The McGraw-Hill Companies

2.5G MOBILE NETWORKS

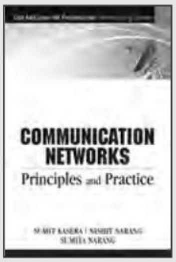
GPRS and EDGE

The McGraw·Hill Companies

McGraw-Hill Professional—*Networking Series*

This series dwells on the various aspects of communication networks. The titles under the series include both introductory texts for beginners as well as technology specific texts in wire line and wireless domain. These are designed to cover the subject comprehensively, yet are easy to understand and free of jargon. Profusely illustrated, these books serve the needs of professional engineers and students.

Titles in this series:



Communication Networks: *Principles and Practice*

Sumit Kasera, Nishit Narang and *Sumita Narang*

This is an introductory text on networking concepts. In a novel approach adopted by the authors, the book is organized in a manner such that each chapter first provides details of a particular aspect of communication networks and then elaborates how the concept applies to a contemporary networking technology like TCP/IP, ATM, or even Third Generation (3G) mobile network.

ATM Networks: *Concepts and Protocols*

Sumit Kasera

This is a single stop reference on ATM, covering a wide range of topics including ATM protocol reference model, core concepts (like Routing, Signalling, and Switching), ATM network architecture, and applications of ATM network (e.g. 3G and MPLS).

3G Networks: *Architecture, Protocols and Procedures*

Sumit Kasera and *Nishit Narang*

This describes the Third Generation (3G) mobile networks for UMTS WCDMA system based on 3GPP specifications. The book comprehensively covers UMTS network architecture (including access network and core network), protocols (including access network and core network protocols), procedures (including radio resource management, mobility management, call/session handling and security management), and services (including supplementary services and value-added services).

2G Mobile Networks: *GSM and HSCSD*

Nishit Narang and Sumit Kasera

This book exhaustively covers fundamentals of GSM technology (including cellular concepts, network and protocol architecture of GSM), GSM Air Interface (including GSM frame hierarchy, burst structure, physical and logical channels), Mobile Station, Access and Core Networks, and finally service aspects of GSM networks (including voice transfer, SMS, cell broadcast service, local services, and circuit switched data and high-speed circuit switched data).

The McGraw·Hill Companies

2.5G MOBILE NETWORKS

GPRS and EDGE

Sumit Kasera

Principal Systems Engineer, Aricent,
Gurgaon, India

Nishit Narang

Principal Systems Engineer, Aricent,
Gurgaon, India

A.P. Priyanka

Senior Testing Engineer, Aricent
Gurgaon, India

Tata McGraw-Hill Publishing Company Limited

NEW DELHI

McGraw-Hill Offices

New Delhi New York St Louis San Francisco Auckland Bogotá Caracas
Kuala Lumpur Lisbon London Madrid Mexico City Milan Montreal
San Juan Santiago Singapore Sydney Tokyo Toronto

Tata McGraw-Hill

Published by Tata McGraw-Hill Publishing Company Limited,
7 West Patel Nagar, New Delhi 110 008.

ISBN (13): 978-0-07-065692-5
ISBN (10): 0-07-065692-4

Managing Director: *Ajay Shukla*
Head—Professional and Healthcare: *Roystan La'Porte*
Publishing Manager—Professional: *R Chandra Sekhar*
Jr Sponsoring Editor (Science & Technology): *Nimisha Goswami*
Asst Manager—Production: *Sohan Gaur*
Manager—Sales and Marketing: *S Girish*
Product Manager (Services, Technology and Computing): *Rekha Dhyani*
Controller—Production: *Rajender P Ghansela*
Asst General Manager—Production: *B L Dogra*

Typeset at Tej Composers, WZ-391, Madipur, New Delhi 110063, and printed at Adarsh Printers, C-50-51, Mohan Park, Naveen Shahdara, Delhi-110 032.

Cover Design: Kapil Gupta
Cover Printer: Rashtriya Printers

RCZQCRXXRABQD

To

My wife Manisha for loving me as I am.

• Sumit Kasera

My wife Sumita, for being a constant source of encouragement.

• Nishit Narang

My parents, Rajni and A. Prakash for making me what I am.

• A.P. Priyanka

Foreword

It gives me great pleasure to write the foreword of 2.5*G Mobile Networks: GPRS and EDGE*. This book culminates the trilogy of books written by Sumit Kasera and Nishit Narang on three key wireless technologies, namely 2G Networks (GSM and HSCSD), 2.5G Networks (GPRS and EDGE) and 3G Networks (UMTS and IMS). Achieved in a span of about 6 years, this feat is remarkable and deserves to be commended. These books brought by McGraw-Hill are cogently written yet are comprehensive and cover a wide array of topics.

Delving a bit in evolution of cellular networks, the first 2G GSM networks were commercially deployed in Europe in early nineties while the first commercial launch of 3G UMTS networks happened in Japan in 2001. Somewhere in between the two, the GPRS network was launched. Thus, both 2G and 2.5G networks, as compared to the 3G networks, are now more than a decade old. In this context, one sometimes wonders at the relevance of bringing about a book on 2G (launched in 2006) and 2.5G (being launched in early 2008). If one wants the answer, then he or she does not have to go too far and just look at the current corporate battle for the 2G spectrum in India. The battle is fought between various companies that include incumbents as well as aspirants, between camps (GSM versus CDMA), government ministers, telecom regulators and the appellate. Without any formal auction of spectrum, the bids of scarce air interface have started outwitting each other. The spectrum war indicates the growth potential of these cellular technologies including GSM as well as GPRS.

As a matter of fact, some of the leading GSM equipment providers are developing low cost GSM/GPRS network elements specifically for India and other such economies and trying to spread the GSM/GPRS networks in parts of the world where mobile penetration is low. EDGE, with its ability to deliver higher data rates with minimum incremental cost, is also a key weapon in the hands of operators to drive growth in the "data" segment. In this backdrop where emerging economies including China and India and other developing or underdeveloped countries show huge potential for 2G/2.5G technology, a book by an Indian author is certainly a welcome step. In any case, the available literature on these technologies is inadequate, costly, full of jargon and complex equations and many times unavailable.

In contrast, this new book derives its strength from lucid procedural descriptions and detailed illustrations, which both novice as well as experienced professionals can appreciate equally. The book start with a reasonably comprehensive overview of GSM networks that essentially summarizes their earlier work on the topic. The next chapter then talks about the GPRS network architecture—both transmission as well as control plane architecture. Air interface is discussed next where the key topics like logical channels and channel coding are nicely elaborated. The next two chapters cover the most important air interface protocols, namely Medium Access Protocol and Radio Link Control protocol. There is a healthy mix of concepts and the protocol in these chapters. The presentation then moves onto core network protocols including GPRS Mobility Management and Session Management. Both these chapters are well illustrated with flow diagrams to drive the point. GPRS user plane aspects are further detailed in Chapter 8. The book ends with a final chapter on EDGE where the key differences between GPRS and EDGE are summarized, along with coverage on the important aspects of EDGE.

As one of the leaders in creating communication software for wireless networks, we at Aricent are constantly ramping up engineers to create a strong foundation on this basic wireless technology. The need for a greater variety of introductory and reference reading material from authors with practical experience has been long felt. The authors of this book are actively involved in design, system integration, testing and piloting of GSM/GPRS network elements and have brought out this book from their rich experience. Thus, I once again congratulate my colleagues Sumit Kasera and Nishit Narang, as well as the relatively junior yet no less competent A. P. Priyanka, who through their deep understanding of the subject and penchant for sharing their knowledge have presented yet another of their works. I am quite hopeful that this book will be a valuable guide to both students and engineers working on GPRS and EDGE. The book will also be a great follow up for those who have read GSM and want to know more on the intermediate technologies before jumping to 3G UMTS.

N. Mohanrangan
Assistant Vice President
Aricent

Preface

Raison d'être

Mobile telecommunications have left the POTS (plain old telephone system) far behind. Today, one in every three has a mobile connection. In parts of Europe and Asia, close to 80 percent of the population carries a mobile—developed countries are close to getting saturated. The situation in developing countries will be encouragingly similar in a few years—countries like China and India are experiencing exciting growth opportunities for the mobile telecommunication industry. Given the scenario, certain trends are evident: as the subscriber base gets saturated, growth will come from new services and not new subscribers. For example, multi-media services that provide integration of voice and data will become critical. Moreover, the bandwidth limitation of existing wireless networks will force operators to deploy next generation networks. This will result in end-user benefit—users can enjoy bandwidth-intensive multimedia applications that are being developed but limited by data carrying capacities of mobile networks.

The pressing need for greater bandwidth and service capabilities led International Telecommunications Union – Telecommunication Standardization (ITU-T) to form a vision of Third Generation (3G) networks. In sync with this development our book on Third Generation (3G) networks (2004) was very well received. Even though 2G and 2.5G in comparison to 3G are older technologies, the following, current, strong outlook for these two technologies has necessitated development of books on 2G and 2.5G mobile networks:

1. GSM (*Global System for Mobile Communications*) is the most popular Second Generation (2G) network for wireless voice and low data rate applications, such as SMS (Short Message Service). With an existing subscriber base of two billion, large parts of the world, particularly China and India, hold a very strong growth potential for 2G GSM and 2.5G GPRS networks in the coming five years. As of mid-2006, 2G GSM networks accounted for more than 75% of the mobile telecommunications market—3G UMTS networks even after five years are no where even near. Even for services that require high bandwidth, simpler options, such as GPRS and EDGE are preferred.
2. Technological improvements have resulted in 2G and 2.5G networks that are capable of higher data rates. Among these are: circuit-switching techniques, such as *High-Speed Circuit Switched Data* (HSCSD), and packet-multiplexing techniques including *General Packet Radio Services* (GPRS) and *Enhanced Data Rates for Global Evolution* (EDGE).
3. 3G UMTS networks are backward compatible with 2G GSM and 2.5G GPRS networks. Over the next few years, GSM and GPRS operators would move towards UMTS. During

the transition, which is likely to span many years, they will have to support GSM and GPRS. GSM-3G and GPRS-3G integration will be important during this time, in which multi-mode handsets will appear. This makes it imperative to know about GSM and GPRS.

4. Several mobile telecommunication companies around the world are still involved in the development of GSM and GPRS products. And this trend is projected for another five to seven years, until 2015.

To cover the journey from 2G to 2.5G to 3G we present a series of three books (Ref. Fig. P.1). As mentioned earlier the first book *3G Generation Networks* was published in 2004. The second book *2G Mobile Networks* was published in late 2006. This book completes the trilogy of books on the three generation of mobile cellular technologies.

About the Book

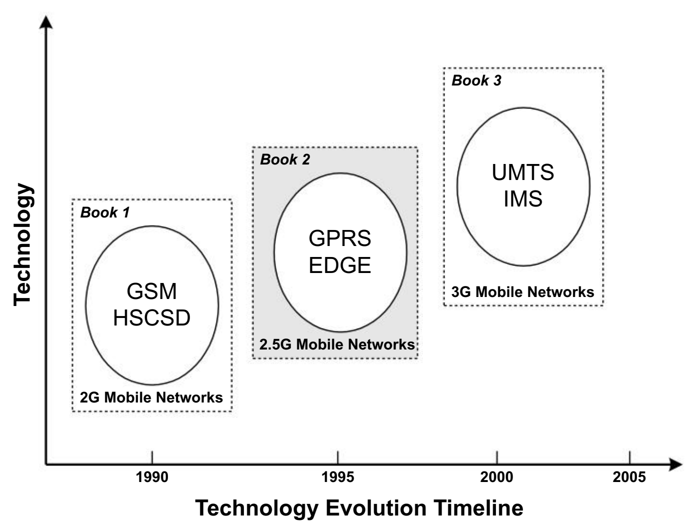
This book presents GPRS concepts in a simple, yet effective manner. All the important topics have been given adequate attention. The overall idea is to first provide introduction material and then take the discussion from air interface to core network.

The book starts with a detailed overview of GSM in **Chapter 1**. This chapter essentially summarizes the contents of *2G Mobile Networks*. Readers well versed with GSM may skip this chapter. Nonetheless, it is recommended to go through this chapter as GPRS uses GSM as its underlying foundation.

Chapter 2 describes GPRS network architecture. All the network entities and the interfaces between them are described in this chapter. A brief description of the relevant protocols for the interfaces is also presented.

FIGURE P.1

Three Book Series on Wireless Technologies



GPRS Air Interface is described in **Chapter 3**. It covers GPRS logical channels, channel organization and various schemes of channel coding (CS1 to CS4). The chapter also covers important air interface concepts like timing advance and power control.

The twin protocols of Medium Access Control (MAC) and Radio Link Control (RLC) are described in **Chapter 4** and **Chapter 5,** respectively. The main function of MAC layer is establishment and release of Temporary Block Flow (TBF). These are described in detail in Chapter 4. Other important functions of MAC including System Information Broadcast, Paging and Cell Re-selection are also covered in this chapter. The next chapter on RLC primarily deals with block structure of both control blocks and data blocks. The RLC operations for acknowledgement and reliable delivery are also discussed.

The core network procedures of GPRS Mobility Management (GMM) and Session Management (SM) are covered in **Chapter 6** and **Chapter 7,** respectively. In particular, Chapter 6 covers important GMM procedures including Attach/Detach, Routing Area Update, Authentication, and Combined Attach/Routing Area Update procedures. Chapter 7 covers SM procedures including PDP Context Management procedures. It also covers activation, modification and deactivation of PDP Context.

Chapter 8 provides a summarized view of GPRS user plane. The role of key user-plane protocols including Subnetwork Dependent Convergence Protocol (SNDCP), Logical Link Control (LLC), BSS GPRS Protocol (BSSGP), Network Service (NS) and GPRS Tunelling Protocol for User Plan (GTP-U) are covered in this chapter. Other user plane protocols like RLC/MAC are not discussed as they are already covered earlier.

