase 1 and Phase 2 for communication between visited and home networks (GMSC, VLR, HLR), new signaling interfaces are needed for communication between basic switching and service logic in the visited and home networks. For this signaling, a new application part of SS#7 is being specified, the *CAMEL Application Part* (CAP), which assumes the functions similar to INAP. These functions and protocols represent the basic structure for the realization of intelligent services and their flexible introduction.

The prerequisites for CAMEL are the definition of a standardized extended call model with appropriate trigger points, and the specification of the generic range of services which must be provided by the SSP. More precisely, the new extended call model must also include a model of subscriber behavior, because besides normal call processing aspects, it also contains events like *location updating*. For each subscriber, this behavior model is stored

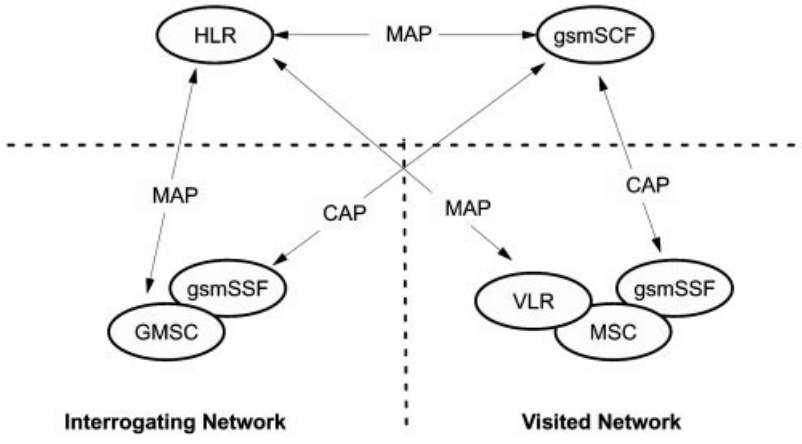

Figure 12.9: Functional architecture for CAMEL

in the HLR and supplied from the home network to the currently visited SSP/MSC. In this way, the SSP/MSC has a set of trigger points with corresponding SCP addresses for each subscriber roaming in its area. When a trigger condition is satisfied, the call and transaction processing in the SSP/MSC is interrupted, and the SCP is notified. The SCP can now analyze the context and, according to the service implementation, give instructions to the SSP to perform particular functions. Typical functions the SSP has to implement are call forwarding, call termination, or other stimuli to the subscriber [45]. Based on this behavior model and the corresponding control protocol between mobile SSP and home SCP, which are connected through a set of generic SSP functions, we can expect to see a large variety of operator-specific services in the future.

12.5.2 Service Platforms on the Terminal Side

12.5.2.1 SIM Application Toolkit (SAT)

The *SIM Application Toolkit* (SAT) has been a further step toward provider-specific vertical services. The GSM SIM card is provided completely by the network operator, in particular because it contains security functions. From this fact, the basic approach arose to equip the SIM card with additional, operator-specific functions. Without a standardized interface to the mobile equipment, this was only possible in a very limited manner and only in close cooperation with equipment manufactures. The SIM Application Toolkit removes these restrictions by defining a standardized interface between mobile equipment and SIM card. In this way, operator-specific applications can run on the SIM card and can thereby control clearly defined, selected functions of the terminal. Corresponding applications can be carried out in the PLMN or even outside the PLMN on dedicated servers making it possible to implement completely new services. The communication between the SIM card application and its counterpart in the network is currently implemented over SMS, but in the near future other bearer services (in particular GPRS) are also possible. The functions defined in the SAT framework can be categorized into *SIM data download* and

proactive SIM. The functional interface between SIM card and terminal is done with proactive SIM mechanisms. They include:

- display of text
- transmission of SMS messages
- connection setup (speech and data) triggered by the SIM card
- playing of sounds in the mobile equipment
- read-out of local information from the equipment into the SIM card.

With these mechanisms, a broad variety of new features can be offered, for example download of data to the SIM card. This includes the download of new or existing commands and applications to be installed. With the toolkit, the SIM card is able to display new, operator-specific menu options to the user, and to read out user actions from the mobile equipment. Most far-reaching are the functions for call control, where each number typed in can be analyzed by the SIM card. This allows for operator-specific treatment of telephone numbers, e.g. the mapping of numbers or barring functions. In a further standardization step, the functions of SAT have been enhanced with security and encryption mechanisms. SAT-capable mobile stations have been available for a few years.

12.5.2.2 Mobile Station Application Execution Environment (MExE)

Of similar scope is the *Mobile Station Execution Environment* (MExE), which implements a generic application platform in the terminal. The most important components are a virtual machine for execution of Java code and the *Wireless Application Protocol* (WAP). Both techniques open the door for a variety of new services and applications. With a virtual machine running on the mobile terminal, applications can be uploaded and executed. This demands a high computational effort in the mobile stations. The WAP is explained in the following.

12.6 Wireless Application Protocol (WAP)

> *WAP is a major step in building the wireless Internet, where people on-the-go can access the Internet through their wireless devices to get information such as e-mails, news headlines, stock reports, map directions and sports scores when they need it and where they need.*
> Chuck Parish, Founding Member and Chairman (1998–1999) WAP forum

WAP is regarded as an important step of today's GSM networks toward a "mobile Internet." During the last few years, WAP has been developed and standardized by the WAP Forum [61,63]. This industry consortium was founded by Nokia, Ericsson, Phone.com (formerly Unwired Planet), and Motorola in December 1997 and has several hundred members today.

The philosophy of WAP is to transfer Internet content and other interactive services to mobile stations to make them accessible to mobile users. For this purpose, WAP defines a system architecture, a protocol family, and an application environment for transmission and display of WWW-like pages for mobile devices.

The motivation for the development of WAP were the fundamental restrictions posed by mobile equipment and cellular networks in comparison to PCs and fixed wired networks. These are in particular the limited opportunities for display and input (small displays, number keypad, and no mouse) as well as the limited memory and processing power. Furthermore, a mobile station's power consumption should be as low as possible. On the network side, it is clear that the wireless transmission has less bandwidth, a higher bit error probability, and less stable connections than wired networks.

The protocols and the application environment defined for WAP consider these limitations. The protocols of the WAP architecture are basically a modification, optimization, and enhancement of the *Internet Protocol* (IP) stack used in the World Wide Web for use in mobile and wireless environments. WAP focuses on applications tailored to the capabilities of cellular phones and the needs of mobile users. One can say that WAP "creates an information Web for cellular phones, distinct from the PC-centric Web" [24].

12.6.1 Wireless Markup Language (WML)

With respect to the mentioned requirements, the *Wireless Markup Language* (WML) has been developed. It represents a pendant to the *Hypertext Markup Language* (HTML) used in the World Wide Web. WML is defined as a document type of the meta language *Extensible Markup Language* (XML). It contains some phone-specific tags and requires only a phone keypad for input. For display of monochrome graphics the *Wireless Bitmap* (WBMP) format has been defined.

A microbrowser, which is running on each WAP terminal, interprets the received WML documents and displays their content (text, pictures, links) to the user. Such a microbrowser is also referred to as WML browser and is the pendant to a Web browser used in PCs.

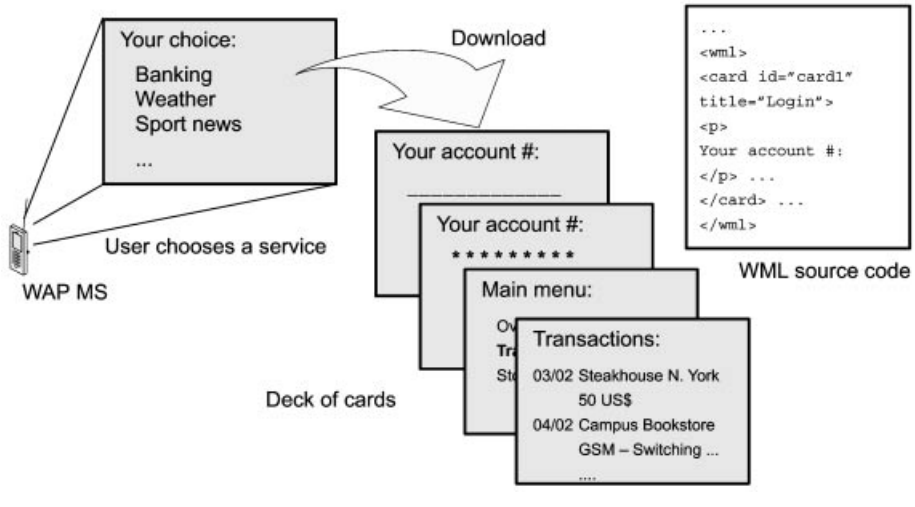

Figure 12.10: WAP example

In doing so, the presentation of WML documents is not limited to classical mobile telephones, but there also exist WML browsers for other devices, such as for *Personal Digital Assistants* (PDAs) under the operating systems PalmOS, Windows CE, or EPOC systems. These devices may be linked over infrared or Bluetooth [6] with a GSM mobile station, or they have their own GSM/GPRS air interface.

WML documents are organized in *cards* and *decks* (see Figure 12.10). When a subscriber chooses a service, a deck of cards is download to the mobile station. The user can then view these cards with his or her WML browser, make inputs, and navigate between the cards. Each card is designed for one user interaction.

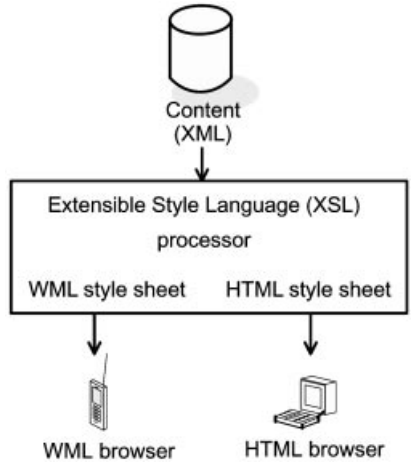

Figure 12.11: Generation of WML and HTML documents

Figure 12.11 illustrates how an automatic parallel creation of WML and HTML documents may look like [62]. The *World Wide Web Consortium* (W3C) currently specifies the *Extensible Style Language* (XSL) [59]. Using XSL style sheets, WML and HTML documents can be automatically generated from content written in XML.