Chapter 9 is the concluding chapter of the book that describes the enhancements done in GPRS to support Enhanced Data Rates for Global Evolution (EDGE). The new modulation and coding schemes for EDGE (MCS 1 to 9) are described in this chapter. Important changes in RLC/MAC block structure and TBF establishment/release for EDGE are also described in this chapter.

Sumit Kasera

Nishit Narang

A. P. Priyanka

Acknowledgements

First off, we thank our organization Aricent, Gurgaon, India where we have our professional career. The organization has provided us a solid foundation to write technical books on different technologies in communication networks.

We remain obligated to our senior colleagues at Aricent, N. Mohanrangan, Prabhat Kumar Sahoo, Sudatta Kar, Rajinder Kumar, Ravikanth Pasumarthy, Manish Gupta and Dhrubajyoti Das for their support and guidance.

We would also like to thank all personnel at Nokia Siemens Networks (NSN) who were a part of the collaborative software project on GSM/GPRS/EDGE. Special thanks are due to Juha Matturi, Heikki Herva, William Brockington, Andy Tabor, Rade Luburic and Eric Jones for sharing their expertise that helped us become proficient with the three wireless technologies.

Our acknowledgements would not be complete without appreciating our families for their support and patience while finalizing the book.

We also thank all our readers who time and again have provided valuable feedback and motivated us to write new and better books.

Apart from above, individually, we would like to acknowledge specific persons/institutions.

Both Sumit and Nishit would thank their alma mater and all their professors for providing them a strong engineering platform. While Sumit is an alumnus of Indian Institute of Technology (IIT) Kharagpur, Nishit completed his graduation from the prestigious IIT Delhi.

Priyanka would thank good fortune and the almighty for being present in the right place at the right time. As she puts it: "Luck is when opportunity meets preparation". In particular, she would like to thank Sumit and Nishit for showing faith in her capabilities and providing her with this opportunity of being part of the authoring team. She expresses special gratitude to Sumit for his tireless mentoring and patience as he guided her through this process of book-authoring. She owes her preparation, not just for this book, but for the whole life to her parents without whom she could not justify any opportunity. She would also like to thank her sister and husband for their unconditional love, support and patience. Last, but not the least, she would like to thank all colleagues at Aricent for making each day a learning experience.

Website

To supplement the material in this book a website has been created: *http://gprsbook.tripod.com*. Readers are strongly encouraged to access this site, where they will find:

- The Preface
- Table of contents
- Errata
- Feedback and review comments
- References
- Other related material

Suggestions

We solicit reader comments, feedback and criticism that will help us add more value to the subsequent editions of this book. Please write to us at: nwbook@lycos.com.

Sumit Kasera
Nishit Narang
A. P. Priyanka

Contents

Abbreviations

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
8-PSK	Eight-state Phase-Shift Keying
AC	Address Control
AGCH	Access Grant Channel
AN	Access Network
APN	Access Point Name
ARFCN	Absolute Radio Frequency Channel Number
ARQ	Automatic Repeat reQuest
AuC	Authentication Center
BC	Bearer CHannel
BCCH	Broadcast Control CHannel
BCS	Block Check Sequence
BEC	Backward Error Correction
BEP	Bit Error Probability
BG	Border Gateway
BH	Block Header
BSC	Base Station Controller
BSIC	Base Station Identity Code
BSN	Block Sequence Number
BSS	Base Station Subsystem
BSSAP	BSS Application Part
BSSGP	BSS GPRS Protocol
BSSMAP	BSS Management Application Part
BTS	Base Transceiver Station
BVC	BSSGP Virtual Connection
BVCI	BSSGP Virtual Connection Identifier
CA	Cell Allocation
CBCH	Cell Broadcast CHannel
CBS	Cell Broadcast Service
CC	Call Control
CC	Country Code

CCCH	Common Control CHannel
CCN	Cell Change Notification
CCU	Channel Codec Unit
CDMA	Code Division Multiple Access
CGI	Cell Global Identity
CI	Cell Identity
C/I	Carrier/Interference
CM	Connection Management
CN	Core Network
CPS	Coding and Puncturing Scheme
CS	Coding Scheme
CU	Cell Update
CV	Countdown Value
D	Direction
D-AMPS	Digital-Advanced Mobile Phone Services
DL	Downlink
DLCI	Data Link Connection Identifier
DNS	Domain Name System
DRX	Discontinuous Reception
DTAP	Direct Transfer Application Part
DTM	Dual Transfer Mode
ECSD	Enhanced Circuit Switched Data
EDGE	Enhanced Data Rates for Global Evolution
EGPRS	Enhanced GPRS
EIR	Equipment Identity Register
ETSI	European Telecommunications Standards Institute
FACCH	Fast Associated Control CHannel
FBI	Final Block Indicator
FCCH	Frequency Correction Channel
FCS	Frame Check Sequence
FDMA	Frequency Division Multiple Access
FEC	Frame Error Checking/Forward Error Correction
FH	Frame Header
FN	Frame Number
FPB	First Partial Bitmap
FQDN	Fully Qualified Domain Name
FR	Frame Relay
FS	Final Segment
FSM	Finite State Machine
GERAN	GSM/EDGE Radio Access Network
GGSN	Gateway GPRS Support Node
GMM	GPRS Mobility Management
GMSC	Gateway Mobile Switching Center

GMSK	Gaussian Minimum Shift Keying
GPRS	General Packet Radio Services
GSM	Global System for Mobile Communication
GTP	GPRS Tunneling Protocol
GTP-C	GPRS Tunneling Protocol for Control Plane
GTP-U	GPRS Tunneling Protocol for User Plane
HCS	Header Check Sequence
HCS	Hierarchical Cell Structures
HLR	Home Location Register
HPLMN	Home PLMN
HSCSD	High-Speed Circuit Switched Data
ICMP	Internet Control Message Protocol
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity
IP	Internet Protocol
IR	Incremental Redundancy
IWMSC	Interworking MSC
LA	Location Area
LA	Link Adaptation
LAC	Location Area Code
LAI	Location Area Identity
LAPD	Link Access Protocol on the D channel
LAPDm	Link Access Protocol on the Dm channel
LCS	Location Services
LI	Length Indicator
LLC	Logical Link Control
LMSI	Local Mobile Station Identity
LR	Location Registration
LSP	Locally Significant Part
M	More
MAC	Medium Access Control
MAP	Mobile Application Part
MCC	Mobile Country Code
MCS	Modulation and Coding Scheme
ME	Mobile Equipment
MM	Mobility Management
MNC	Mobile Network Code
MNRG	MS Not Reachable for GPRS
MNRR	MS Not Reachable Reason
MO	Mobile Originated
MS	Mobile Station
MSC	Mobile-services Switching Center
MSIN	Mobile Subscriber Identification Number

MSISDN	Mobile Subscriber ISDN
MSRN	Mobile Station Roaming Number
MT	Mobile Termination
MT	Mobile Terminated
MTP	Message Transfer Part
NBP	Next Partial Bitmap
NDC	National Destination Code
NMO	Network Mode of Operation
NMS	Network Management System
NMSI	National Mobile Station Identity
NS	Network Service
NSAPI	Network-layer SAPI
NSC	Network Service Control sublayer
NSE	Network Service Entity
NSEI	Network Service Entity Identifier
NSS	Network and Switching Subsystem
NS-VC	Network Service Virtual Connection
NS-VCI	Network Service Virtual Connection Identifier
PACCH	Packet Associate Control CHannel
PAGCH	Packet Access Grant CHannel
PBCCH	Packet BCCH
PC	Power Control
PCCCH	Packet Common Control CHannel
PCH	Paging CHannel
PCU	Packet Control Unit
PDC	Personal Digital Communications
PDCH	Packet Data CHannel
PDN	Packet Data Network
PDP	Packet Data Protocol
PDTCH	Packet Data TCH
PDU	Protocol Data Unit
PFC	Packet Flow Context
PFI	Packet Flow Identifier
PI	Packet Idle
PI	PFI Indicator
PLL	Physical Link Layer
PLMN	Public Land Mobile Network
PNCH	Packet Notification CHannel
PPCH	Packet Paging CHannel
PPP	Point-to-Point Protocol
PR	Power Reduction
PRACH	Packet Random Access CHannel
PSI	Packet System Information

PT	Packet Transfer
PTCCH	Packet Timing advance Control CHannel
P-TMSI	Packet Temporary Mobile Subscriber Identity
PVC	Permanent Virtual Circuit
QoS	Quality of Service
R	Retry
RA	Routing Area
RACH	Random Access CHannel
RAI	Routing Area Identity
RAND	Random Number
RBB	Received Block Bitmap
RBSN	Reduced Block Sequence Number
RFL	Radio Frequency Layer
RLC	Radio Link Control
RNR	Receiver Not Ready
RR	Radio Resource
RRBP	Relative Reserved Block Period
RRC	Radio Resource Control
RTI	Radio Transaction Identifier
SACCH	Slow Associated Control CHannel
SACCH/TF	Slow Associated Control CHannel for full rate traffic channel
SACCH/TH	Slow Associated Control CHannel for half rate traffic channel
SAP	Service Access Point
SAPI	Service Access Point Identifier
SCCP	Signaling Connection Control Part
SCH	Synchronization CHannel
SDCCH	Standalone Dedicated Control CHannel
SDU	Service Data Unit
SGSN	Serving GPRS Support Node
SI	System Information
SI	Stall Indicator
SIM	Subscriber Identity Module
SM	Session Management
SME	Short Message Entity
SMS	Short Message Service
SM-SC	Short Message Service Center
SN	Subscriber Number
SNDCP	Sub-Network Dependent Convergence Protocol
SNR	Serial Number
SNS	Sub-Network Service sublayer
SNS	Sequence Number Space
SPI	Security Parameter Index
SS	Supplementary Service

TA	Timing Advance
TAC	Type Allocation Code
TAI	Timing Advance Index
TBF	Temporary Block Flow
TCAP	Transaction Capability Application Part
TCH	Traffic CHannel
TCH/F	Full Rate Traffic CHannel
TCH/H	Half Rate Traffic CHannel
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TE	Terminal Equipment
TEID	Tunnel Endpoint IDentifier
TFI	Temporary Frame Identity
TFT	Traffic Flow Template
TI	TLLI Indicator
TLLI	Temporary Logical Link Identifier
TMSI	Temporary Mobile Subscriber Identity
ToS	Type of Service
TRX	Transceiver
TS	Timeslot
TSC	Training Sequence Code
UDP	User Datagram Protocol
UL	Uplink
UMTS	Universal Mobile Telecommunication System
UNI	User to Network Interface
USF	Uplink State Flag
USSD	Unstructured Supplementary Service Data
UUS	User to User Signaling
VBS	Voice Broadcast Service
VGCS	Voice Group Call Service
VLR	Visitor Location Register
VPLMN	Visited PLMN
WS	Window Size
WWW	World Wide Web

Authors' Profiles

Sumit, Nishit and Priyanka (Left to right).

Sumit Kasera is Principal Systems Engineer at Aricent, India. He did his B Tech in Computer Science and Engineering from IIT, Kharagpur and MS in Software Systems from Birla Institute of Technology and Science, Pilani. His current area of interest is software development for GSM, GPRS and EDGE access network. He also has experience in software development for networking protocols like ATM, TCP/IP and 3G UMTS. Sumit has been an active participant in various technical forums like ITU-T, ATM Forum and 3GPP. He has presented papers and conducted seminars in these forums and also participated in the review of 3GPP specifications. Prior to this book, Sumit has authored *Communication Networks*, *ATM Networks*, *2G Networks* and *3G Networks*.

Nishit Narang is Principal Systems Engineer at Aricent, India. He did his B Tech in Computer Science and Engineering from IIT Delhi and MS in Software Systems from Birla Institute of Technology and Science, Pilani. His current areas of interest include software development for GSM/GPRS and 3G UMTS. He also has prior experience in software development for networking protocols like ATM and TCP/IP. Nishit has been an active participant in various technical forums—IEEE ACM, ATM Forum and 3GPP, where he has presented papers and conducted technical seminars. Prior to this book, Nishit has authored *2G Networks*, *3G Networks* and *Communication Networks*.

A. P. Priyanka is Senior Engineer (Testing) at Aricent, India. She did her M Sc in Informatics from Institute of Informatics and Communication, Delhi University. Currently, her areas of interest are testing GSM, GPRS and EDGE network entities. She also has experience in development of GPRS protocol stack for mobile stations and various mobile applications.

Chapter 1

GSM Overview

1.1 Introduction

Commercial mobile networks were launched in the mid-1980s. Since then the mobile communication world has been witnessing rapid changes marked by significant improvement in the services being offered. Satisfying consumer demands for better and improved services, and generating more revenue for the operator have been the areas of focus. No wonder, by the year 2002, the number of mobile users in the world had already exceeded the one billion mark. In many countries, the rate of growth of mobile subscribers has far outstripped the growth in users of fixed wireline networks. These facts indicate the high acceptability that mobile services have found with the consumers.

On the other hand, the communication world has witnessed an equally significant growth in the Internet arena. Internet popularity has grown manifolds over the same period, and therefore it comes as no surprise that these two technological marvels—the Internet and the mobile networks—are witnessing a merger of sorts. Evolution of current GSM networks to GSM/GPRS networks is a means to provide mobile subscribers access to packet networks, like the Internet over a mobile communication network. The General Packet Radio Service (GPRS) network does not exist in isolation but in conjunction with a GSM network. While the GSM network provides the conventional circuit-switched services (voice and circuit-switched data services), GPRS network provides an efficient means to support packet-data services. The GPRS network can therefore be used to provide an existing GSM subscriber an efficient mechanism to access the Internet.