12.6.2 Protocol Architecture

The WAP protocol architecture is shown in Figure 12.12. As mentioned before, WAP is based on the WWW protocol stack and adjusts those protocols to the requirements of wireless transmission and small portable devices.

For applications, a uniform microbrowser environment has been specified: the *Wireless Application Environment* (WAE). It comprises the following functionality and formats:

- the *Wireless Markup Language* (WML),
- a simple script language *WMLScript*, which is based on JavaScript,
- programming interfaces for control of telephony services (*Wireless Telephony Application* (WTA) interface), and
- data formats for pictures, electronic business cards (vCard), and entries of the phone directory and calendar (vCalendar).

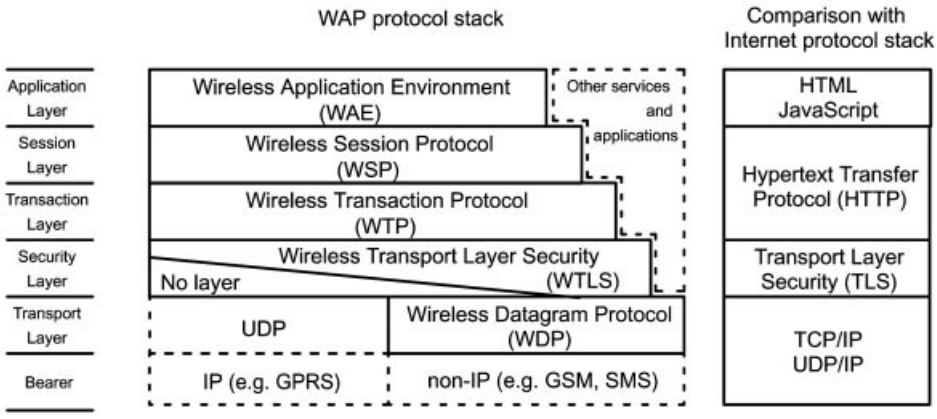

Figure 12.12: *Wireless Application Protocol* (WAP) – protocol architecture

The WTA interface allows the microbrowser to interact with telephony functions. For example, it specifies how calls are initiated from the microbrowser or how entries from the phone directory are sent out.

The main task of the *Wireless Session Protocol* (WSP) is the establishment and termination of a session between the mobile station and the WAP gateway (Figure 12.14). Thereby, connection-oriented (over WTP) as well as connectionless (over datagram services, e.g. WTP) sessions are defined. In case a radio connection breaks down, the session can be stopped for a certain period of time and resumed later.

The *Wireless Transaction Protocol* (WTP) is a lightweight transaction-oriented protocol. Its task is to guarantee the reliable exchange of the mobile station's request and the WAP gateway's response messages (also see Figure 12.14). It thus constitutes the basis for interactive browsing. WTP includes functions for acknowledgement of messages, retransmission of erroneous or lost messages, and the removal of duplicate messages. In addition, an acknowledged and an unacknowledged datagram service is defined for push services, where the server can send content to a mobile station without an initiating request from the mobile user. The server may send an emergency warning, for example.

Optionally, the *Wireless Transport Layer Security* (WTLS) protocol may be employed. It is based in the protocol *Transport Layer Security* (TLS), which is used in the Internet and was formerly known as *Secure Socket Layer* (SSL). WTLS offers basic security functions, such as data integrity, encryption, user identity confidentiality, and authentication between server and mobile station. Moreover, protection against denial-of-service attacks is provided. The functionality of WTLS can be made effective (or not) according to the application and security of the used network. For example, if an application already uses strict security techniques, the complete scope of WTLS functions will not be needed. It is worth mentioning that WTLS can also be used for secure data transfer between two mobile stations (e.g. for authenticated exchange of electronic business cards).

The WAP transport protocol is known as *Wireless Datagram Protocol* (WDP). It is for example used instead of UDP for bearer services that are not based on IP (see Figure 12.12). GSM bearer services for WAP can be either circuit switched data services (e.g.

SMS) or the *General Packet Radio Service* (GPRS), where GPRS of course offers faster data transfer and volume-based billing.

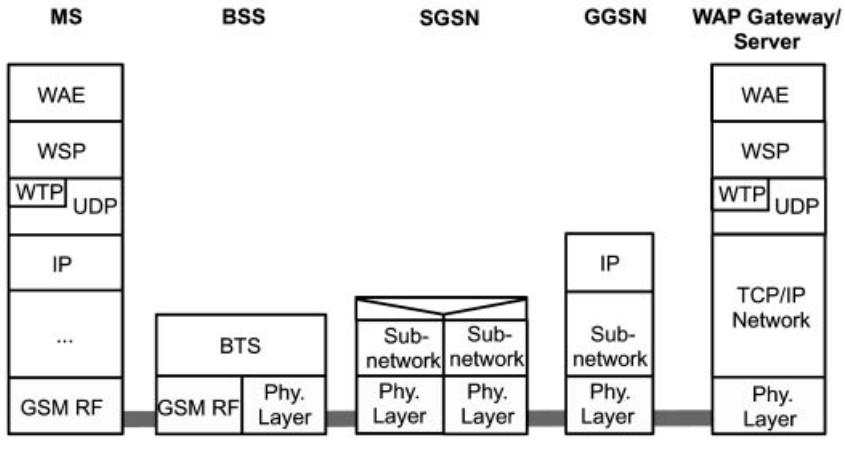

Figure 12.13: Protocol architecture WAP over GPRS (compare with Figure 11.7)

Figure 12.12 indicates that non-WAP protocols can also access specific layers of the WAP stack. Furthermore, not all WAP protocols must be used always. Certain applications may for example only require the services of WTP and underlying layers. Figure 12.13 gives an example, where WAP comes into operation over GPRS as a bearer service.

12.6.3 System Architecture

Figure 12.14 gives a schematic illustration of a typical WAP system architecture. The principle how content is stored in a distributed way within the network and finally offered to the user is similar to the principle of the WWW. Servers store content directly as a WML document or content is generated with scripts. Mobile stations download these contents from the server to their microbrowser, which then presents them to the user. In theory, it is possible to store content in HTML and subsequently convert it to WML, however, in practice, applications and contents directly offered in WML are suited much better [63].

As shown in Figure 12.14, a WAP Gateway acts as an interface between external servers and the mobile stations. Its main tasks include:

- conversion of requests from the WAP protocol stack to the WWW protocol stack (HTTP over TCP/IP) and vice versa (i.e. a protocol gateway functionality)
- encoding and decoding of WML documents into a binary format

The WAP Gateway also represents a proxy server and acts as a cache for frequently requested contents.

The following example illustrates the transaction procedure between a mobile user, the WAP Gateway, and an external server: A subscriber intends to view a document which is offered on a server. His or her WML browser sends a WSP REQUEST to the appropriate

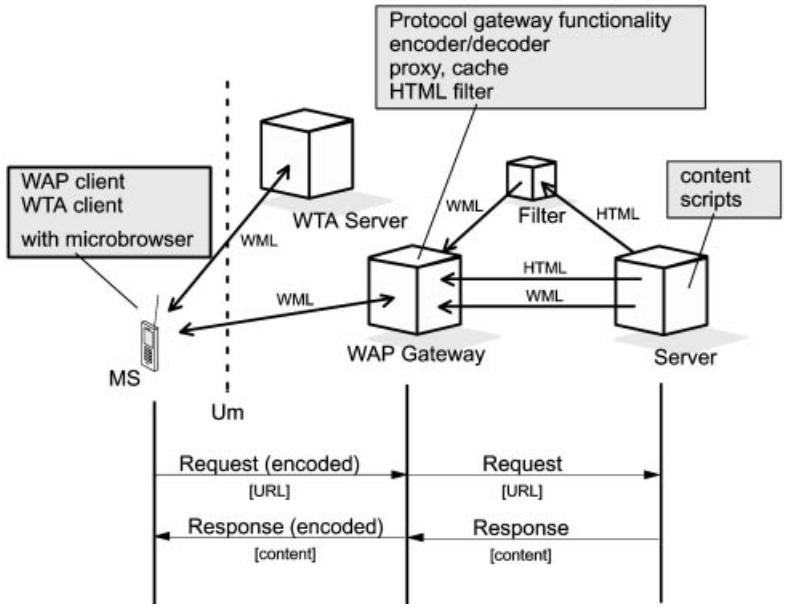

Figure 12.14: WAP system architecture and REQUEST/RESPONSE transaction

address of the server. The request is forwarded to the WAP Gateway, which then converts it into an HTTP REQUEST and contacts the server. Next, the server transmits the requested content in WML format to the WAP Gateway, which writes the content into its cache and sends it in binary-encoded form to the mobile station. The latter presents the first card of the deck on the microbrowser to the user. If the external server transmitted the document in HTML format, the gateway would convert it into WML format.

12.6.4 Services and Applications

The first specification of WAP has been released by the WAP Forum in April 1998. Version 1.1 followed in June 1999 and Version 1.2 in December 1999. WAP terminals have been introduced around February 1999 for the first time, and today there exists a broad variety of WAP products: mobile equipment, gateways, development tools, WML browsers and editors.

Besides the technical implementation in the network and the development of new WAP-capable mobile equipment, innovative WAP services are in particular in demand. These days, several information services are offered over WAP. Subscribers can retrieve news, weather forecasts, stock reports, and local restaurant and event guides with their WAP phone. Furthermore, mobile e-commerce services (e.g. ticket reservation, mobile banking and online auctions) become more and more popular. There is much scope left for new applications. Push services, for example, may transmit important information to mobile stations without the need to request them actively. Highly interesting are so-called location-based services, in which the service knows the current physical location of the user and may use this information in a specific way. Navigation services with displayed maps

on the browser or virtual tourist guides ("I would like to have information about the building on the left side.") are two examples.

The next few years will show whether WAP will win recognition or whether – with the next generation of cellular networks with higher data rates – enhanced mobile stations (with larger displays, etc.) will communicate over an HTTP-over-TCP/IP protocol stack and finally have access to the worldwide open Internet as we know it from the wired world.

12.7 Beyond GSM: On the Road to UMTS

It could be argued that with all its features and coupled with satellite interworking and near-global roaming capabilities, GSM will soon fulfil all the goals of the planned third generation system.
William Webb, Smith System Engineering

With all its enhancements, GSM will represent the mainstream of mobile communication systems for the next several years. However, it is obvious that due to technical and economical reasons, GSM will be followed by a third generation mobile communication system. This system, called *Universal Mobile Telecommunication System* (UMTS) within ETSI/Europe, is aimed to support a wide range of voice and data services, focussing on mobile packet switched data services based on IP technology. An important strategic goal is wireless access to the Internet (see Figure 12.15). Moreover, UMTS will give the mobile user performance similar to the fixed network and will stimulate the development of new mobile multimedia applications. On an international level, i.e. within ITU, the worldwide family of 3G mobile networks is known as *International Mobile Telecommunication 2000* (IMT-2000).

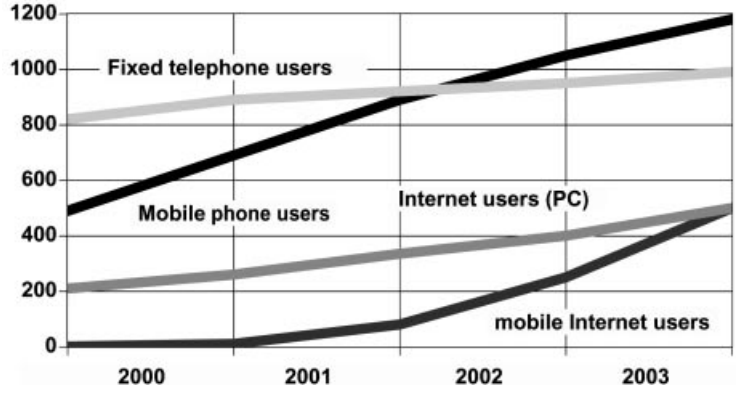

Figure 12.15: Million users worldwide (source: Ericsson, August 2000)

Looking at the rapidly growing number of GSM subscribers (Figure 1.2), it can also be predicted that any future system must support a very high number of subscribers. For the year 2002, it is expected that the worldwide number of mobile phones exceeds the number of fixed telephones (Figure 12.15). In some countries, e.g. Finland and Japan, this is

already reality today. With respect to the radio spectrum needed for the evolving mass market and considering the bandwidth requirements of the envisaged broadband services (up to 2 Mbit/s), the radio interface has to become more spectrum-efficient than today. Therefore, European countries and others have devoted considerable effort to developing the concepts for a flexible and efficient next generation of mobile communication systems.

Taking into account the worldwide success of GSM, UMTS will build, as much as possible, on the existing GSM infrastructure and technology. Overall, it is the intention of all participants in the UMTS standardization process to enable a smooth transition from second-generation GSM to third-generation UMTS/IMT-2000 systems. In particular the GSM service platforms will play a prominent role.

However, UMTS will have a new radio interface, the *UMTS Terrestrial Radio Access* (UTRA), using the frequency bands around 2 GHz and new multiple access techniques. At the World Radio Conference (WRC) in 1992, the decision on the frequency band for IMT-2000 was made (Figure 12.16): The spectrum 1885–2025 MHz and 2110–2200 MHz has been reserved. Europe will allocate 1900–1980 MHz, 2010–2025 MHz, and 2110–2170 MHz, which is 155 MHz in total. In addition, 60 MHz (the bands 1980–2010 MHz and 2170–2200 MHz) can be used for the satellite component of UMTS, denoted as *Mobile Satellite System* (MSS).

The basic decision on the UMTS multiple access technology was made by ETSI in January 1998. Two techniques are provided:

- For operation in paired bands and *Frequency Division Duplex* (FDD), UMTS will use *Wideband-CDMA* (W-CDMA).
- For operation in an unpaired band, using *Time Division Duplex* (TDD), the UMTS system adopts the radio access technique called TD-CDMA, essentially a combination of TDMA and CDMA.

The UTRA proposal is a compromise between the participating companies, and it has worldwide support among equipment manufactures and network operators. According to ETSI, "the agreed solution offers a competitive continuation for GSM to UMTS." A detailed description can be found in [30].

In parallel to activities in Europe, respective standardization bodies around the world elaborated additional proposals for IMT-2000. Altogether, a number of 10 proposals were submitted to ITU for the terrestrial part of IMT-2000. In Japan (ARIB), Korea (TTA), and USA (T1P1) work was done on different W-CDMA systems. In December 1998, those bodies and ETSI joined to form the *Third Generation Partnership Project* (3GPP) [1], which then agreed on a common W-CDMA mode. The IMT-2000 proposals can thus be grouped into three FDD proposals (W-CDMA, cdma2000 (USA, Korea), and UWC-139 (USA)) and three TDD proposals (TD-CDMA, TD-SCMDA (China), and DECT).

In July 1999, the *Operators Harmonization Group* (OHG) agreed upon a *Multi-Carrier* (MC) mode, which also enables operators of IS-95 CDMA networks in North America to migrate to UMTS. Furthermore, the TD-CDMA and TD-SCDMA specifications have been harmonized, and, finally it has been achieved that there will be one worldwide CDMA standard with three different modes:

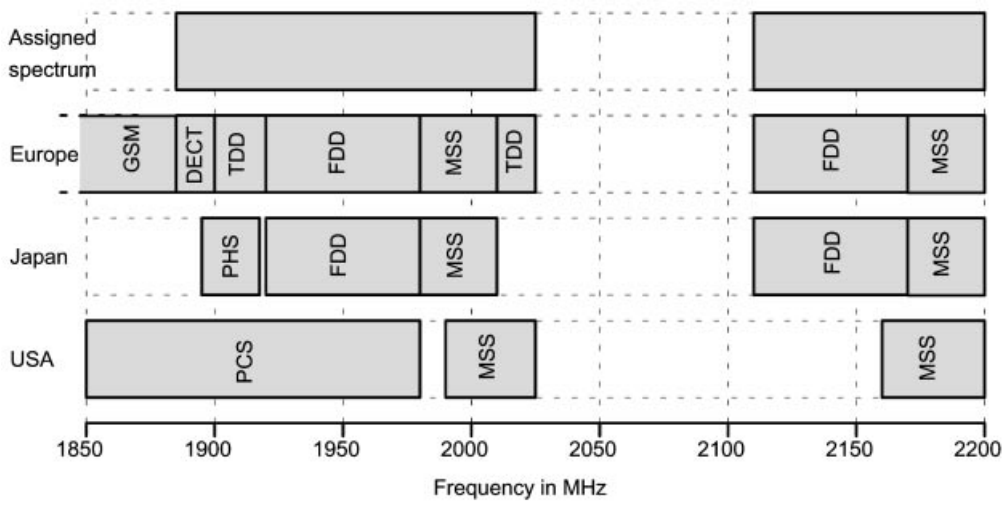

Figure 12.16: Frequency bands for UMTS/IMT-2000

Table 12.3: Third generation cellular network family (without DECT)

Type	**CDMA**			**TDMA**
Name	**W-CDMA**	**cdma2000**	**TD-CDMA**	**EDGE/UWC-136**
Standardization body	3GPP	3GPP2	3GPP (harmonized with Chin. TD-SCDMA)	ETSI and UWCC
Multiple access	Direct sequence CDMA	Multicarrier CDMA	TD-CDMA	TDMA
Duplex	FDD	FDD	TDD	FDD

- a *direct sequence* mode based on W-CDMA (UTRA FDD),
- a *multi-carrier* mode based on cdma2000, and
- a *TDD* mode based on the UTRA proposal TD-CDMA.

Furthermore, the introduction of EDGE (Section 12.3.3) in GSM and in the American system TDMA-136 results in a harmonized and consistent EDGE/UWC-136 system, which offers 3G functionality but operates in the "old" frequency spectrum. Table 12.3 gives an overview of the 3G mobile systems.

Equally important as the radio interface will be the service concept of UMTS. With respect to the service aspects, the standard will provide [18] two sorts of mechanisms:

- Mechanisms to enable the creation of supplementary services including the creation and execution of appropriate MMI (man–machine interface) procedures to the user's terminal.
- Mechanisms to permit the definition of interworking functions, appropriate for the

creation of teleservices and/or end-user applications, including the downloading and execution of these functions in the user's terminal and in appropriate network elements.

These techniques will be similar to the mechanisms in GSM, such as CAMEL, MExE, and SAT.