Since the GPRS network is built over an existing GSM network, understanding how a GSM network operates really helps interpret the deeper principles associated with a GPRS network. This chapter therefore provides an overview of the GSM technology, which is the foundation for GPRS networks. The objective of this chapter is to familiarize the reader with some of the basic concepts of GSM, which is undoubtedly one of the most prominent second-generation mobile network technologies. The chapter begins with a brief discussion on the history of mobile networks, starting from the first-generation networks. This is followed by a general discussion on cellular concepts and then by a discussion on the GSM network concepts, which are based on these cellular concepts. The discussion around GSM focuses on the GSM network and protocol architecture, and the services that GSM offers.

1.2 History of Mobile Communication

The initial mobile networks of the mid-1980s, referred to as the first generation (1G) networks, were based on analog communication in the radio path. These networks had limited regional scope, mostly confined to national boundaries, and were an outcome of agreements between the regional telecom operator and the domestic industry. The early 1990s saw the replacement of these analog networks with the digital second generation (2G) networks. GSM became the most popular of the 2G standards. The success of 2G technologies provided the necessary thrust to mobile wireless communications and paved the way for enhanced networks in the future.

There were a couple of important changes that marked the shift from 1G to 2G standards. First, unlike the analog communication in 1G, the 2G standards were based on digital communication, both in the radio path and between network entities. Second, 2G standardization processes were aimed at making the notion of global roaming more realistic. Standardization in 1G was never elaborate, leading to national standards, which offered no roaming beyond national boundaries. However, 2G standards brought about semiglobal acceptance and, consequently, roaming to larger regions. Primarily, there are four competing 2G technologies, as explained next:

- **Global system for mobile communication.** The European Telecommunications Standards Institute (ETSI) developed the GSM specifications in 1989. The early GSM systems used a 25-MHz frequency spectrum in the 900-MHz band. This 25-MHz frequency spectrum was then divided into 124 carrier frequencies of 200-kHz each. A single 200-kHz radio channel was shared between eight users by allocating a unique time slot to each of them. The GSM is therefore viewed as a combination of two multiplexing techniques, namely frequency division multiple access (FDMA) and time division multiple access (TDMA). Of late, there are two other variants of GSM operating at the frequency of 1800 MHz (1.8 GHz) and 1900 MHz (1.9 GHz).
- **IS-95.** While most of the 2G networks were based on FDMA and TDMA, a company by the name Qualcomm designed a code division multiple access (CDMA) scheme. This scheme uses separate codes to distinguish between data transmitted by different users on the same frequency. The IS-95 is popular in South Korea and the United States apart from some other countries. This technology can be seen as a worthy competitor of the GSM technology.

- **Personal digital communications (PDC).** While a majority of countries adopted the GSM technology, PDC gained immense popularity in Japan. However, outside Japan, PDC, which operates at 800 and 1500 MHz frequencies, is not very popular.
- **US-TDMA (or D-AMPS).** Another 2G system popular in North America is the digital-advanced mobile phone services (D-AMPS). It is backward compatible with AMPS, which is one of the popular 1G mobile technology used in North America. D-AMPS is a digital version of the 1G AMPS technology.

Amongst all the 2G technologies mentioned here, GSM is the most popular 2G mobile network technology, when looking at the market share of these technologies. By the end of 2002, GSM accounted for over 66% of the world's total 2G network market.

1.3 Cellular Concepts

In any wireless communication, one of the most expensive and scarce resources is the frequency spectrum. Mobile network operators generally pay huge amount as licensing fees to obtain the frequency spectrum that is used to provide mobile networking services. Frequency spectrum being an expensive resource, it is important to design the mobile network in a fashion so as to use the frequency spectrum most efficiently.

Mobile networks are mostly modeled as cellular networks, as depicted in Fig. 1.1. A cellular network consists of a collection of geographical areas/regions, called cells. These cells can range in size from a few hundred meters to tens of kilometers in radius. Each cell has a central node that acts as the radio transceiver (transmitter and receiver), commonly also called the radio tower. The radio tower is responsible for setting up the communication channels with the mobile handsets over the radio interface. The capacity of the radio tower, in terms of the simultaneous channels it can support, should largely be consistent with the simultaneous number of mobile users who are expected to be active (i.e., in a call) within the region serviced by the radio tower.

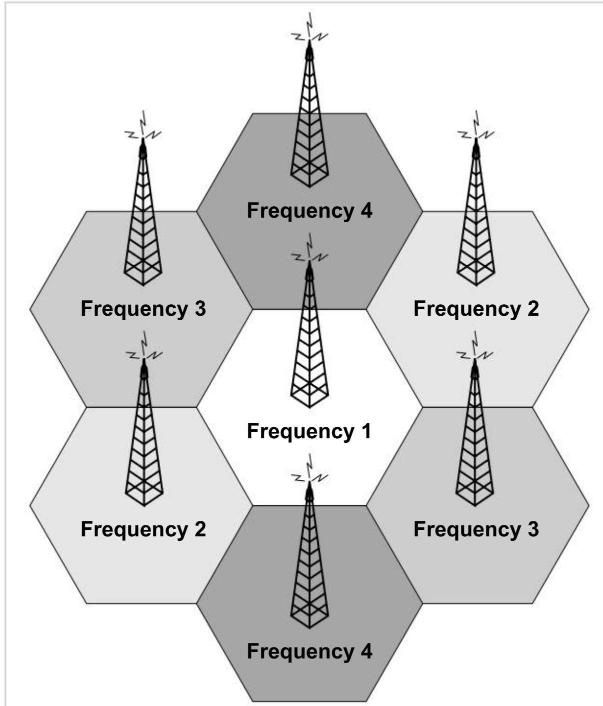
In this sense, a cell can also be defined as the region or area that is serviced by the radio tower. The size and shape of the cell would thus depend upon the reach of the signals transmitted from the radio tower. Each cell or radio tower uses a particular frequency sub-band to service the handsets within its region (see Fig. 1.1).

Designing a mobile network as a cellular network has multiple advantages as follows:

- **Reuse of frequency.** In a cellular network, the entire frequency spectrum allocated to the mobile network operator is first divided into smaller subbands. Each cell in the network uses a particular frequency subband to provide services within its geographical area. Cells are assigned distinct frequency subbands so as to prevent interference between communications happening in these cells. This is the principle behind FDMA. Further, multiple users within the same cell time-share the frequency subband allocated to the cell. Thus, while FDMA is used to prevent interference between neighboring cells, TDMA is used within a cell to serve multiple active mobile handsets.

FIGURE 1.1

Cellular Network



To make most optimum utilization of the available frequency spectrum, a cellular network allows reuse of frequency subbands in cells that are not neighbor. This is shown in Fig. 1.1, where non-neighboring cells are depicted to reuse the frequency subband while maintaining distinct frequency subbands with immediate neighbor cells.

- **Location identification for handsets.** The prime advantage of any mobile network is the mobility that it provides to users while still maintaining connectivity with the network. Here, the term connectivity basically means that the user can initiate or receive calls from the network, when he/she is moving within the network. To achieve this connectivity, it is important to know the location of the user within the mobile network so that incoming call for the user can be delivered. Cells offer the finest level of granularity to describe the current location of the mobile user within the entire mobile network.

  If the exact cell of the mobile user is known, an incoming call for him/her can be routed to the radio tower serving that particular cell. The radio tower can then setup a communication path with the mobile handset over the radio interface to deliver the incoming call. Thus, designing a mobile network as a collection of cells, offers a means to identify the location of the mobile user within the network.

  While designing a mobile network as a cellular network has multiple advantages, there are some associated disadvantages as well. Some of these are discussed next:

- **Cost of infrastructure.** As mentioned earlier, each cell has its own central radio tower. With each new cell, the infrastructure cost increases to establish the radio tower itself and also to establish the access lines from the rest of the network to the radio tower. Thus, the more the number of cells within a cellular network, the more the infrastructure cost.
- **Call handover issues.** An active mobile user involved in an ongoing call can also be mobile at the same time. Being mobile allows the mobile user to travel across multiple cells during the call lifetime. To prevent the call from being disconnected when the mobile user enters a new cell, the call must be rerouted from the current radio tower handling the call to the new radio tower. This is the principle behind call handovers. It is obvious that with smaller cells, the average number of such handover required per call will increase, thus increasing the signaling load on the network.
- **Location update signaling load.** To maintain connectivity with the mobile network, the mobile handset must indicate its current cell location to the network. This implies that the mobile handset must send a location update indication to the network, each time it changes its location to a new cell. Thus, any cellular network will need to handle this increased signaling load due to location update indications from mobile handsets.

While a cellular network has its own associated disadvantages, mobile networks are still designed as cellular networks, especially due to the advantages that such networks provide. Mobile networks, however, are designed in a manner so as to reduce the disadvantages cited here as much as possible. For example, extensive network planning is done at the time of mobile network establishment or enhancement to work out the most optimum size of the cell and the location of the radio tower. Further, to minimize the signaling load on the network due to location update indications on each cell change, a mechanism called paging is used to reduce the load. These concepts are further discussed in more detail in [GSM N. Narang], which discusses the protocols and procedures in GSM mobile networks.

1.4 GSM Network Architecture

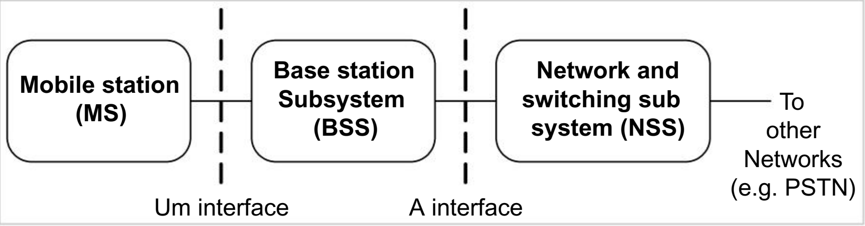
A typical GSM network can be modeled on its three basic parts or subsystems, namely mobile station (MS), base station subsystem (BSS), and the network and switching subsystem (NSS). In literature, the subsystem BSS is also sometimes referred to as the GSM access network (AN) and the subsystem NSS is also referred to as the GSM core network (CN). This terminology is also interchangeably used in this book. This basic model of the GSM network is depicted in Fig. 1.2.

The MS is the radio handset used by a subscriber/user to access the services provided by the mobile network. To connect to the network, an MS interfaces with the GSM BSS (or AN) using the GSM air interface. The GSM air interface is also referred to as the Um interface in the GSM specifications. The GSM air interface is discussed in more detail in Section 1.6.

The GSM BSS is an access network and it performs functions specific to the radio access technique used in GSM. The GSM BSS has two different types of entities: the base transceiver

FIGURE 1.2

GSM Network Architecture



station (BTS) that terminates the radio connection with the MS, and a base station controller (BSC) that controls the resources of the BTS. The BTS is the central radio tower within a cell, while the BSC is an intelligent node that controls the radio resources of the BTS. BSC and one or more BTS collectively form the GSM BSS. The BSC itself interfaces with the GSM NSS/CN over the A-interface.

The GSM NSS performs the core functions of the network, which include mobility management (MM), call control (CC), switching, and routing. The CN also manages the subscription information of a subscriber and provides services based on this information.

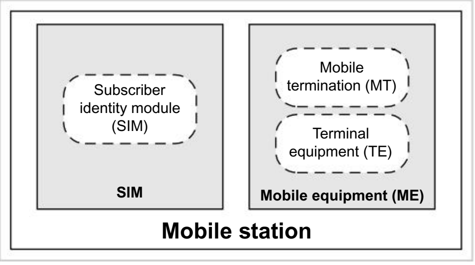
Next, the details of the MS, BSS, and NSS have been provided.

1.4.1 Mobile Station

Mobile Station is a device used by a subscriber/user to access the network services. To give it a modular design, this device is divided into two components: mobile equipment (ME) and subscriber identity module (SIM) (Fig. 1.3).

FIGURE 1.3

GSM Mobile Station



The SIM is a smart card that contains the logic required to unambiguously and securely identify the user. Thus, the SIM is the user-dependent part of the MS and is given by the service provider. It can be easily inserted into an ME or removed from it.

The SIM can contain multiple application software modules. One of these application software modules holds the logic required to identify the user. In particular, the SIM contains the permanent identity of the user (called the IMSI), the shared secret key (used for authentication), the user phone book, and a host of other information. The service provider of the mobile subscriber provides this user-dependent information (IMSI, authentication keys, etc.) contained in the SIM.

The ME, on the other hand, is the user-independent part of MS. An equipment manufacturer, who is normally an entity distinct from the service provider, manufactures the ME. The ME is further divided into two distinct logical components, namely mobile termination (MT) and terminal equipment (TE).

The MT is that part of the ME that performs functions like radio transmission termination, authentication, and mobility management. The TE component of ME manages the hardware (e.g., speaker, microphones, and user display) and end-user applications (e.g., Web browser). It may also control the mobile termination by using the modem control command set (AT commands). The division of ME into MT and TE is also referred to as MT–TE functionality split.

[3GPP TS 24.002] describes the access reference configuration and depicts the split of the ME into its two components, MT and TE. [3GPP TS 51.011] describes the SIM–ME interface. This specification defines the transmission protocols for the communication stack between the SIM and the ME. [3GPP TS 42.017] describes the SIM functional characteristics and also includes information about the mandatory and optional data stored in the SIM.

1.4.2 Base Station Subsystem

The GSM BSS resides between the MS and the GSM NSS. The GSM BSS is an access network and performs functions specific to the radio access technique used in GSM. The BSS network architecture is depicted in Fig. 1.4. The AN in the GSM, also called the BSS, comprises one BSC and one or more BTS. The primary function of BSS is to communicate with the MS in a certain area as viewed by the CN (i.e., MSC/VLR). The radio equipment of a BSS may cover one or more cells.

The BTS provides services in a cell. A BTS works on the instructions of BSC. The interface between BTS and BSC is based primarily on proprietary solutions. The important functions of BTS include channel coding, encryption/decryption on the radio interface, and transcoding and rate adaptation.

As the name suggests, the base station controller (BSC) controls one or more BTS. The primary task of BSC is to control the radio resources of GSM radio access network. The important functions of BSC include radio resource (RR) management, control of BTS, intercell handovers, and power control.

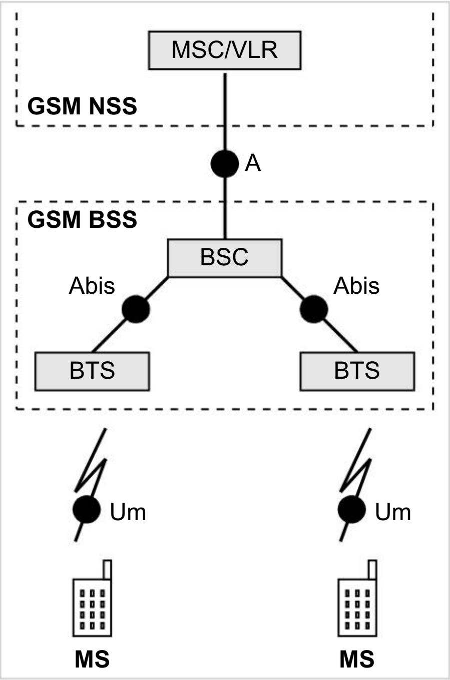
Details of the GSM BSS as well as the protocols used within the BSS are discussed in more detail in Section 1.8.

1.4.3 Network Switching Subsystem

The GSM NSS (also called the GSM core network) is essentially the heart of the GSM network, and it provides few key functions, namely mobility management, call handling, subscriber

FIGURE 1.4

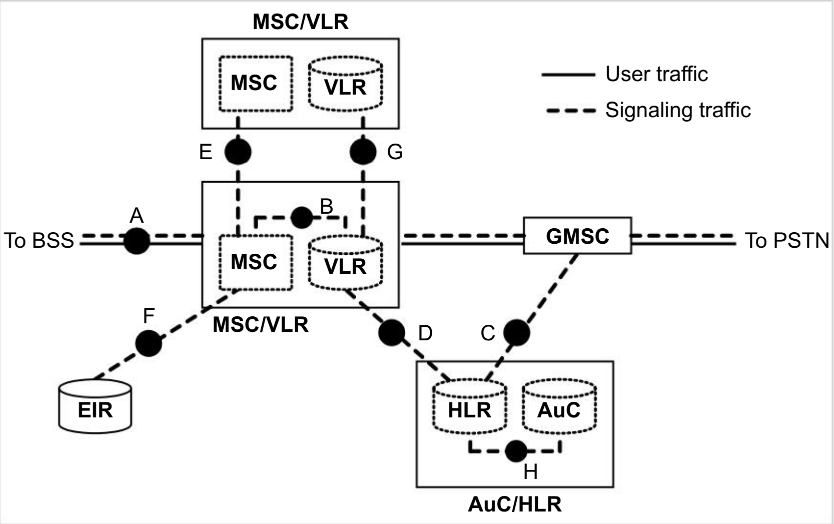
GSM BSS Network Architecture



data management, and switching. Apart from these functions, CN also provides supporting functions like authentication and equipment validation. For providing these functions, the CN has two types of entities: databases (or registers) and switches (see Fig. 1.5).

FIGURE 1.5

GSM NSS Network Architecture



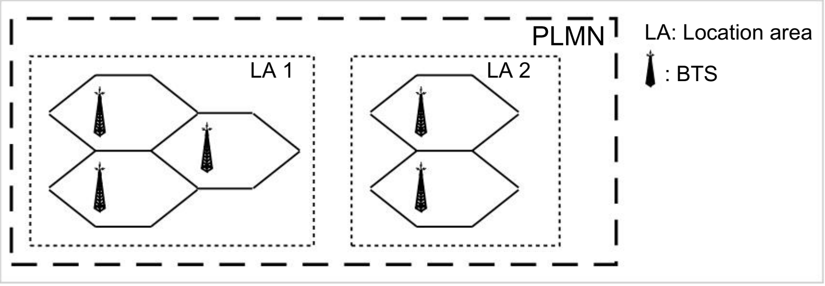
Databases within the GSM NSS include the permanent master database called the home location register (HLR), the temporary database visitor location register (VLR) catering to the subscribers that are currently visiting a mobile network, the authentication center (AuC) for subscriber authentication, and the equipment identity register (EIR) for handset verification. Switches within the GSM NSS include the mobile-services switching center (MSC) and the gateway mobile switching center (GMSC) that provide the switching services.

1.5 Hierarchical Structure of GSM Network

The GSM network architecture is organized as a multi-tier hierarchical structure. The hierarchical structure of the GSM network is depicted in Fig. 1.6. At the lowest level of the GSM hierarchy is the cell. As discussed in Section 1.3, a cell is defined as the geographical area that is served by one BTS.

Figure 1.6

Hierarchical Structure of GSM Network



At the next level of the hierarchy comes the location area (LA), which is a collection of one or more cells. The concept of LA is introduced to reduce the signaling load in the network due to location update indications from the mobile stations. As discussed in Section 1.3, location update indication from mobile stations on each cell change introduces a high signaling load on the network. To reduce this load, the procedure is modified to only send a location update indication if the MS changes an LA. As long as the MS remains within the LA, no cell change indication is required to be sent to the network. In conjunction with this modified procedure, the procedure of paging is used to identify the exact cell of the MS within the LA, in case there is an incoming call for the mobile user. The procedure of paging involves transmission of a broadcast message within the LA where the MS is known to exist. On receipt of a response to this paging message from the MS, the network gets to know the exact cell where the MS exists.

At the highest level of the hierarchy is a public land mobile network (PLMN), which consists of multiple LAs. The PLMN is the telecommunications network providing mobile services.

The hierarchical structure of the GSM network is explained in more detail next.

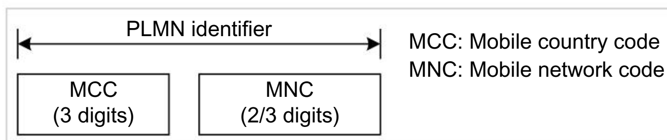
1.5.1 Public Land Mobile Network

In GSM, at the highest level of the hierarchy is a PLMN. A PLMN is defined as a telecommunications network providing mobile cellular services. A PLMN is uniquely identified by its PLMN identifier.

The PLMN identifier comprises of mobile country code (MCC) and mobile network code (MNC), as shown in Fig. 1.7. The MCC is of three digits and identifies the country to which the Public Land Mobile Network (PLMN) belongs. The next two or three digits of the PLMN identifier form the MNC. The MNC identifies a particular PLMN within a country. It is recommended that within a country identified by the MCC, all PLMN either use only two or three digits for MNC. A mixture of the two schemes is not recommended.

Figure 1.7

Structure of PLMN Identifier



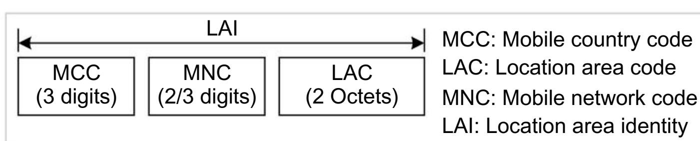
1.5.2 Location Area

Location area is defined as an area in which an MS may move freely without updating its current location with the network. In case an MS moves outside its LA, it informs the network its current location through a location update indication.

Each LA is uniquely identified by a location area identity (LAI). The structure of LAI is shown in Fig. 1.8. The MCC and MNC are the same as MCC and MNC of the PLMN to which the LA belongs. The last two octets of the LAI are the location area code (LAC) that identifies a location with a PLMN. Collectively, the LAI forms a unique identifier for an LA across all PLMNs.

Figure 1.8

Structure of Location Area Identifier



1.5.3 Cell Global Identity

At the lowest level of GSM hierarchy is the cell. Each cell is identified by the cell identity (CI). A CI is unique within an LA. To identify a cell uniquely across PLMNs, an identity called the cell global identity (CGI) is defined. CGI is obtained by the concatenation of LAI and the CI. The structure of CGI is depicted in Fig. 1.9.

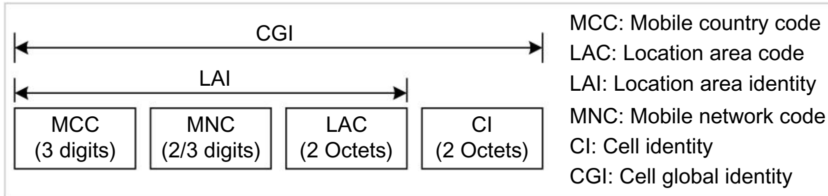


FIGURE 1.9

Structure of Cell Global Identifier

1.6 GSM AIR INTERFACE

The beauty of GSM lies in its extensive standardization whereby different parts of the network can interoperate even if they are provided by different equipment vendors. While some interfaces of the network are not completely standardized (e.g., BSC-BTS interface), the air interface is well specified with a number of companies involved in making MS and/or BTS/BSC. Owing to the interoperability requirements, an unambiguous specification of GSM air interface is vital. Apart from this, the GSM air interface (or for that matter any air interface) occupies importance owing to scarce frequency spectrum and the need to have a high spectral efficiency.

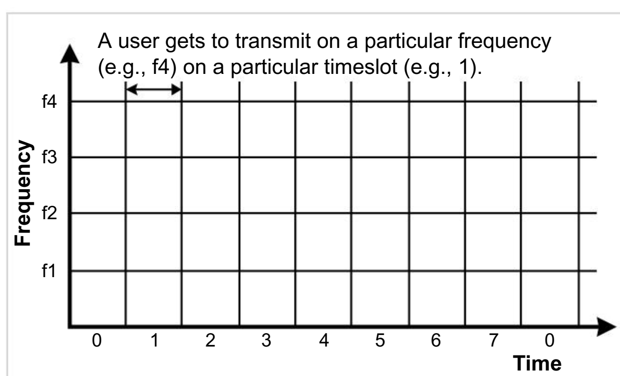
1.6.1 Multiple Access Scheme

As there are no dedicated wires in a mobile network, the wireless interface, by definition, provides a shared medium. This medium must have means to provide access to multiple subscribers. GSM uses a mix of FDMA and TDMA to provide multiple access. Here, "multiple access" refers to sharing of radio interface across different frequencies where different communication paths are simultaneously supported between different mobile stations.

More precisely, the available spectrum for GSM radio interface is separated into bands of 200 kHz. Here, the band separation refers to a separation between central frequencies of different frequency bands at which transmission takes place. As shown in Fig. 1.10, the GSM spectrum is divided into different frequencies (f1, f2, f3, and so on) in the frequency domain and into a recurring pattern of eight timeslots in the time domain numbered from 0 to 7. The

FIGURE 1.10

FDMA/TDMA Scheme Used in GSM



collection of eight timeslots that recur after a fixed time period is referred to as a TDMA frame. Each user operates on a particular frequency and on a particular timeslot. Thus, a user could use every second timeslot (timeslot 1) of frequency f4 for transmission.

The above description presents a rather simplified view of things. First of all, a user not only has to transmit but must also receive data. Thus, frequencies always exist in pairs, one for uplink (MS to BTS) and another for downlink (BTS to MS). Table 1.1 shows the uplink/downlink frequencies for different GSM systems. As seen, the downlink frequency is higher than uplink frequency because higher frequency requires higher power and the BTS can transmit at higher power as compared to MS. Different GSM system supports different number of channels. For example, GSM-900 supports 124 channels where the first uplink frequency is 890.2 MHz and the last uplink frequency is 914.8 MHz. The corresponding downlink frequencies can be calculated by adding 45 MHz to the downlink frequency. Each frequency pair is assigned a unique number called the absolute radio frequency channel number (ARFCN). For example, ARFCN 55 refers to uplink 901 MHz and downlink of 946 MHz.

1.6.2 GSM Frame Hierarchy

A simplified structure of GSM air interface is a recurring TDMA frame that comprises eight timeslots (numbered modulo 8). This timeslot that recurs every TDMA frame and operates over a frequency pair is viewed as the "physical channel" that carries user/signaling information. This physical channel lasts for 15/26 ms or 576.9 µs. The information is carried at a modulation rate of 270.833 Kbps resulting in time duration of 156.25 bit duration, where a bit duration (also called symbol duration) is 48/13 µs. Given that duration of a physical channel is not a whole number, some wonder how a fraction of information can be carried? The answer to this is that the guard period that forms part of the channel is represented in time and not in terms of the actual information carried. The information contents are always represented in terms of integral number of bits.

The physical content of a timeslot is called a *burst*. There are different types of bursts (normal, access, frequency correction, and synchronization), and each burst has a different structure. Bursts then form the lowest level of GSM frame hierarchy as shown in Fig. 1.11. Each burst directly maps to a timeslot. The TDMA frame, which has eight timeslots, forms the next level in the hierarchy.

At the next level, there is a collection of TDMA frames with two distinct flavors: 26-frame and 51-frame multiframes. The 26-frame multiframe is used for traffic channels where 24 channels are used for actual traffic exchange, one is used for low-rate signaling that accompanies any traffic exchange, and one channel is left idle (to give mobile some rest). The 51-frame multiframe is used for signaling.

The difference at multiframe level is removed at superframe level where, irrespective of type of multiframe, a superframe always has 1326 frames (26 x 51 or 51 x 26). The last level of the hierarchy is the hyperframe, which comprises 2048 superframes. Thus, the GSM frame hierarchy repeats itself after an interval of 3 hr 28 min 53 sec 760 ms.