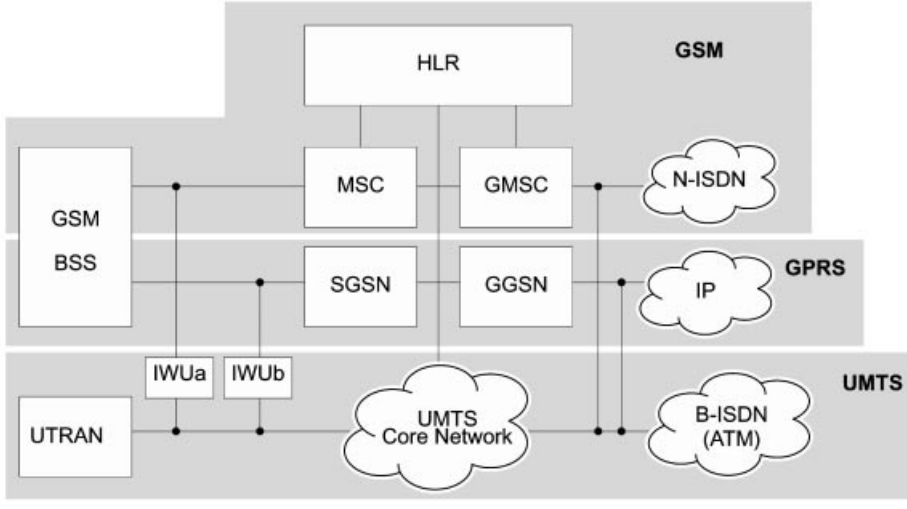

Figure 12.17: Evolution steps from GSM to UMTS

Figure 12.17 shows an evolution scenario for a soft migration from GSM to UMTS. On the basis of the existing circuit and packet switched infrastructure (GSM/GPRS) and entities for mobility management (HLR, MAP), the *UMTS Terrestrial Radio Access Network* (UTRAN) with the new air interface can be introduced as needed (UMTS Phase 1). Here, UTRANs can be installed in parallel to BSSs of GSM, probably by re-using existing locations. In a further step (UMTS Phase 2) – which is probably the first step for new network operators without GSM – a future-proof fixed infrastructure, the *UMTS Core Network* (CN), can be built up. It will be implemented with ATM- and IP-based transport technology.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

References

[1] 3GPP (Third Generation Partnership Project). http://www.3gpp.org.

[2] W. D. Ambrosch, A. Maher, and B. Sassceer. *The Intelligent Network*. Berlin: Springer, 1989.

[3] G. Begin and D. Haccoun. Performance of sequential decoding of high-rate punctured convolutional codes. *IEEE Transactions on Communications*, vol. 42, no. 3, pp. 966–987, 1994.

[4] D. Bertsekas and R. Gallager. *Data Networks*. Englewood Cliffs, NJ: Prentice Hall, 1987.

[5] Ch. Bettstetter, H.-J. Vögel, and J. Eberspächer. GSM phase 2+ General Packet Radio Service GPRS: Architecture, protocols, and air interface. *IEEE Communications Surveys, Special Issue on Packet Radio Networks*, vol. 2, no. 3, 1999.

[6] Bluetooth SIG. http://www.bluetooth.com.

[7] P. Bocker. *ISDN – das diensteintegrierende digitale Nachrichtennetz. 3rd edition*, Berlin: Springer, 1990.

[8] M. Bossert. *D-Netz-Grundlagen – Funkübertragung in GSM-Systemen, Teil 1 und 2. Funkschau*, vols. 22 and 23, 1991.

[9] M. Bossert. *Channel Coding for Telecommunications*. John Wiley & Sons, 1999.

[10] G. Brasche and B. Walke. Concepts, services, and protocols of the new GSM Phase 2+ General Packet Radio Service. *IEEE Communications*, August, pp. 94–104, 1997.

[11] D. Bruhn, E. Ekudden, and K. Hellwig. Adaptive multi-rate: a new speech service for GSM and beyond. In: *Proceedings 3rd ITG Conference Source and Channel Coding*, Munich, January, 2000 (vde-verlag, ITG Technical Report 159).

[12] S. Buckingham. *Mobile Positioning – An Introduction*, http://www.mobilepositioning.-com/, December, 1999.

[13] CCITT Recommendation M.3010. *Principles for a Telecommunications Management Network*, Genf, 1992.

[14] CCITT Recommendation M.3020. *TMN Interface Specification Methodology*, Genf, 1992.

[15] K. David and T. Benkner. *Digitale Mobilfunksysteme*. Stuttgart: B.G. Teubner, 1996.

[16] P. Decker and U. Pertz. Simulative Leistungsbewertung der nichttransparenten Fax-Übertragung im GSM-System. In: B. Walke (ed.). *Informationstechnische Gesellschaft im VDE: Mobile Kommunikation*, Lectures of the ITG-Fachtagung, Sept. 1993, Neu-Ulm. vde-verlag, Berlin, 1993 (ITG Technical Report 124).

[17] J. Eberspächer (ed.). *Vertrauenswürdige Telekomunikation. Proceedings of the Münchner Kreis*, Heidelberg: Hüthig, 1999.

[18] ETSI. http//www.etsi.org.

[19] ETSI. *Annual Report and Activity Report* 1999.

[20] S. Faccin, L. Hsu, R. Koodli, K. Le, and R. Purnadi. GPRS and IS-136 intregration for flexible network and services evolution. *IEEE Personal Communications*, vol. 6, no. 3, pp. 48–54, June 1999.

[21] W. Fuhrmann, V. Brass, U. Janßen, F. Kühl, and W. Roth. Digitale Mobilkommunikationsnetze. In: N. Gerner, H.-G. Hegering, and J. Swoboda (eds.). *Kommunikation in verteilten Systemen*, Tutorium anläßlich der ITG/GI-Fachtagung Kommunikation in verteilten Systemen KiVS, March 1993, Munich. Lehrstuhl f. Datenverarbeitung, Technische Universität München, 1993.

[22] A. Furuskär, J. Näslund, and H. Olofsson. EDGE – Enhanced Data Rates for GSM and TDMA/ 136 Evolution. *Ericsson Review*, no. 1, 1999.

[23] R. H. Glitho and S. Hayes. Telecommunications management network: vision vs. reality. *IEEE Communications*, vol. 33, no. 3, pp. 47–52, 1995.

[24] D. J. Goodman. The wireless internet: promises and challenges. *IEEE Computer*, July 2000.

[25] H. Gottschalk. Zeichengabetechnische Anbindung digitaler Mobilfunknetze an das Festnetz der Telekom. In: B. Walke (ed.). *Informationstechnische Gesellschaft im VDE: Mobile Kommunikation*, Lectures of the ITG-Fachtagung, September, 1993, Neu-Ulm. Berlin: vde-verlag, 1993 (ITG Technical Report 124).

[26] H. Granbohm and J. Wiklund. GPRS – General Packet Radio Service. *Ericsson Review*, no. 2, 1999.

[27] GSM Association. http://www.gsmworld.com.

[28] J. Hagenauer and N. Seshadri. The performance of rate compatible punctured convolutional codes. *IEEE Transactions on Communications*, vol. 38, pp. 966–980, 1990.

[29] Ch. Hartmann and H.-J. Vögel. Teletraffic analysis of SDMA-systems with inhomogeneous MS location distribution and mobility. *Wireless Personal Communications, Special Issue on Space Division Multiple Access*. Kluwer Academic Press, 1999.

[30] *IEEE Communications Magazine, Special Issue on ACTS Mobile Program in Europe*. vol. 36, no. 2, pp. 80–136, 1998.

[31] ISO/IEC 33091991. Information Technology – Telecommunications and information exchange between systems – High-level data link control (HDLC) procedures – Frame structure.

[32] ITU-T Recommendation E.164. Numbering Plan for the ISDN Era.

[33] ITU-T Recommendation Q.735. Multi-Level and Preemption (MLPP).

[34] ITU-T Recommendation V.110. Support of Data Terminal Equipment (DTEs) with V-Series Type Interfaces by an Integrated Services Digital Network (ISDN).

[35] R. Johannesson and K. Sh. Zigangirov. *Fundamentals of Convolutional Codes*. IEEE Press, 1999.

[36] M. Junius and X. Marger. Simulation of the GSM handover and power control based on propagation measurements in the German D1 network. In: *Proceedings of the fifth Nordic Seminar on Digital Mobile Radio Communications (DMR V)*, Helsinki, pp. 367–372, 1992.

[37] R. Kalden, I. Meirick, and M. Meyer. Wireless internet access based on GPRS. *IEEE Personal Communications*, April, pp. 8–18, 2000.

[38] S. Kallel. Complementary punctured convolutional (CPC) codes and their applications. *IEEE Transactions on Communications*, vol. 43, no. 6, pp. 2005–2009, 1995.

[39] K. D. Kammeyer. *Nachrichtenübertragung*. Stuttgart: B.G. Teubner, 1996.

[40] L. Kleinrock. *Queueing Systems – Vol. 1. Theory*. New York: John Wiley & Sons, 1975.

[41] M. Laitinen and J. Rantale. Integration of intelligent network services into future GSM networks. *IEEE Communications Magazine*, June, pp. 76–86, 1995.

[42] W. C. Y. Lee. *Mobile Cellular Telecommunication Systems*. New York: McGraw-Hill, 1989.

[43] Y.-B. Lin. OA&M for the GSM Network. *IEEE Network Magazine*, vol. 11, no. 2, pp. 46–51, 1997.

[44] W. Mende. *Bewertung ausgewählter Leistungsmerkmale von zellularen Mobilfunksystemen*. Dissertation, Hagen Fernuniversität, 1991.

[45] M. Mouly and M.-B. Pautet. Current evolution of the GSM systems. *IEEE Personal Communications Magazine*, October, pp. 9–19, 1995.

[46] E. Natvig. Evaluation of six medium bitrate coders for the pan-European digital mobile radio system. *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 2, pp. 324–334, 1988.

[47] M. Nilsson. Third-generation radio access standards. *Ericsson Review*, no. 3, 1999.

[48] C. E. Perkins. Mobile IP. *IEEE Communications Magazine*, May, pp. 84–99, 1997.

[49] C. E. Perkins. *Mobile IP – Design Principles and Practices*. Reading, MA: Addison-Wesley, 1998.

[50] J. G. Proakis. *Digital Communications, 2nd edition*. New York: McGraw-Hill, 1989.

[51] L. H. Sahasrabuddhe and B. Mukherjee. Multicast routing algorithms and protocols: a tutorial. *IEEE Network, Special Issue on Multicasting: Empowering the Next-Generation Internet*, January 2000.

[52] V. Sahin. Telecommunications management metwork – principles, models and applications. In: S. Aidarous and T. Plevyak (eds.). *Telecommunications Network Management into the 21st Century*. New York: IEEE Press, pp. 72–121, 1993.

[53] S. Schmidt. Management (Operation & Maintenance) von GSM Base Station Subsystemen. In: B. Walke (ed.). *Informationstechnische Gesellschaft im VDE: Mobile Kommu-*

nikation, Lectures of the ITG-Fachtagung, September 1993, Neu-Ulm. Berlin: vde-verlag, 1993 (ITG Technical Report 124).