TABLE 1.1 GSM Uplink/Downlink Frequency Bands

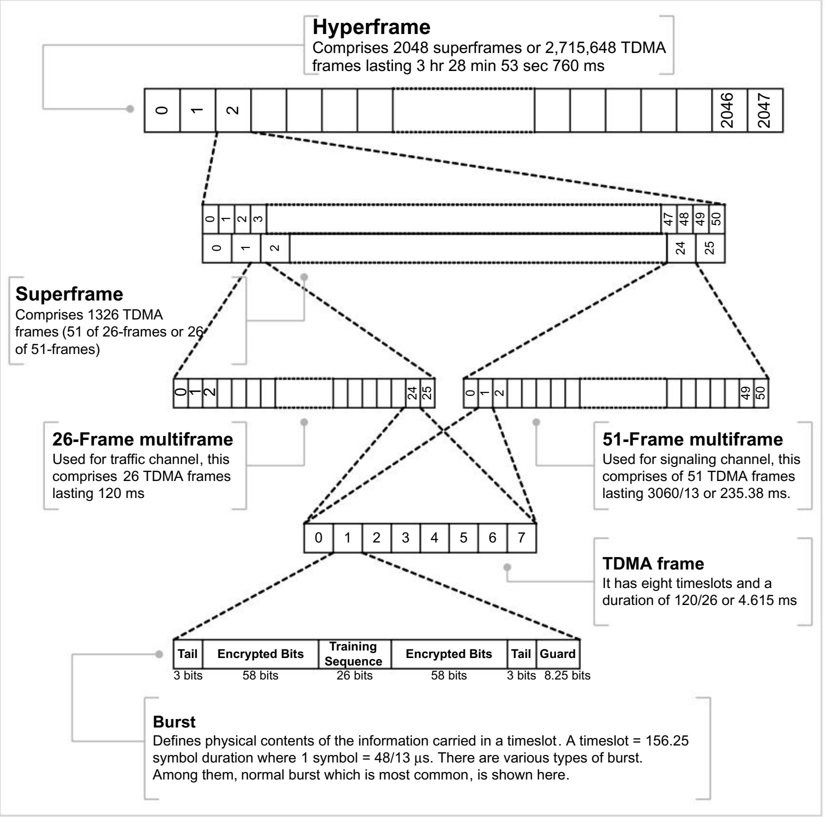
Type[*]	*Uplink (MHz)*	*Downlink (MHz)*	*Uplink frequency*	*ARFCN*
GSM-900	890–915	935–960	$F(n) = 890 + 0.2n$	$1 \le n \le 124$
GSM-900 (ext)	880–915	925–960	$F(n) = 890 + 0.2n$	$0 \le n \le 124^{+}$
			$F(n) = 890 + 0.2 \times (n-1024)$	$975 \le n \le 1\,023$
DCS-1800	1710–1785	1805–1880	$F(n) = 1710.2 + 0.2 \times (n-512)$	$512 \le n \le 885$
PCS-1900	1850–1909.6	1930–1889.6	$F(n) = 1850.2 + .2 \times (n-512)$	$512 \le n \le 810$
GSM-450	450.4–457.6	460.4–467.6	$F(n) = 450.6 + 0.2 \times (n-259)$	$259 \le n \le 293$
GSM-480	478.8–486	488.8–496	$F(n) = 479 + 0.2 \times (n-306)$	$306 \le n \le 340$
GSM-850	824–849	869–894	$F(n) = 824.2 + 0.2 \times (n-128)$	$128 \le n \le 251$

*: Frequency band for Railway-GSM not shown in the table.

+: For extended 900 band, n = 0 can be used as (890, 915) pair is a part of the available frequency.

FIGURE 1.11

GSM Frame Hierarchy



1.6.3 GSM Bursts and their Types

In a very simple TDM scheme, each timeslot contains a bit or a set of bits. However, the information carried on a timeslot of the GSM air interface is an elaborate structure. This structure, or the physical content of the information carried in a timeslot, is called a burst. A burst lasts for 156.25 bit duration or a period of 577 µs. During this period, the signal reaches from amplitude of 0 to a specified value. Thereafter, the signal is modulated and transmitted. At the end of transmission, the amplitude reaches back to 0. Here, the transmitted signal is viewed to be the modulated signal obtained after modulating a signal that comprises a series of 1's followed by the information to be sent which is again followed by a series of 1's. To ensure efficient demodulation, three guard bits (all 0's) are added before and after the information contents to ensure transition of the signal (from 1 to 0 and back from 0 to 1) during the transfer.

The exact structure of a burst depends upon the application of the burst. For all common applications that include exchange of signaling and user traffic, the normal burst is used (see Section 1.6.3.1). Then there is the access burst used by mobile to make the first attempt to communicate (see Section 1.6.3.2). However, before these bursts can be used, the BTS uses two special bursts called the frequency correction burst (see Section 1.6.3.3) and synchronization burst (see Section 1.6.3.4) to allow the mobile to synchronize in frequency and time.

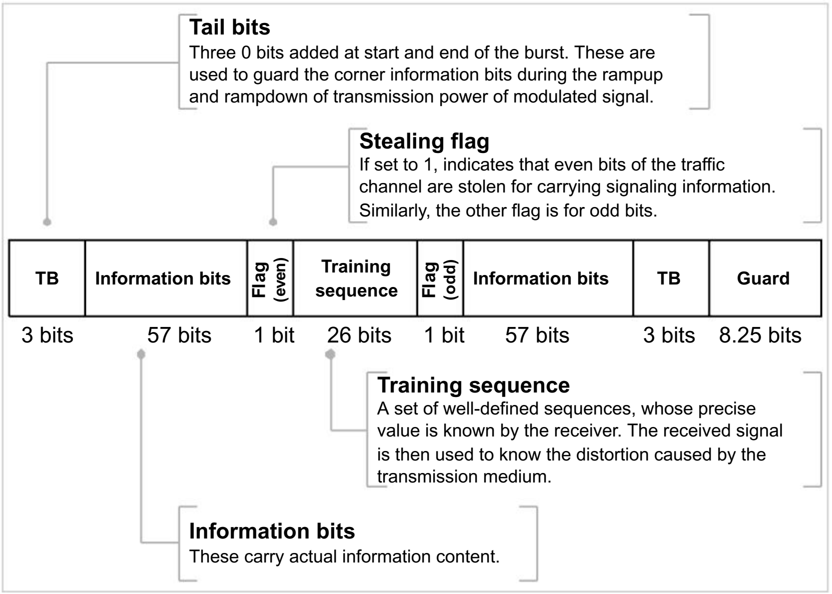
1.6.3.1 Normal Burst

The structure of the normal burst is shown in Fig. 1.12. The use of all 0 tail bits has been explained already. This is followed by the information contents, which the specifications refer to as "encrypted bits". Actually, these bits may or may not be encrypted because not all information exchanged over the air interface is encrypted (e.g., broadcast signaling messages are not encrypted). The 58-bit information is subdivided into 57 bits and a stealing flag. This flag is used when the timeslot is allocated for a traffic channel but has to be stolen for fast signaling (e.g., handover). Under such circumstances, the user traffic is discarded with temporary loss of information to carry out the high-priority signaling procedure. Now one may wonder the need for two such flags. The answer lies in the elaborate interleaving procedure that allows two different messages of a user to be carried in the same burst (one occupying odd positions and other occupying even positions). Thus, two flags are used to explicitly communicate which set of bits (odd or even) is stolen.

Another important field used in normal burst is the 26-bit training sequence. The receiver knows the value of this field a priori (through signaling). Thus, based on the signal that is actually received (as against the signal that is expected), the receiver can estimate the distortion caused by the transmitted medium. Given this, the receiver knows the transmitted training sequence, the received training sequence, and the received information signal. Based

Figure 1.12

Normal Burst



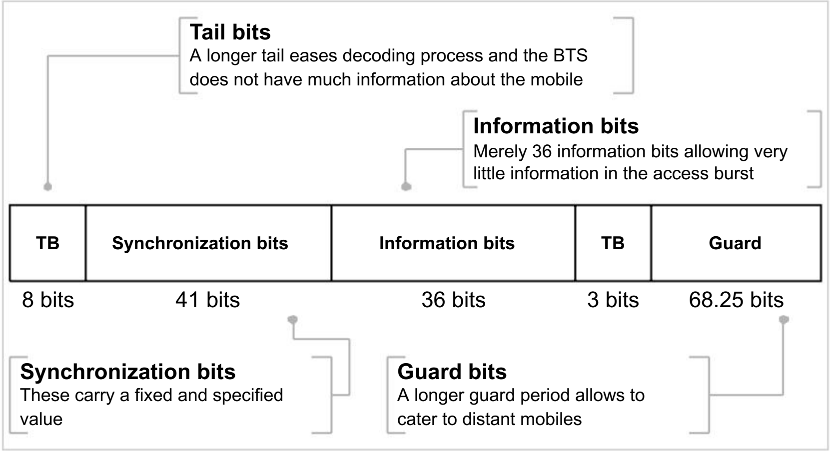
on these inputs, the receiver can predict the information that was presumably sent by the sender. The training sequence is also used to find the exact location of the useful information within the expected reception window. Since it lies in the middle of the burst, it is also called midamble. The benefit of keeping it in the center is that the distortion caused could be assumed to be uniform across the whole burst, as against keeping it in the beginning or at the end. It may be noted that even though thousands of values are possible for this field, only eight sequences are valid sequences. These sequences have been chosen in such a way that they have low correlation between one another. This helps in distinguishing signals coming from cells using same frequencies (but different training sequences). An important property of how normal burst is used is evident from the low payload that it provides. As 114 bits are inadequate to carry user information, 4 normal bursts are used to carry a message. This set of four bursts is referred to as a block. Examples of how these blocks are organized will come later in this chapter.

1.6.3.2 Access Burst

Owing to the synchronized nature of GSM air interface, there are special requirements when the mobile first tries to communicate with the network. The key challenge is to find the distance of the mobile from the base station in such a way that the mobile can compensate for delays in propagation of the signal till the base station. To help mobile find this distance, a special burst is defined called the access burst. There are certain peculiarities of this burst. First of all, its useful contents are much shorter than normal burst to allow for proposal delays (see Fig. 1.13). Thus, only 36 information bits can be carried, this too after adding redundancy. The original information thus can be merely 8 or 11 bits. Thus, when the mobile first attempts to access the network, it does not provide its own identity; instead, it seeks a channel on which it can send further information needed for authentication and resource allocation. Another peculiarity is the long guard period that allows distant mobiles to make first attempt and yet fall within the reception window of the base station. There is no training sequence in this burst to reduce signaling complexity.

FIGURE 1.13

Access Burst



The access burst is used to determine the distance of the mobile from the base station. This scenario arises in two basic cases. First, this happens when mobile makes first attempt for signaling (e.g., location update) or for communication (e.g., mobile-originating call or SMS). The second scenario happens when mobile moves from one cell to another (i.e., handover). In this case too, the first burst to the new base station is an access burst where it is important to synchronize the mobile and the base station.

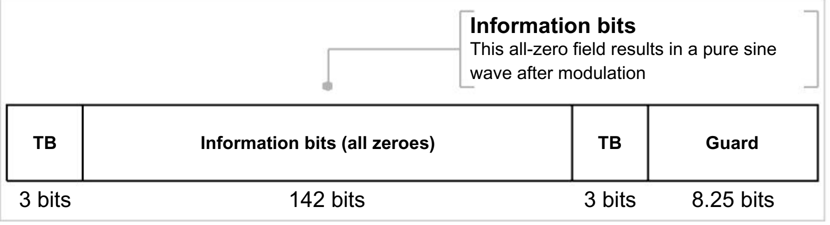
1.6.3.3 Frequency Correction Burst

This is the simplest of the burst in which the only field present is the all zero 142-bits information field (see Fig. 1.14). The result is a pure sine wave that allows the mobile to find the broadcast frequency (also referred to as the broadcast control channel [BCCH] frequency). The importance of the broadcast frequency is that all operations of the mobile follow the detection of the broadcast frequency. Once the BCCH frequency is detected, the mobile can then read the contents of the synchronization burst, which always follows the frequency correction burst.

The frequency correction burst is used by what is called the frequency correction channel (FCCH) on the broadcast frequency.

Figure 1.14

Frequency Correction Burst



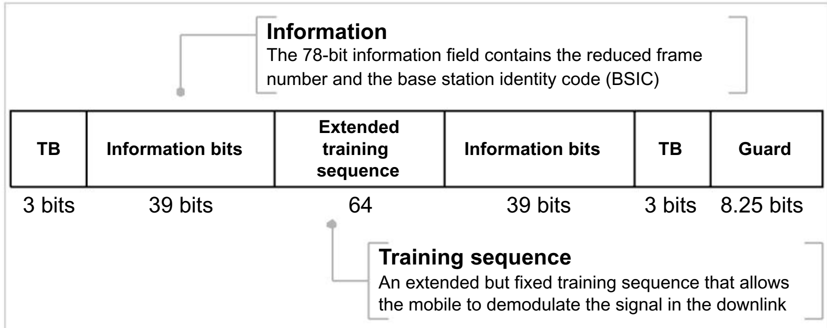
1.6.3.4 Synchronization Burst

The synchronization burst always follows the frequency correction burst (i.e., they always come in pairs) (see Fig. 1.15). While the frequency correction burst allows the mobile to detect the broadcast frequency, the synchronization burst allows the mobile to know the current counters in the GSM frame hierarchy. It is important to know the current counters so that all subsequent communication is based on the same counters. Another important function of this burst is that it carries the 6-bit Base Station Identity Code (BSIC). The synchronization burst has an extended training sequence that allows the mobile to demodulate the signal in the downlink.

The synchronization burst is used by the synchronization channel (SCH) on the broadcast frequency.

FIGURE 1.15

Synchronization Burst



1.6.4 GSM Logical Channels

The concept of burst defined in the previous section provides means to package information at the lowest layer before the signal is modulated and transmitted. However, for the purpose of allocation and management of the air interface, the burst is not adequate. Another level of abstraction is required which defines how these bursts can be used for some meaningful purposes. This abstraction in GSM is the concept of "logical channel". A logical channel can be viewed as a portion of the air interface, which has well-defined transmission characteristics (i.e., how much can be transmitted and when it can be transmitted).