[54] R. Steele. *Mobile Radio Communications*. London: Pentech Press, 1992.

[55] A. S. Tanenbaum. *Computer Networks, 3rd edition*. Prentice Hall, 1996.

[56] P. Tran-Gia. *Analytische Leistungsbewertung verteilter Systeme*. Berlin: Springer, 1996.

[57] T. S. Towle. TMN as applied to the GSM Network. *IEEE Communications Magazine*, vol. 33, no. 3, pp. 68–73, 1995.

[58] H.-J. Vögel, H. Johr, and A. Grom. Messung und verbesserte Markov-Modellierung transparenter GSM-Datendienste. In: B. Walke (ed.). *Mobile Kommunikation*, Lectures of the ITG-Fachtagung, September 1995, Neu-Ulm. Berlin: vde-Verlag (ITG Technical Report 135), pp. 279–287, 1995.

[59] W3C (World Wide Web Consortium). http://www.w3.org.

[60] B. Walke. *Mobile Radio Networks: Networking and Protocols*. John Wiley & Sons, 1999.

[61] WAP Forum. *Official Wireless Application Protocol – The Complete Standard*. New York: John Wiley & Sons, 1999.

[62] WAP Forum. *White Paper: Wireless Application Protocol*, October, 1999.

[63] WAP Forum. http://www.wapforum.org.

[64] C. Watson. Radio equipment for GSM. In: D. M. Balston and R. C. V. Macario (eds.). *Cellular Radio Systems*. Norwood, MA: Artech House, 1993.

[65] G. Xu and S.-Q. Li. Throughput multiplication of wireless LANs for multimedia services: SDMA protocol design. In: *Proceedings of Globecom 94*, November/December, 1994, San Francisco, CA. New York: IEEE, pp. 1326–1332, 1994.

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

Appendix A: GSM Standards

[1] GSM 01.02, General Description of a GSM PLMN

[2] GSM 01.04, Abbreviations and Acronyms

[3] GSM 02.02, Bearer services (BS) supported by a GSM PLMN

[4] GSM 02.03, Teleservices supported by a GSM PLMN

[5] GSM 02.04, General on supplementary services

[6] GSM 02.09, Security aspects

[7] GSM 02.16, International MS Equipment Identities

[8] GSM 02.17, Subscriber Identity Modules – Functional Characteristics

[9] GSM 02.22, Personalisation of GSM Mobile Equipment (ME); Mobile functionality specification

[10] GSM 02.30, Man-Machine Interface (MMI) of the Mobile Station (MS)

[11] GSM 02.34, High Speed Circuit Switched Data (HSCSD), Stage 1

[12] GSM 02.40, Procedure for call progress indications

[13] GSM 02.42, Network Identity and Timezone (NITZ); Service description, Stage 1

[14] GSM 02.53, Tandem Free Operation (TFO); Service description; Stage 1

[15] GSM 02.57, Mobile Station Application Execution Environment (MExE) – Service description, Stage 1

[16] GSM 02.60, General Packet Radio Service (GPRS), Service Description, Stage 1

[17] GSM 02.63, Packet Data on Signaling channels service (PDS), Stage 1

[18] GSM 02.66, Support of Mobile Number Portability (MNP); Service description, Stage 1

[19] GSM 02.67, Enhanced Multi-Level Precedence and Pre-emption Service (eMLPP), Stage 1

[20] GSM 02.68, Voice Group Call Service (VGCS), Stage 1

[21] GSM 02.69, Voice Broadcast Service (VBS), Stage 1

[22] GSM 02.71, Location Services (LCS) – Service description, Stage 1

[23] GSM 02.72, Call Deflection; Service description, Stage 1

[24] GSM 02.78, Customized Applications for Mobile network Enhanced Logic (CAMEL); Service definition, Stage 1

[25] GSM 02.79, Support of Optimal Routing (SOR); Service definition, Stage 1

[26] GSM 02.81, Line Identification supplementary services, Stage 1

[27] GSM 02.82, Call Forwarding (CF) supplementary services, Stage 1

[28] GSM 02.83, Call Waiting (CW) and Call Hold (HOLD) Supplementary Services, Stage 1

[29] GSM 02.84, Multi Party (MPTY) supplementary services, Stage 1

[30] GSM 02.85, Closed User Group (CUG) Supplementary Services, Stage 1

[31] GSM 02.86, Advice of Charge (AoC) Supplementary Services, Stage 1

[32] GSM 02.87, User-to-User Signalling (UUS) Service Description, stage 1

[33] GSM 02.88, Call Barring (CB) Supplementary Services, Stage 1

[34] GSM 02.90, Unstructured Supplementary Service Data (USSD), Stage 1

[35] GSM 02.91, Explicit Call Transfer (ECT)

[36] GSM 02.93, Completion of Calls to Busy Subscriber (CCBS); Service description, Stage 1

[37] GSM 02.95, Support of Private Numbering Plan (SPNP); Service description, Stage 1

[38] GSM 02.96, Name identification supplementary services; Stage 1

[39] GSM 02.97, Multiple Subscriber Profile (MSP) Service description; Stage 1

[40] GSM 03.01, Network functions

[41] GSM 03.02, Network Architecture

[42] GSM 03.03, Numbering, Addressing and Identification

[43] GSM 03.04, Signalling Requirements Relating to Routing of Calls to Mobile Subscribers

[44] GSM 03.05, Technical performance objectives

[45] GSM 03.07, Restoration procedures

[46] GSM 03.08, Organization of Subscriber Data

[47] GSM 03.09, Handover Procedures

[48] GSM 03.10, GSM PLMN Connection Types

[49] GSM 03.11, Technical realization of supplementary services

[50] GSM 03.12, Location registration Procedures

[51] GSM 03.13, Discontinuous Reception (DRX) in the GSM system

[52] GSM 03.14, Support of Dual Tone Multi-Frequency signalling (DTMF) via the GSM system

[53] GSM 03.15, Technical realization of Operator Determined Barring (ODB)

[54] GSM 03.16, Subscriber data management; Stage 2

[55] GSM 03.18, Basic call handling; Technical realization

[56] GSM 03.19, Subscriber Identity Module Application Programming Interface (SIM API); SIM API for Java Card; Stage 2

[57] GSM 03.20, Security Related Network Functions

[58] GSM 03.22, Functions related to Mobile Station (MS) in idle mode

[59] GSM 03.26, Multiband operation of GSM/DCS 1800 by a single operator

[60] GSM 03.30, Radio network planning aspects

[61] GSM 03.32, Universal Geographical Area Description (GAD)

[62] GSM 03.34, High Speed Circuit Switched Data (HSCSD), Stage 2

[63] GSM 03.38, Alphabets and language-specific information

[64] GSM 03.39, Interface protocols for the connection of Short Message Service Centers (SMSCs) to Short Message Entities (SMEs)

[65] GSM 03.40, Technical realization of Short Message Service (SMS) Point-to-Point (PP)

[66] GSM 03.41, Technical realisation of the short message service cell broadcast (SMSCB)

[67] GSM 03.42, Compression algorithm for text messaging services

[68] GSM 03.43, Support of Videotex

[69] GSM 03.44, Support of teletex in a GSM PLMN

[70] GSM 03.45, Technical Realization of Facsimile Group 3 Service – transparent

[71] GSM 03.46, Technical Realization of Facsimile Group 3 Service – non transparent

[72] GSM 03.47, Example protocol stacks for interconnecting Service Center(s) (SC) and Mobile-services Switching Center(s) (MSC)

[73] GSM 03.48, Security Mechanisms for the SIM application toolkit; Stage 2

[74] GSM 03.49, Example protocol stacks for interconnecting Cell Broadcast Center (CBC) and Base Station Controller (BSC)

[75] GSM 03.50, Transmission planning aspects of the speech service in the GSM PLMN system

[76] GSM 03.53, Tandem Free Operation (TFO); Service description; Stage 2

[77] GSM 03.54, Description for the use of a Shared Inter Working Function (SIWF) in a GSM PLMN; Stage 2

[78] GSM 03.57, Mobile Station Application Execution Environment (MExE) Functional description, stage 2

[79] GSM 03.58, Characterization, test methods and quality assessment for handsfree Mobile Stations (MSs)

[80] GSM 03.60, General Packet Radio Service (GPRS), Service Description, Stage 2

[81] GSM 03.63, Packet Data on Signalling Channels Service (PDS) Service Description, stage 2

[82] GSM 03.64, General Packet Radio Service (GPRS), Overall Description of the Air Interface, Stage 2

[83] GSM 03.66, Support of Mobile Number Portability (MNP) Technical Realization, Stage 2

[84] GSM 03.67, Enhanced Multi-Level Precedence and Pre-emption Service (eMLPP), Stage 2

[85] GSM 03.68, Voice Group Call Service (VGCS), Stage 2

[86] GSM 03.69, Voice Broadcast Service (VBS), Stage 2

[87] GSM 03.70, Routing of calls to/from PDNs

[88] GSM 03.71, Location Services (LCS) – functional description, Stage 2

[89] GSM 03.72, Call Deflection (CD) Supplementary Service, Stage 2

[90] GSM 03.73, Support of Localised Service Area (SoLSA), Stage 2

[91] GSM 03.78, Digital cellular telecommunications system (Phase 2+); Customized Applications for Mobile network Enhanced Logic (CAMEL), Stage 2

[92] GSM 03.79, Support of Optimal Routeing (SOR)

[93] GSM 03.81, Line Identification Supplementary Services, Stage 2

[94] GSM 03.82, Call Forwarding (CF) Supplementary Services, Stage 2

[95] GSM 03.83, Call Waiting (CW) and Call Hold (HOLD) Supplementary Services, Stage 2

[96] GSM 03.84, Multi Party (MPTY) Supplementary Services, Stage 2

[97] GSM 03.85, Closed User Group (CUG) Supplementary Services, Stage 2

[98] GSM 03.86, Advice of Charge (AoC) Supplementary Services, Stage 2

[99] GSM 03.87, User-to-User Signalling (UUS) Supplementary Service, Stage 2

[100] GSM 03.88, Call Barring (CB) Supplementary Services, Stage 2

[101] GSM 03.90, Unstructured Supplementary Service Data (USSD), stage 2

[102] GSM 03.91, Explicit Call Transfer (ECT) supplementary service, stage 2

[103] GSM 03.93, Technical realization of Completion of Calls to Busy Subscriber (CCBS), Stage 2

[104] GSM 03.96, Name Identification Supplementary Services, stage 2

[105] GSM 03.97, Multiple Subscriber Profile (MSP) Phase 1, Stage 2

[106] GSM 04.01, MS-BSS interface General aspects and principles

[107] GSM 04.02, GSM PLMN Access Reference Configuration

[108] GSM 04.03, MS-BSS Interface, Channel Structures and Access Capabilities

[109] GSM 04.04, MS-BSS Layer 1 General Requirements

[110] GSM 04.05, MS-BSS Data Link Layer – General Aspects

[111] GSM 04.06, MS-BSS Data Link Layer Specification

[112] GSM 04.07, Mobile radio interface signalling layer 3 General aspects

[113] GSM 04.08, Mobile Radio Interface Layer 3 Specification

[114] GSM 04.10, Mobile Radio Interface Layer 3 Supplementary Services Specification – General Aspects

[115] GSM 04.11, Point-to-point short message service support on mobile radio interface

[116] GSM 04.12, Cell broadcast short message service support on mobile radio interface

[117] GSM 04.13, Performance Requirements on Mobile Radio Interface

[118] GSM 04.14, Individual equipment type requirements and interworking; Special conformance testing functions

[119] GSM 04.18, Mobile radio interface layer 3 specification Radio Resource Control Protocol

[120] GSM 04.21, Rate Adaptation on the MS-BSS Interface

[121] GSM 04.22, Radio Link Protocol for Data and Telematic Services on the MSBSS Interface

[122] GSM 04.30, Location Services (LCS) Supplementary service operations, Stage 3

[123] GSM 04.31, Location Services (LCS) – Mobile Station (MS) – Serving Mobile Location Centre (SMLC) Radio Resource LCS Protocol (RRLP)

[124] GSM 04.35, Location Services (LCS) – Broadcast Network Assistance for Enhanced Observed Time Difference (E-OTD) and Global Positioning System (GPS)

[125] GSM 04.53, Inband Tandem Free Operation (TFO) of Speech Codecs Service Description, stage 3

[126] GSM 04.56, GSM Cordless Telephony System (CTS), Phase 1 – CTS radio interface layer 3 specification

[127] GSM 04.57, GSM Cordless Telephony System (CTS), Phase 1 – CTS supervising system layer 3 specification

[128] GSM 04.60, General Packet Radio Service (GPRS), MS-BSS Interface, RLC/ MAC Protocol

[129] GSM 04.63, Packet Data on Signalling Channels Service (PDS) service description, stage 3

[130] GSM 04.64, General Packet Radio Service (GPRS), MS-SGSN, Logical Link Control (LLC) Layer

[131] GSM 04.65, General Packet Radio Service (GPRS), MS-SGSN, Subnetwork Dependent Convergence Protocol (SNDCP)

[132] GSM 04.67, Enhanced Multi-Level Precedence and Pre-emption Service (eMLPP), Stage 3

[133] GSM 04.68, Group Call Control (GCC) Protocol

[134] GSM 04.69, Broadcast Call Control (BCC) Protocol

[135] GSM 04.71, Mobile radio interface layer 3, Location Services (LCS) specification

[136] GSM 04.72, Call Deflection (CD) Supplementary Service

[137] GSM 04.80, Mobile Radio Interface Layer 3 Supplementary Services Specification – Formats and Coding

[138] GSM 04.81, Line Identification Supplementary Services, Stage 3

[139] GSM 04.82, Call Forwarding (CF) Supplementary Services, Stage 3

[140] GSM 04.83, Call Waiting (CW) and Call Hold (HOLD) Supplementary Services, Stage 3
[141] GSM 04.84, Multi Party (MPTY) Supplementary Services, Stage 3
[142] GSM 04.85, Closed User Group (CUG) Supplementary Services, Stage 3
[143] GSM 04.86, Advice of Charge (AoC) Supplementary Services, Stage 3
[144] GSM 04.87, User-to-User Signaling (UUS) Supplementary Service, Stage 3
[145] GSM 04.88, Call Barring (CB) Supplementary Services, Stage 3
[146] GSM 04.90, Unstructured Supplementary Service Data (USSD), Stage 3
[147] GSM 04.91, Explicit Call Transfer (ECT) supplementary service, Stage 3
[148] GSM 04.93, Completion of Calls to Busy Subscriber (CCBS), Stage 3
[149] GSM 04.96, Name Identification Supplementary Services, Stage 3
[150] GSM 05.01, Physical Layer on the Radio Path (General Description)
[151] GSM 05.02, Multiplexing and multiple access on the radio path
[152] GSM 05.03, Channel Coding
[153] GSM 05.04, Modulation
[154] GSM 05.05, Radio Transmission and Reception
[155] GSM 05.08, Radio Sub-System Link Control
[156] GSM 05.09, Link Adaptation
[157] GSM 05.10, Radio Subsystem Synchronization
[158] GSM 05.22, Radio link management in hierarchical networks
[159] GSM 05.50, Background for Radio Frequency (RF) requirements
[160] GSM 05.56, GSM Cordless Telephony System (CTS) Phase 1, CTS-FP Radio subsystem
[161] GSM 05.90, GSM Electro-Magnetic Compatibility (EMC) Considerations
[162] GSM 06.01, Full rate speech; Processing functions
[163] GSM 06.02, Half rate speech; Processing functions
[164] GSM 06.06, Half rate speech; ANSI-C code for the GSM half rate speech codec
[165] GSM 06.07, Test sequences for the GSM half rate speech codec
[166] GSM 06.08, Performance characterization of the GSM half rate speech codec
[167] GSM 06.10, Full rate speech transcoding
[168] GSM 06.11, Substitution and muting of lost frames for full-rate speech traffic channels
[169] GSM 06.12, Comfort Noise Aspects for full-rate speech traffic channels
[170] GSM 06.20, Half Rate Speech Transcoding
[171] GSM 06.21, Substitution and Muting of Lost Frames for Half Rate Speech Traffic Channels
[172] GSM 06.22, Comfort Noise Aspects for Half Rate Speech Traffic Channels

[173] GSM 06.31, Discontinuous Transmission (DTX) for Full Rate Speech Traffic Channels

[174] GSM 06.32, Voice activity detection (VAD) for Full Rate Speech Traffic Channels

[175] GSM 06.41, Discontinuous Transmission (DTX) for Half Rate Speech Traffic Channels

[176] GSM 06.42, Voice Activity Detection (VAD) for Half Rate Speech Traffic Channels

[177] GSM 06.51, Enhanced Full Rate (EFR) speech processing functions

[178] GSM 06.53, ANSI-C code for the GSM Enhanced Full Rate (EFR) speech codec

[179] GSM 06.54, Test sequences for the GSM Enhanced Full Rate (EFR) speech codec

[180] GSM 06.55, Performance characterization of the SM Enhanced Full Rate (EFR) speech codec

[181] GSM 06.60, Enhanced Full Rate (EFR) speech transcoding

[182] GSM 06.61, Substitution and Muting of lost frames for Enhanced Full Rate (EFR) speech traffic channels

[183] GSM 06.62, Comfort noise aspects for Enhanced Full Rate (EFR) speech traffic channels

[184] GSM 06.71, Adaptive Multi-Rate (AMR) speech processing functions – General description

[185] GSM 06.73, Adaptive Multi-Rate (AMR) speech – ANSI-C code for the AMR speech codec

[186] GSM 06.74, Test sequences for the Adaptive Multi-Rate (AMR) speech codec

[187] GSM 06.75, Performance Characterization of the GSM Adaptive Multi-Rate (AMR) speech codec

[188] GSM 06.81, Discontinuous Transmission (DTX) for Enhanced Full Rate (EFR) speech traffic channels

[189] GSM 06.82, Voice Activity Detector (VAD) for Enhanced Full Rate (EFR) speech traffic channels

[190] GSM 06.85, Subjective tests on the interoperability of the HR/ FR/ EFR speech codecs, single, tandem and tandem free operation

[191] GSM 06.90, Adaptive Multi-Rate (AMR) speech transcoding

[192] GSM 06.91, Substitution and muting of lost frames for Adaptive Multi Rate (AMR) speech traffic channels

[193] GSM 06.92, Comfort noise aspects for Adaptive Multi-Rate (AMR) speech traffic channels

[194] GSM 06.93, Discontinuous Transmission (DTX) for Adaptive Multi-Rate (AMR) speech traffic channels

[195] GSM 06.94, Voice Activity Detector (VAD) for Adaptive Multi Rate (AMR) speech traffic channels – General description

[196] GSM 07.01, General on terminal adaptation functions for MSs

[197] GSM 07.02, Terminal adaptation functions for services using asynchronous bearer capabilities

[198] GSM 07.03, Terminal adaptation functions for services using synchronous bearer capabilities

[199] GSM 07.05, User of DTE-DCE Interface for Short Message Service (SMS) and Cell Broadcast Services (CBS)

[200] GSM 07.06, Use of the V Series DTE-DCE Interface at the MS for Mobile Termination (MT) configuration

[201] GSM 07.07, AT Command Set for GSM Mobile Equipment

[202] GSM 07.08, GSM Application Programming Interface (GSM-API)

[203] GSM 07.10, Terminal Equipment to Mobile Station (TE-MS) multiplexer protocol