Depending upon its purpose, logical channels can be categorized into two broad categories: "traffic channel" and "control channel." Different types of logical channels are depicted in Fig. 1.16.

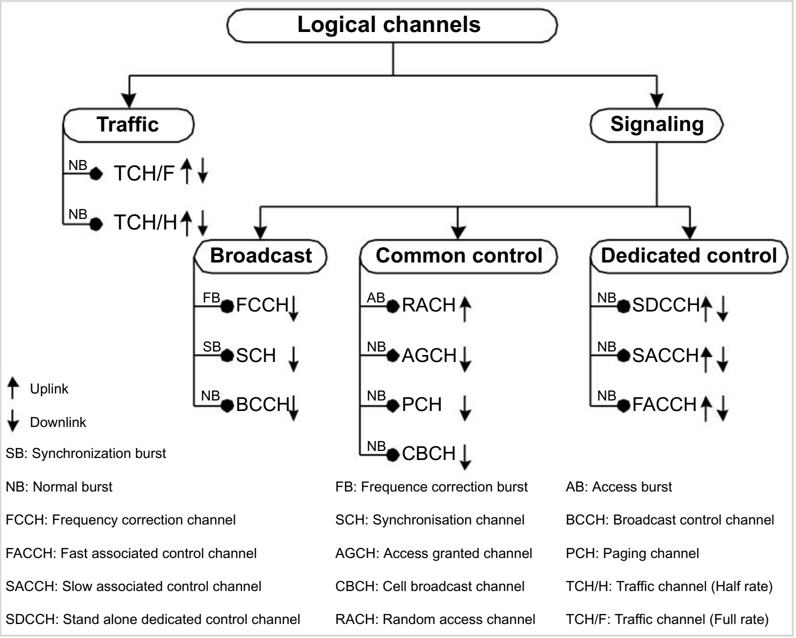
GSM traffic channels consist two types of channels as follows:

- **Full-rate traffic channel (TCH/F).** The TCH/F traffic channel follows the 26-frame multiframe structure where one timeslot is idle and one carries slow rate signaling information. This provides an available data rate of 22.8-Kbps. This 22.8 Kbps channel carried in a 26-frame is the Bm channel or the full rate traffic channel (TCH/F). The low-rate associated signaling channel is called the slow associated control channel for full-rate traffic channel (SACCH/TF).
- **Half-rate traffic channel (TCH/H).** A full-rate channel can be subdivided into two half-rate channels such that two users are allocated one timeslot. In such a case, the two users alternate on a timeslot basis. The data rate available in this case is 11.4 Kbps. Such a half-rate channel is called Lm of half-rate traffic channel (TCH/H). The associated SACCH is called SACCH/TH.

The GSM control/signaling channels are divided into three categories, namely, broadcast, common control, and dedicated control. Broadcast channels are used only in the downlink direction for broadcast of useful information to all mobiles within a cell. Common control and dedicated control channels, on the other hand, are used both in the uplink and downlink. The difference between common control and dedicated control channels stems from the fact that while common control channels are free for use by all mobiles, dedicated control channels are dedicated for a particular user.

Figure 1.16

GSM Logical Channels



GSM broadcast control channels include the following channels:

- **Frequency correction channel (FCCH).** The FCCH uses the frequency correction burst for transmission and has a very simple structure that has all zero 142-bit information field. The result is a pure sine wave that allows the mobiles to find the broadcast frequency (also referred to as the BCCH frequency).
- **Synchronization channel (SCH).** The SCH always follows the FCCH (i.e., they always come in pairs). While the FCCH allows the mobile to detect the broadcast frequency, the SCH allows the mobile to know the TDMA frame number (RFN). Another component of SCH is the 6-bit BSIC. The BSIC is used to avoid ambiguity or interference, which can arise when an MS can receive SCH from two cells that have the same BCCH frequency.
- **Broadcast channel (BCCH).** While the pair of FCCH/SCH gives frequency and timing information to the MS, there are a host of parameters that the network has to convey to the MS using its services. Examples of such parameters are CI, LAI, structure of common control channels, BCCH frequency of neighboring cells, and a host of other parameters. These parameters are packaged in the form of system information messages and are carried over the BCCH channel.

GSM common control channels include the following:

- **Random access channel (RACH).** To get a dedicated channel, the MS must request the network to allocate it such a channel. Such a request made by the MS is called "random access" request. The random access request is a short request made by using the access burst on the RACH.

- **Access grant channel (AGCH).** The AGCH, as the name suggests, is used to grant (i.e., accept) the access request made on the RACH channel. A successful RACH request is necessarily followed by a dedicated channel allocation on the AGCH. The channel so allocated by the network is then exclusively used by the MS to whom the allocation is made.
- **Paging channel (PCH).** Apart from RACH and AGCH, another important common control channel is the channel used for paging called the PCH. The PCH and AGCH are collectively also referred as downlink CCCH channels. The PCH is used for the paging procedure described in Section 1.5.
- **Cell broadcast channel (CBCH).** The CBCH is the common control channel, which is optional to implement. It is used to provide Cell Broadcast Service (CBS). The CBS service provides means to broadcast a number of unacknowledged general CBS messages (short SMS-like messages) to all receivers within a particular region, called the cell broadcast areas.

The GSM dedicated control channels, unlike the common control channels, are allocated for a particular mobile station and are, therefore, dedicated for use of that mobile station. Dedicated control channels are of following types:

- **Stand alone dedicated control channel (SDCCH).** The SDCCH is the basic channel used for exchange of signaling messages between MS and the network. This may be for just signaling purposes or for establishing a traffic channel. The SDCCH channel is allocated to the MS when the MS sends an access request in the RACH, and the network allocates the same using the AGCH channel which carries the details of allocated SDCCH. The SDCCH allocation is shortlived and does not last for long. In case it is used for location update, then after the procedure is complete, the channel is released. In case the channel is used to set up a traffic channel, then also the channel is released as soon as the traffic channel is established.
- **Slow associated control channel (SACCH).** When any dedicated channel is allocated to the MS (a dedicated traffic channel or a dedicated SDCCH channel), another channel is needed for exchanging measurement information regularly. Such a need is fulfilled by SACCH. The data rate requirement for such a channel is low. Its peculiar structure where it gets a fraction of available bandwidth gives it the name “slow“ channel.
- **Fast associated control channel (FACCH).** When an MS has a traffic channel, then the only channel available for signaling is the slow SACCH channel, which needs 480 ms to deliver a complete message. In emergency situations when a message is to be delivered immediately (e.g., during handover), the SACCH is not useful. Under such a scenario, the behavior of the existing traffic channel is altered temporarily to transmit signaling instead of user data. This is the basic principle behind the FACCH. In other words, a traffic channel and the FACCH channels are one and the same channel and differ merely in the usage of the channel.

1.6.5 GSM Channel Structure

While the previous section elaborated upon different channel types, there are certain rules that govern how they are used in different timeslots. These rules define the mapping between physical channels and logical channels. Different channel combination can be used in a timeslot (TS). Among these, the most important timeslot is the TS 0 of the BCCH TRX. The BCCH information is always carried on TS 0 and has two possible channel structures:

- **Noncombined configuration on 51-frame multiframe.** In noncombined configuration, the FCCH + SCH + BCCH + CCCH channels are present on TS 0 and SDCCH/8 on TS 1. It is noncombined in the sense that the dedicated signaling channels (SDCCH) are not combined with BCCH/CCCH and thus require a separate timeslot (TS 1).
- **Combined configuration on 51-frame multiframe.** In combined configuration, the FCCH + SCH + BCCH + CCCH channels are present along with SDCCH on TS 0. It is combined in the sense that the dedicated signaling channels (SDCCH) are combined with BCCH/CCCH on the same timeslot. It may be noted that even when SDCCH is combined with CCCH on TS 0, the SDCCH/8 can occupy TS 1. In fact, there is no restriction from specifications and the SDCCH can be located on any timeslot.

While the above channel combinations are used for the signaling channels, the traffic channel has a different channel configuration, which is based on 26-frame multiframe comprising TCH and SACCH. Each of these channel combinations are explained next.

1.6.5.1 Noncombined Channel Configuration

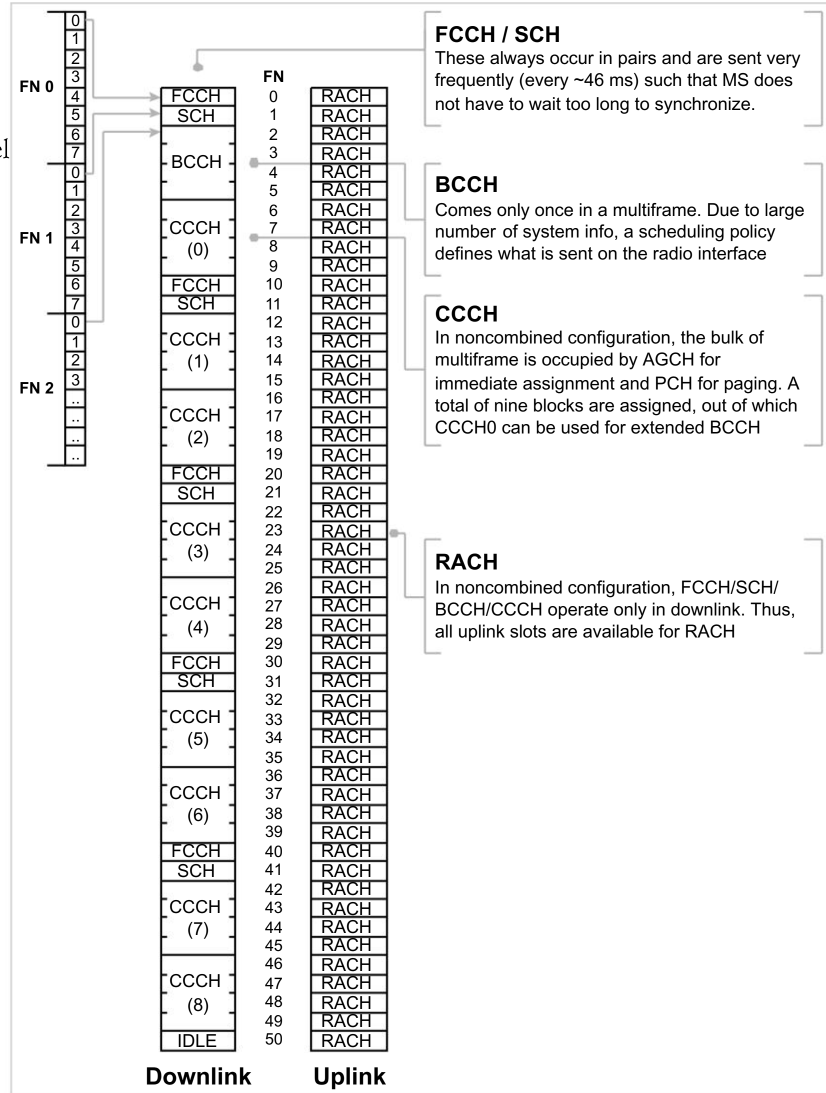
Figure 1.17 shows the structure of noncombined channel combination for TS 0. The reason for this complicated structure is that for signaling purposes, it is not prudent to allocate the entire timeslot for one particular purpose. Since each channel requires a very small portion of the bandwidth, the timeslot is organized in a well-defined structure.

In the downlink, one can see FCCH, SCH, BCCH, and CCCH. FCCH and SCH always come in pairs. Owing to their importance in easy and quick synchronization and identification of BCCH frequency, they come five times in the multiframe (0/1, 10/11, 20/21, 30/31, and 40/41). The next important component is the BCCH channel that carries the system information. It is given by the slots 2 to 5. Note that while FCCH and SCH are carried in their respective burst types, BCCH information is carried over normal burst that requires four bursts to deliver complete information (this is also true for other signaling channels like CCCH, SDCCH, FACCH, and SACCH). The last channel in this structure is the CCCH that can carry paging messages (on PCH) or immediate assignment messages (on AGCH). In noncombined mode, there is capacity to carry nine CCCH messages indicated in the figure as CCCH(0) to CCCH(9). If the system information is more than what can be sent on BCCH block, the first CCCH block is taken for system information and then referred to as extended BCCH.