[204] GSM 07.60, Mobile Station (MS) supporting GPRS

[205] GSM 08.01, BSS-MSC Interface – General Aspects

[206] GSM 08.02, BSS/ MSC Interface Principles

[207] GSM 08.04, BSS-MSC layer 1 specification

[208] GSM 08.06, Signalling transport mechanism specification for the BSS-MSC interface

[209] GSM 08.08, BSS-MSC: Layer 3 Specification

[210] GSM 08.14, GPRS BSS-SGSN interface (Gb interface) – Layer 1

[211] GSM 08.16, GPRS BSS-SGSN Interface, Network Service

[212] GSM 08.18, GPRS BSS-SGSN interface – BSS GPRS Protocol (BSSGP)

[213] GSM 08.20, Rate Adaptation on the BSS-MSC Interface

[214] GSM 08.31, Location Services (LCS) – Serving Mobile Location Center – Serving Mobile Location Center (SMLC-SMLC) – SMLCPP specification

[215] GSM 08.51, BSC-BTS interface, general aspects

[216] GSM 08.52, BSC-BTS interface principles

[217] GSM 08.54, BSC-TRX layer 1: structure of physical circuits

[218] GSM 08.56, BSC-BTS layer 2 specification

[219] GSM 08.58, BSC-BTS layer 3 specification

[220] GSM 08.59, BSC-BTS O&M signaling transport

[221] GSM 08.60, Inband control of Remote Transcoders and Rate Adaptors (for EFR and full rate traffic channels)

[222] GSM 08.61, Inband Control of Remote Transcoder and Rate Adaptors (Half Rate)

[223] GSM 08.62, Inband Tandem Free Operation (TFO) of Speech Codecs – Service Description, Stage 3

[224] GSM 08.71, Location Services (LCS): Serving Mobile Location Center – Base Station System (SMLC-BSS) interface layer 3 specification

[225] GSM 09.01, General Network Interworking Scenarios

[226] GSM 09.02, Mobile Application Part (MAP) specification

[227] GSM 09.03, Signaling Requirements on Interworking between the ISDN or PSTN and the PLMN

[228] GSM 09.04, Interworking between the PLMN and the CSPDN

[229] GSM 09.05, Interworking between the PLMN and the PSPDN for PAD Access

[230] GSM 09.06, Interworking between a PLMN and a PSPDN/ISDN for Support of Packet Switched Data Transmission Services

[231] GSM 09.07, General Requirements on Interworking between the PLMN and the ISDN oder PSTN

[232] GSM 09.08, Application of the Base Station System Application Part (BSSAP) on the E-interface

[233] GSM 09.09, Interworking between Phase 1 infrastructure and Phase 2 Mobile Stations (MS)

[234] GSM 09.10, Information element mapping between MS-BSS/BSS-MSC signalling procedures and MAP

[235] GSM 09.11, Signaling interworking for supplementary services

[236] GSM 09.13, Signaling interworking between ISDN supplementary services; Application Service Element (ASE) and Mobile Application Part (MAP) protocols

[237] GSM 09.16, GPRS, Serving GPRS Support Node (SGSN) – Visitor Location Register (VLR); Gs interface network service specification

[238] GSM 09.18, GPRS, Serving GPRS Support Node (SGSN) – Visitor Location Register (VLR); Gs interface layer 3 specification

[239] GSM 09.31, Location Services (LCS) – Base Station System Application Part LCS Extension (BSSAP-LE)

[240] GSM 09.60, GPRS, GPRS Tunnelling Protocol (GTP) across the Gn and Gp Interface

[241] GSM 09.61, GPRS, Interworking between the PLMN and PDN

[242] GSM 09.78, CAMEL Application Part (CAP) specification

[243] GSM 09.90, Interworking between Phase 1 Infrastructure and Phase 2 Mobile Stations (MS)

[244] GSM 09.91, Interworking Aspects of the SIM/ ME Interface between Phase 1 and Phase 2

[245] GSM 10. xx, Project schedules and open issues

[246] GSM 11.10, Mobile station conformity specifications

[247] GSM 11.11 + 12 + 18, Subscriber Identity Module – Mobile Equipment (SIM-ME) interface specification

[248] GSM 11.17, Subscriber Identity Module (SIM) conformance test specification

[249] GSM 11.20-26, Base Station System (BSS); Equipment specification

[250] GSM 11.30, Mobile Services Switching Center

[251] GSM 11.31, Home Location Register specification

[252] GSM 11.32, Visitor Location Register specification

[253] GSM 11.40, System simulator specification
[254] GSM 12.00, Objectives and Structures of Network Management
[255] GSM 12.01, Common Aspects of GSM Network Management
[256] GSM 12.02, Subscriber, Mobile Equipment and Service Data Administration
[257] GSM 12.03, Security Management
[258] GSM 12.04, Performance Data Measurement
[259] GSM 12.05, Subscriber Related Event and Call Data
[260] GSM 12.06, GSM Network Change Control
[261] GSM 12.07, Operations and Performance Management
[262] GSM 12.08, Subscriber and equipment trace
[263] GSM 12.10, Maintenance Provisions for Operational Integrity of Mobile Stations
[264] GSM 12.11, Maintenance of the Base Station System
[265] GSM 12.13, Maintenance of the Mobile-services Switching Center
[266] GSM 12.14, Maintenance of Location Registers
[267] GSM 12.15, GPRS, GPRS Charging
[268] GSM 12.20, Base Station System (BSS) management information
[269] GSM 12.21, Network Management Procedures and Messages on the Abis Interface
[270] GSM 12.22, Interworking of GSM Network Management (NM) Procedures and Messages at the BSC
[271] GSM 13.xx, Attachment requirements

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

Appendix B: GSM Addresses

BCC	Base Transceiver Station Color Code
BSIC	Base Transceiver Station Identity Code
(M)CC	(Mobile) Country Code
CI	Cell Identifier
FAC	Final Assembly Code
GCI	Global Cell Identity
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity
LAC	Location Area Code
LAI	Location Area Identity
LMSI	Local Mobile Subscriber Identity
MNC	Mobile Network Code
MSIN	Mobile Subscriber Identification Number
MSISDN	Mobile Station ISDN Number
MSRN	Mobile Station Roaming Number
NCC	Network Colour Code
NDC	National Destination Code
NMSI	National Mobile Subscriber Identity
SN	Subscriber Number
SNR	Serial Number
TAC	Type Approval Code
TMSI	Temporary Mobile Subscriber Identity

	TAC	FAC	SNR	SP	(M)CC	MNC	MSIN	SN	LAC	CI	NCC	BCC
IMEI	■	■	■	■								
IMSI					■	■	■					
MSISDN					■	■		■				
MSRN					■	■		■				
LAI					■	■			■			
CI										■		
BSIC											■	■

Additional Addresses for GPRS:

APN	Access Point Name
GSN Address	GPRS Support Node Address (e.g. IP address of SGSN and GGSN)
GSN Number	GPRS Support Node Number (for SGSN and GGSN for communication with e.g. HLR and VLR)
IP Address	Internet Protocol address
NSAPI	Network layer Service Access Point Identifier
P-TMSI	Packet Temporary Mobile Subscriber Identity
PDP Address	PDP Address (e.g. IP address, X.25 address)
RAC	Routing Area Code
RAI	Routing Area Identity
TID	Tunnel Identifier (=IMSI+NSAPI)
TLLI	Temporary Logical Link Identifier

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer,
Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

Appendix C: Acronyms

3GPP	3rd Generation Partnership Project
3GPP2	3rd Generation Partnership Project 2
3PTY	Three Party Service
8-PSK	8 Phase Shift Keying
A3, A5, A8	Encryption Algorithms
AB	Access Burst
Abis	BTS-BSC Interface
ACELP	Algebraic Code Excitation Linear Prediction
ACSE	Association Control Service Element
AGCH	Access Grant Channel
AMR	Adaptive Multi-Rate (codec)
AOC	Advice of Charge
ARQ	Automatic Repeat Request
ASCI	Advanced Speech Call Items
ASE	Application Service Element
ATM	Asynchronous Transfer Mode
AUC	Authentication Center
BAIC	Barring of All Incoming Calls
BAOC	Barring of All Outgoing Calls
BCC	Base Transceiver Station Color Code
BCCH	Broadcast Control Channel
BCH	Broadcast Channel
BCS	Block Check Sequence
BFI	Bad Frame Indication
BG	Border Gateway
BIC-Roam	Barring of Incoming Calls when Roaming Outside the Home PLMN
B-ISDN	Broadband ISDN
Bm	Mobile B Channel
BN	Bit Number
BOIC	Barring of Outgoing International Calls
BOIC-exHC	Barring of Outgoing International Calls except those to Home PLMN
BSC	Base Station Controller
BSIC	Base Tranceiver Station Identity Code

BSS	Base Station Subsystem
BSSAP	Base Station System Application Part
BSSAP+	Base Station System Application Part +
BSSGP	Base Station System GPRS Application Protocol
BSSMAP	Base Station System Management Application Part
BTS	Base Transceiver Station
BTSM	Base Transceiver Station Management
CA	Cell Allocation
CAMEL	Customized Applications for Mobile Network Enhanced Logic
CAP	CAMEL Application Part
CBCH	Cell Broadcast Channel
CC	Country Code
CCBS	Completion of Call to Busy Subscriber
CCCH	Common Control Channel
CDMA	Code Division Multiple Access
CELP	Code Excited Linear Prediction Coding
CFB	Call Forwarding on Mobile Subscriber Busy
CFNRc	Call Forwarding on Mobile Subscriber Not Reachable
CFNRy	Call Forwarding on No Reply
CFU	Call Forwarding Unconditional
CI	Cell Identifier
CLIP	Calling Number Identification Presentation
CLIR	Calling Number Identification Restriction
CLNS	Connectionless Network Service
CM	Connection Management
CMISE	Common Management Information Service Element
CN	Core Network
CODEC	Coder/Decoder
COLP	Connected Line Identification Presentation
COLR	Connected Line Identification Restriction
CONF	Conference Calling
CONS	Connection-Oriented Network Service
CRC	Cyclic Redundancy Check
CT	Call Transfer
CUG	Closed User Group
CW	Call Waiting
DAB	Digital Audio Broadcast
DB	Dummy Burst
DCCH	Dedicated Control Channel
DCN	Data Communication Network
DECT	Digital Enhanced Cordless Telecommunication
DHCP	Dynamic Host Configuration Protocol

Dm	Mobile D Channel
DNS	Domain Name Service
DRX	Discontinuous Reception
DSL	Digital Subscriber Line
DTMF	Dual Tone Multiple Frequency
DTX	Discontinuous Transmission
DTAP	Direct Transfer Application Part
DVB	Digital Video Broadcast
ECSD	Enhanced Circuit Switched Data
EDGE	Enhanced Data Rates for GSM Evolution
EFR	Enhanced Full Rate (CODEC)
EGPRS	Enhanced GPRS
EIR	Equipment Identity Register
EMLPP	Enhanced Multi-Level Precedence and Pre-emption Service
E-OTD	Enhanced Observed Time Difference
ERMES	European Radio Messaging Standard
ETSI	European Telecommunication Standards Institute
FA	Fax Adapter
FAC	Final Assembly Code
FACCH	Fast Associated Control Channel
FCAPS	Fault, Configuration, Accounting, Performance, Security
FB	Frequency Correction Burst
FCCH	Frequency Correction Channel
FCS	Frame Check Sequence
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FN	TDMA Frame Number
FPH	Freephone Service
FPLMTS	Future Public Land Mobile Telecommunication System
FTAM	File Transfer Access and Management
GCR	Group Call Register
GEA	GPRS Encryption Algorithm
GGSN	Gateway GPRS Support Node
GMLC	Gateway Mobile Location Center
GMM/SM	GPRS Mobility Management and Session Management protocol
GMSC	Gateway MSC
GMSK	Gaussian Minimum Shift Keying
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSC	GSM Speech Codec
GSM	Global System for Mobile Communication

GSMSS	GSM Satellite System
GSN	GPRS Support Node
GTP	GPRS Tunnelling Protocol
HLR	Home Location Register
HSCSD	High Speed Circuit Switched Data
HSN	Hopping Sequence Number
HTML	Hypertext Markup Language
HTTP	Hypertext Transport Protocol
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity
IMT-2000	International Mobile Telephone System 2000
IN	Intelligent Network
INAP	Intelligent Network Application Part
IP	Internet Protocol
IPv4	Internet Protocol Version 4
IPv6	Internet Protocol Version 6
ISC	International Switching Center
ISDN	Integrated Services Digital Network
IWF	Interworking Function
Kc	Cipher/Decipher Key
Ki	Subscriber Authentication Key
L2R	Layer 2 Relay
L2RBOP	Layer 2 Relay Bit-Oriented Protocol
L2RCOP	Layer 2 Relay Character-Oriented Protocol
LA	Location Area
LAC	Location Area Code
LAI	Location Area ID
LAPDm	Link Access Procedure D mobile
LCN	Local Communication Network
LCS	Location Service
LEO	Low Earth Orbiting satellite
LLC	Logical Link Control layer
LMSI	Local Mobile Subscriber Identity
LPC	Linear Predictive Coding
LTP	Long Term Prediction
MA	Mobile Allocation
MAC	Medium Access Control layer
MAH	Mobile Access Hunting
MAIO	Mobile Allocation Index Offset
MAP	Mobile Application Part
MC	Multi Carrier
MCI	Malicious Call Identification

MD	Mediation Device
MEO	Medium Earth Orbiting satellite
MexE	Mobile Station Application Execution Environment
MHS	Message Handling System
MM	Mobility Management
MMI	Man Machine Interface
MNC	Mobile Network Code
MOS	Mean Opinion Score
MS	Mobile Station
MSC	Mobile Switching Center
MSIN	Mobile Subscriber Identification Number
MSISDN	Mobile Station ISDN Number
MSK	Minimum Shift Keying
MSRN	Mobile Station Roaming Number
MSS	Mobile Satellite System
MT	Mobile Termination
MTP	Message Transfer Part
NB	Normal Burst
NCC	Network Colour Code
NCH	Notification Channel
NDC	National Destination Code
NE	Network Element
NMSI	National Mobile Subscriber Identity
NMT	Nordic Mobile Telephone
NS	Network Service
NSAPI	Network Layer Service Access Point Identifier
OHG	Operators Harmonization Group
OMAP	Operation, Maintenance and Administration Part
OMC	Operation and Maintenance Center
OMSS	Operation and Maintenance Subsystem
OS	Operation System
OSI	Open Systems Interconnection
P-IWMSC	Packet Interworking MSC
PACCH	Packet Associated Control Channel
PAD	Packet Assembler/Disassembler
PAGCH	Packet Access Grant Channel
PBCCH	Packet Broadcast Control Channel
PBX	Private Branch Exchange
PCCCH	Packet Common Control Channel
PCH	Paging Channel
PCN	Personal Communication Network
PCS	Personal Communication System

PDA	Personal Digital Assistant
PDCH	Packet Data Channel
PDN	Public Data Network
PDN	Packet Data Network
PDP	Packet Data Protocol
PDTCH	Packet Data Traffic Channel
PDU	Protocol Data Unit
PLL	Physical Link Layer
PLMN	Public Land Mobile Network
PNCH	Packet Notification Channel
PPCH	Packet Paging Channel
PRACH	Packet Random Access Channel
PSK	Phase Shift Keying
PSPDN	Packet Switched Public Data Network
PTCCH	Packet Timing Advance Control Channel
PTM	Point-to-Multipoint Service
PTM-G	Point-to-Multipoint Service – Group Call
PTM-M	Point-to-Multipoint Service – Multicast
P-TMSI	Packet Temporary Mobile Subscriber Identity
PTP	Point-to-Point Service
QN	Quarter Bit Number
QoS	Quality of Service
RA	Rate Adaptation
RA	Routing Area
RACH	Random Access Channel
RAI	Routing Area Identity
RAND	Random Number (for authentication)
REVC	Reverse Charging
RFCH	Radio Frequency Channel
RFL	Physical RF Layer
RFN	Reduced TDMA Frame Number
RLC	Radio Link Control layer
RLL	Radio in the Local Loop
RLP	Radio Link Protocol
ROSE	Remote Operation Service Element
RPE	Regular Pulse Excitation
RR	Radio Resource Management
SACCH	Slow Associated Control Channel
SAT	SIM Application Toolkit
SATIG	Satellite Interest Group
SB	Synchronization Burst
SCCP	Signaling Connection Control Part

SCH	Synchronization Channel
SCN	Sub Channel Number
SCP	Service Control Point
SDCCH	Stand-alone Dedicated Control Channel
SDMA	Space Division Multiple Access
SGSN	Serving GPRS Support Node
SID	Silence Descriptor
SIM	Subscriber Identity Module
SM-CP	Short Message Control Protocol
SM-RP	Short Message Relay Protocol
SMLC	Serving Mobile Location Center
SMS	Short Message Service
SMS-GMSC	Short Message Service – Gateway MSC
SMS-IWMSC	Short Message Service – Interworking MSC
SMS-SC	Short Message Service – Service Center
SMSCB	Short Message Service Cell Broadcast
SMSS	Switching and Management Subsystem
SN	Subscriber Number
SNDCP	Subnetwork Dependent Convergence Protocol
SNR	Serial Number
SNR	Signal to Noise Ratio
SOSS	Support of Operator Specific Services
SP	Signaling Point
SPC	Signaling Point Code
SRES	Session Key (for authentication)
SS	Supplementary Services
SSL	Secure Socket Layer
SSP	Service Switching Point
TA	Terminal Adaptor
TA	Timing Advance
TAC	Type Approval Code
TACS	Total Access System
TBF	Temporary Block Flow
TCAP	Transaction Capabilities Application Part
TCH	Traffic Channel
TCP	Transmission Control Protocol
TD-CDMA	Time Division CDMA
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TD-SCDMA	Time Division Synchronous CDMA
TETRA	Terrestrial Trunked Radio
TFI	Temporary Flow Identifier

TFO	Tandem Free Operation
TID	Tunnel Identifier
TLLI	Temporary Logical Link Identifier
TLS	Transport Layer Security
TMN	Telecommunication Management Network
TMSI	Temporary Mobile Subscriber Identity
TN	Time Slot Number
TOA	Time of Arrival
TSC	Training Sequence Code
UDI	Unrestricted Digital Information
UDP	User Datagram Protocol
Um	Air/Radio Interfacee
UMTS	Universal Mobile Telecommunication System
UPT	Universal Personal Telecommunication
URAN	UMTS Radio Access Network
URL	Universal Resource Locator
USF	Uplink State Flag
UTRA	UMTS Terrestrial Radio Access
UTRAN	UMTS Terrestrial Radio Access Network
UUS	User to User Signaling
UWCC	Universal Wireless Communications Consortium
VAD	Voice Activity Detection
VBS	Voice Broadcast Service
VGCS	Voice Group Call Service
VLR	Visited Location Register, VLR Nummer
W3C	World Wide Web Consortium
WAE	Wireless Application Environment
WAP	Wireless Application Protocol
WBMP	Wireless Bitmap (Format)
W-CDMA	Wideband CDMA
WDP	Wireless Datagram Protocol
WLL	Wireless Local Loop
WML	Wireless Markup Language
WRC	World Radio Conference
WSP	Wireless Session Protocol
WTA	Wireless Telephony Application
WTLS	Wireless Transport Layer Security protocol
WTP	Wireless Transaction Protocol
WWW	World Wide Web
XML	Extensible Markup Language
XSL	Extensible Style Language

GSM Switching, Services and Protocols: Second Edition. Jörg Eberspächer, Hans-Jörg Vögel and Christian Bettstetter

Print ISBN 0-471-49903-X Online ISBN 0-470-84174-5

Index

D

E

F

G

H

I

N

O

P

Q

R

S

T

U

V

W

X

Z