The internal division of this capacity is controlled by a parameter called the BS_AG_BLKS_RES. This parameter indicates the number of blocks on each common control channel

Figure 1.17

BCCH Multiframe Structure for TS 0 in Noncombined Channel Combination



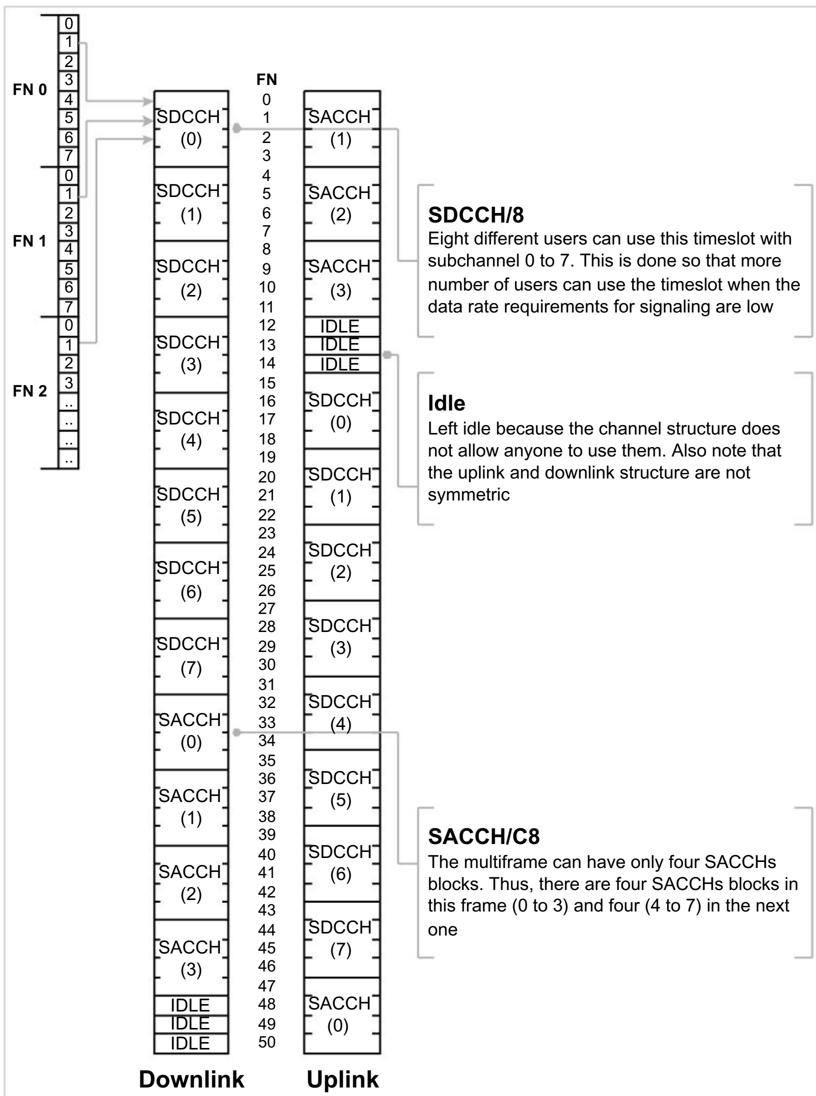
reserved for access grant messages. Its values range from 0 to 7. A value of 0 implies that none of the channels are reserved, while 7 implies that out of nine CCCH blocks, only two is available for paging. It may be noted that a value of 0 does not imply that there will be no AGCH messages. An immediate assignment message (on AGCH) can go over any CCCH, but a paging message (on PCH) cannot go over a channel reserved exclusively for AGCH.

In the uplink, all the channels are used as RACH because there are no other channels that need to go in the uplink.

Since TS 0 is for BCCH/CCCH, the SDCCH gets TS 1 in noncombined mode. The placement of SDCCH in the multiframe is shown in Fig. 1.18. A total of eight different users can share the timeslot. Due to this reason, the channel combination is also called SDCCH/8 combination (where 8 corresponds to eight subchannels, each occupied by a different SDCCH user). Each SDCCH comes along with a SACCH channel. However, in the figure, eight SDCCHs are shown but only four SACCHs are shown. This is because the rate of SDCCH is twice that of SACCH. Thus, if one takes another such multiframe, then one will see that the multiframe has SACCH(4) to SACCH(7). Thus, in two 51-frame multiframes or 102 TDMA

Figure 1.18

BCCH Multiframe Structure for TS 1 in Noncombined Channel Combination



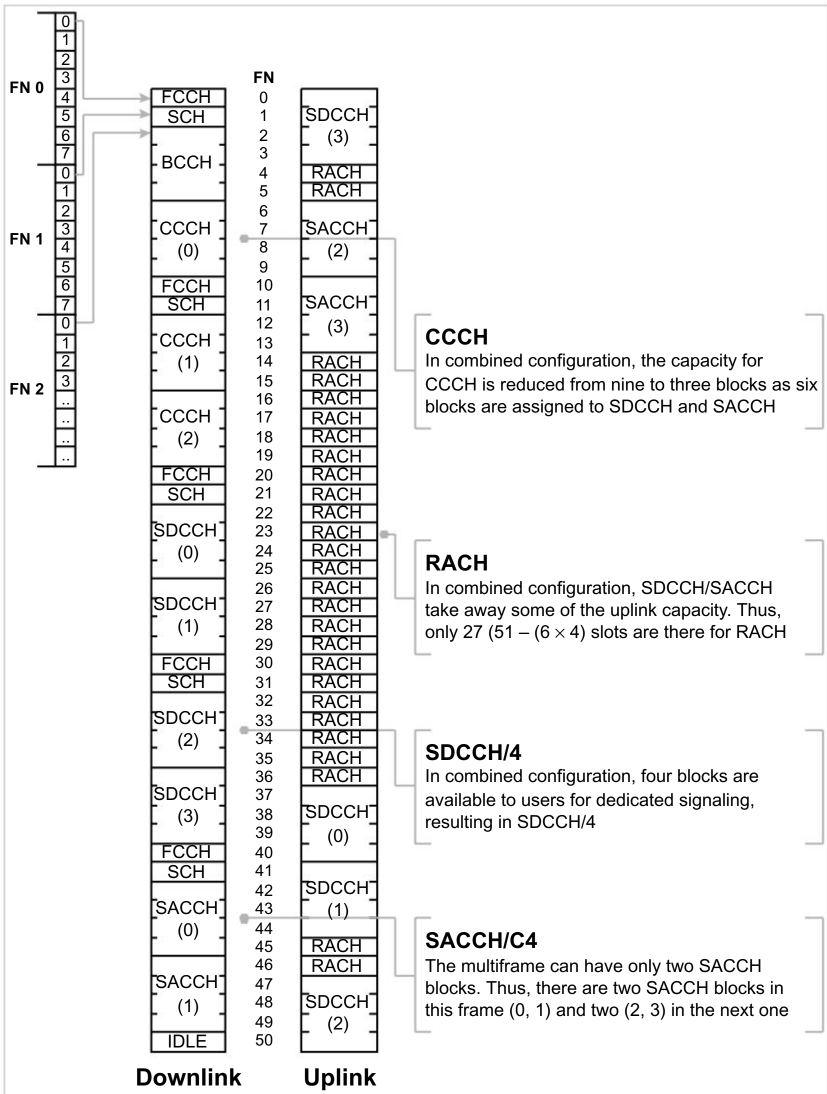
frames, there are two SDCCH blocks and one SACCH block for each SDCCH user. Another observation to be made is the rather asymmetric allocation in uplink and downlink.

1.6.5.2 Combined Channel Configuration

Figure 1.19 shows the structure of combined channel combination for TS 0. In the downlink, one can see FCCH, SCH, BCCH, CCCH, SDCCH, and SACCH. As compared to noncombined mode, the positions of FCCH/SCH (0/1, 10/11, 20/21, 30/31, and 40/41) and BCCH (2–5)

FIGURE 1.19

BCCH Multiframe Structure in Combined Channel Combination



remain unaltered. However, the CCCH capacity is significantly reduced from nine blocks to three blocks. For the remaining six blocks, four is allocated to SDCCH and two to its SACCH. Like the noncombined mode, the rate for SDCCH is twice that of SACCH. Given that four SDCCH users can share this timeslot, this combination is also called SDCCH/4 combination (where 4 corresponds to four subchannels, each occupied by a different SDCCH user).

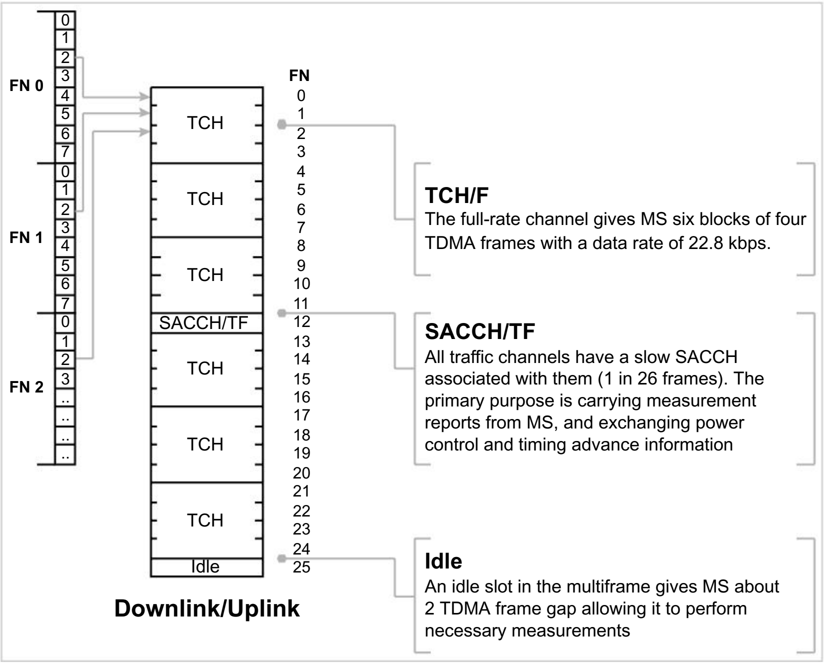
In the uplink too, RACH has to give way to the uplink slots for SDCCH and SACCH. As 24 positions are vacated, the number of RACH slots in combined mode reduces to 27 (from 51 in noncombined mode). The asymmetry in uplink and downlink allocation can be observed here too.

1.6.5.3 Traffic Channel Configuration

Figure 1.20 shows the structure of TCH/F. The structure, which comprises 8×26 slots, has duration of 120 ms and is same in uplink and downlink directions. There are three blocks of four normal bursts followed by a SACCH/Idle slot, followed further by three blocks of four normal bursts followed by an Idle/SACCH slot. As one multiframe structure has only one SACCH slot, and the SACCH carried on normal burst requires four slots, four multiframes are required to carry an SACCH message. This implies that an SACCH message is delivered in 480 ms—the precise reason why it is called a slow channel. For consecutive timeslots, one SACCH slot is on 12th frame and one is on 25th frame.

FIGURE 1.20

26-Frame Structure for Full-rate Traffic Channel



For some applications where the bandwidth requirement is low, a full-rate channel can be further divided such that two users are allocated one timeslot. In such a case, the two users alternate on a timeslot basis. The data rate available in this case is 11.4 kbps. Such a half-rate channel is called Lm of half-rate traffic channel (TCH/H). The associated SACCH is called SACCH/TH. Note that bursts cannot be divided and thus it is not possible for one burst to be shared by two users.

1.6.6 GSM Physical Layer Functions

This section elaborates upon the functional aspects of the GSM physical layer. The physical layer carries various types of information, signaling information, user speech, and user data. There are differences in the exact functions of the physical layer of these data types. Some functions are common, while some are information-specific. This section covers the common physical layer functions and as an example, covers the physical layer functions specific to full-rate speech (see Fig. 1.21).

1.6.6.1 Speech Handling Function

There are various speech handling functions relevant in the GSM system. The input is the common 8-bit Pulse Code Modulation (PCM) sample based on the popular A-law or μ-law. There is a conversion (or transcoding function that converts this input to a 13-bit uniform PCM sample at rate of 8000 samples per second. Using a speech encoder, 160 speech samples in 13-bit uniform PCM format is encoded to blocks of 260 bits. This coding considerably reduces the bandwidth requirement on the air interface.

While the input has a sampling rate of 8000 sample/sec, the encoded bit stream is of 13 kbps (260 bit samples every 20 ms). The coding scheme used in the process is the regular pulse excitation long-term prediction (RPE-LTP) with linear predictive coder (LPC) analysis. The details of the scheme are beyond the scope of this book and the reader is referred to [3GPP TS 46.010] for details.

Apart from transcoding and speech coding, few other functions are also present that have significant importance in GSM. These functions, which provide discontinuous transmission (DTX), are present because during a normal conversation, the participants alternate so that, on the average, each direction of transmission is occupied about 50% of the time. Here, DTX is a mode of operation where the transmitters are switched on only for those frames which contain useful information. DTX is important because it reduces the average interference level in the air interface and also saves battery life as MS need not transmit all the time. To support DTX, a voice activity detector (VAD) is required on the transmit side which can determine whether there is any useful voice that needs to be transmitted. Apart from VAD, another important function is "comfort noise generation." This is necessary because the background voice can suddenly disappear when transmitter does not transmit silence. In order for the receiving entity to continue similar perception of received voice, noise is inserted artificially at the receiver.

Another requirement for speech handling is to tackle lost speech frames or speech frames that are stolen by FACCH. In order to handle such scenarios, and to subside the impact of an isolated lost frame, techniques are used where the lost frame is substituted by a predictable frame based on previous data. This provides some sort of continuity for the receiver.

To know the basic architecture for handling full-rate speech and the references for associated functions like VAD and comfort noise generation, the reader is referred to [39PP TS 46.001].

It may be noted that the speech handling is primarily done at MS and at TRAU. These functions are not relevant to BTS. However, all other functions shown in Fig. 1.21 apart from speech transcoding/encoding are applicable to MS and BTS (and not to TRAU).

1.6.6.2 Channel Coding

Channel coding functionality is used to detect and correct bit errors in the received bit stream. That is, it reduces the bit error rate (BER). In order to achieve this, redundancy is added to the

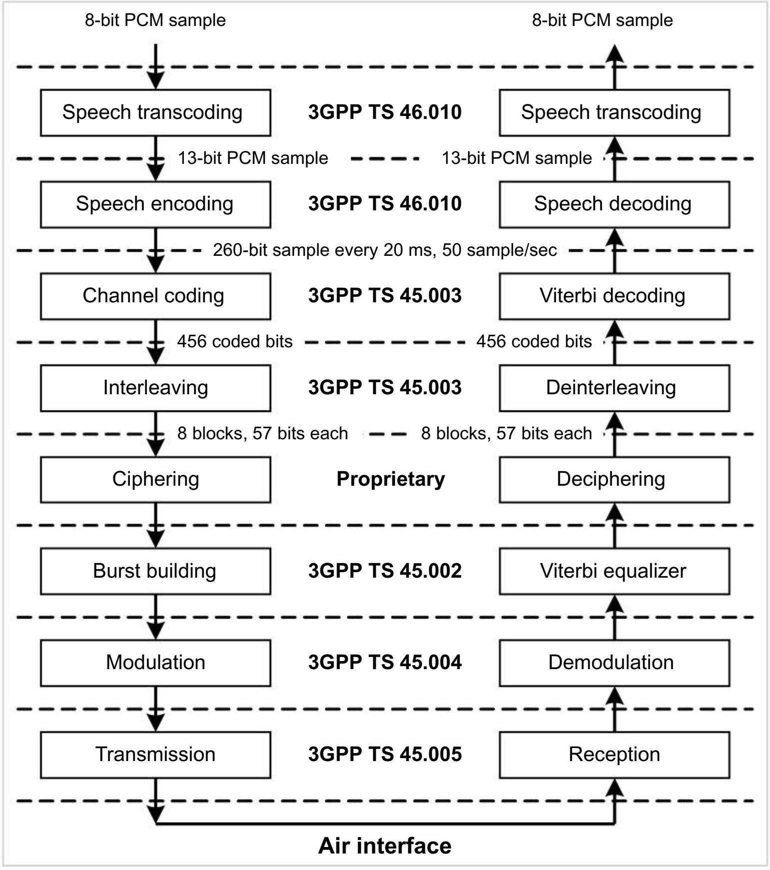


FIGURE 1.21

Physical Layer Functions with Reference to Full-rate Speech transmitted information during coding. During decoding, the redundant information is used to detect the presence of errors or to guess the most probable transmitted bits. Errors are detected when transmitted redundancy is different from the one calculated with received data.

GSM uses various forms of channel-coding schemes, including parity coding, convolution coding, and fire coding. The first two coding schemes are discussed here.

Parity codings are simple cyclic codes that are capable of detecting one-bit errors.

Convolution coding is derived from the word convolution, which means modify or distort. In other words, the coding is performed by modifying the inputs. This is done through a memory-based scheme where the transmitted bit is a convoluted sequence of input. The scheme is based on three parameters. The first two parameters are the number of input bits and the number of output bits. The ratio of input and output bits is the rate for convolution coding. For example, the full-rate speech uses 1/2 coding whereby the output of coder is double the input. The third parameter is the extent of history maintained. In other words, this represents the number of input bits used to generate an output bit. For example, the polynomials used for full-rate channel coding are $G0 = 1 + D^3 + D^4$ and $G1 = 1 + D + D^3 + D^4$. Here, D refers to delay. So, for G0, the source sequence is generated using the current source sequence (1), the source sequence delayed by three-bit period (D^3), and the source sequence delayed by four-bit periods (D^4). Similarly, the source sequence using G1 is generated and the output is then based on the output of both the sequences. At the receiver, the convolution code is decoded using the Viterbi decoder.

Figure 1.22 shows the channel-coding function for full-rate speech. As already told in previous section, the speech encoding results in sample of 260 bits that are channel coded. These 260 bits are divided into three classes based on their priority: class 1a having 50 bits which are most sensitive to bit errors, class 1b having 132 bits which are moderately sensitive to bit errors, and class 2 having 78 bits which are least sensitive to bit errors. As there are three levels of severity, the coding applied to these classes of bits also differs. Class 1a bits undergo a block coding which results in a 3-bit cyclic redundancy code (CRC) added for error detection. If on reception an error is detected, the frame is declared unusable and discarded. In such cases, an attenuated version of the previous sample is used. The class 1a bits, along with CRC, 132 class 1b bits, and 4 tail bit, are fed to a half-rate convolutional encoder resulting in 378 bits of output. The class 2 bits are then added unprotected giving a total of 456 bits in all.

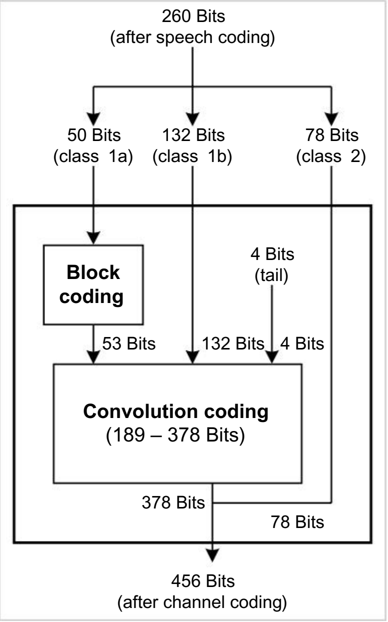
Note that while 456 bits can be sent over four bursts, with each burst carrying 114 bits each, this is not how it is done. Instead of four bursts, the bits are interleaved and spread over eight bursts. The interleaving principles are discussed in the next section.

1.6.6.3 Interleaving

The basic principle of interleaving is that it helps in spreading the effect of consecutive errors to different parts of the message such that the error correction mechanisms can be applied. Without interleaving, consecutive errors would have rendered the error correction

Figure 1.22

Channel Coding for Full-rate Speech



mechanisms useless. In order to prevent this, the message is transmitted in such a way that the consecutive bits carried in a burst are not consecutive bits of the source message. Thus, successive bit errors get distributed across different parts.

Going back to the case of full-rate speech, the 456 output bits are first spread across time such that instead of sending them as 0, 1, 2, 3, …, 455, they are arranged in eight blocks such that no two consecutive bits belong to the same block. Furthermore, these eight blocks are sent such that four blocks occupy even-numbered bits along with four blocks of previous sample that occupy odd-numbered bits; the remaining four blocks occupy odd-numbered bits along with four blocks of next sample that occupy even-numbered bits (see Fig. 1.23).

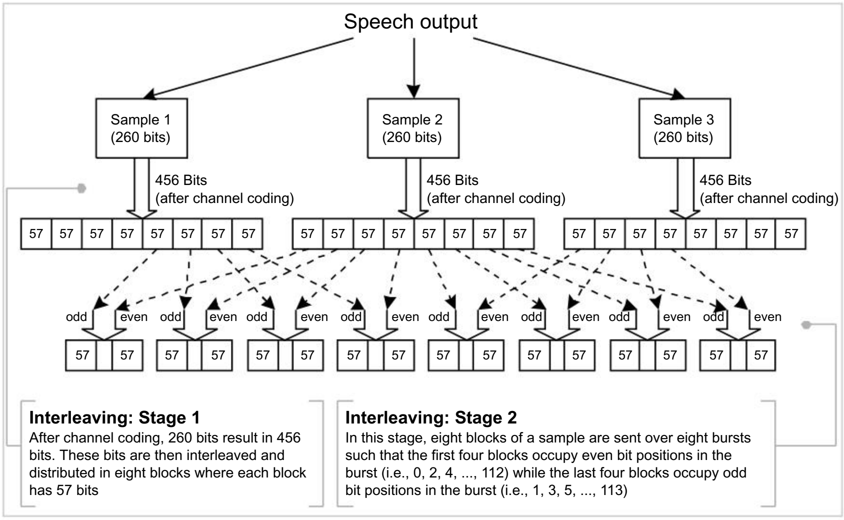
1.6.6.4 Ciphering

Ciphering and the associated authentication is a multistep process. The process typically starts when the MS performs location update procedure. Then, MSC/VLR fetches authentication data from the AuC and authenticates the MS through a challenge–response protocol. This is discussed further in Chapter 7. After the MS is authenticated, the ciphering key (Kc) is sent by MSC/VLR to BTS via the BSC. This key, and a similar key generated by MS, is used for ciphering signaling and user information over the air interface.

The process entails generating a ciphering sequence and XORing it with the user payload that is to be transmitted. The sequence is generated using an algorithm (e.g., A5/1 and A5/2) to which the current frame number (FN) and Kc are the inputs. This ciphering sequence is

FIGURE 1.23

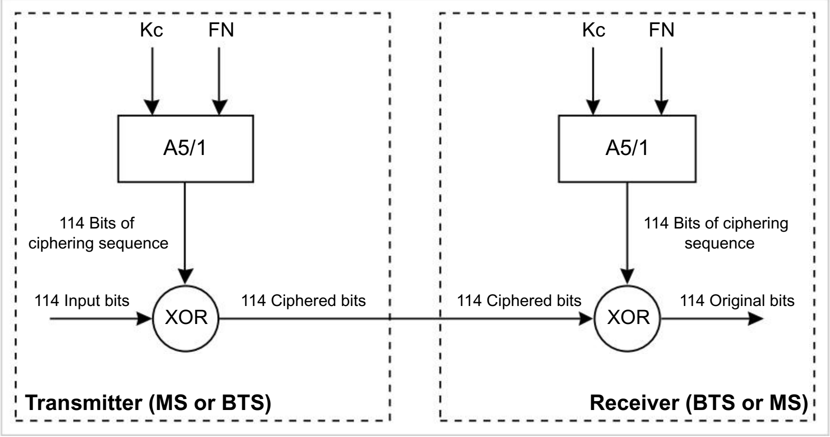
Interleaving in Full-rate Speech



XORed with the input bits to give the cipher text. At the receiving end, the same procedure is applied to give the original bits (see Fig. 1.24).

FIGURE 1.24

Ciphering and Deciphering



1.6.6.5 Modulation

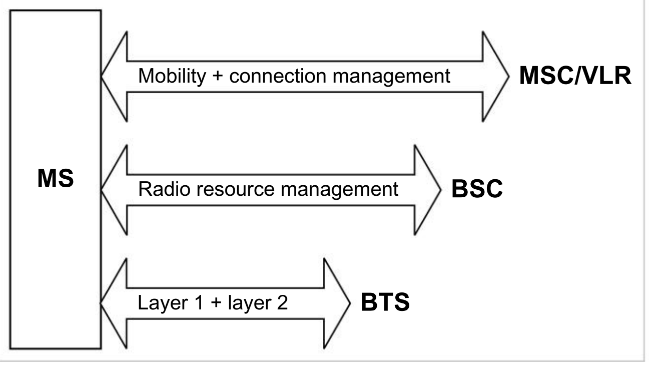
On the air interface, the signal is modulated using a carrier frequency. The modulation technique used in GSM is the Gaussian minimum shift keying (GMSK). The GMSK is a special type of minimum shift keying (MSK) which itself belongs to the class of frequency shift keying techniques. The basic idea is to use two frequencies to represent binary 0 and 1. While the details of GMSK are beyond the scope of this book, suffice to say here that GMSK was chosen over MSK because unlike MSK, which has hard shift between frequencies for 0 and 1, the GMSK has a smoother transition. Further, GMSK offers a nice balance between spectral efficiency, implementation complexity of transmitter and receiver, power consumption at the MS, extent of spurious emission, and the ability to transmit within a frequency band of 200 kHz.

1.7 GSM Communication Paradigm

The GSM protocol architecture is designed such that the MS communicates with different protocol entities at different levels of abstraction. While the detailed protocol architecture is discussed in the next section, two levels of communication are covered here, which are important in understanding the way the overall communication takes place (see Fig. 1.25). These two levels of communication are the ones that the MS has with the GSM BSS and the GSM NSS, respectively.

Figure 1.25

Protocol Layers as Seen by MS



As the MS does not have dedicated resources to communicate with the network, all communication first requires that the MS establishes an RR connection. The RR connection allocated and supervised by BSC gives MS the exclusive right to use a part of the air interface for communication. Since the establishment of RR connection would also require some means for MS and BSC to communicate, shared resources are made available to the whole community of MS to establish RR connections. Once the RR connection is established, the MS moves from idle connection state to dedicated connection state. The RR connection is not permanent but stays as long as MS is engaged in active signaling with the network (e.g., for registration with the network) or is engaged in exchange of user data (e.g., voice call). During the course of establishing the RR connection, the layer 1 (physical layer) and layer 2 (link layer) connection is established between the MS and the BTS. This is one level of communication where the MS establishes the RR connection to communicate with the GSM BSS.

Besides an RR connection, an MS also maintains a MM connection with the MSC/VLR (part of GSM NSS). The MM connection is established by the MM protocol layer in the MS and is used by the MS to communicate with the GSM NSS. As part of the MS MM layer procedures, whenever an MS comes in an area served by a given MSC/VLR, it registers with it through location update procedure. Thereafter, it periodically sends location update messages to notify MSC/VLR of its presence to the network. As long as the MS is known to the MSC/VLR, the MSC/VLR notifies the MS of any incoming activity like a mobile-terminated voice call or a mobile-terminated short message service (SMS). The MSC/VLR also allows any outgoing call to the MS in this state.

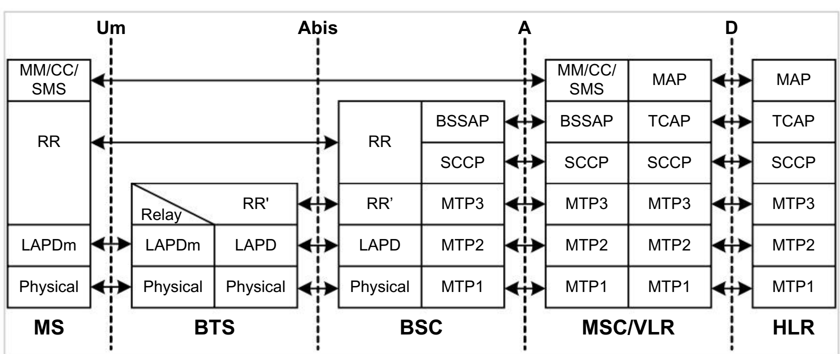
Above the MM layer in the MS is the connection management (CM) layer that is used to communicate with entities in the GSM core network, primarily for call control and supplementary service management. Thus, while MM layer in the MS is responsible for managing mobility through location update procedures, the CM layer is responsible for actual call control and supplementary services management. The MM and CM layer communication is thus the second level of communication, whereby an MS directly communicates with the GSM core network entities.

1.8 GSM Protocol Architecture

While abstract protocol architecture for GSM was discussed in previous section, this section provides a more detailed view of GSM protocol architecture as depicted in Fig. 1.26. Note that this figure is for the GSM control plane, which is used for signaling. The user plane protocol architecture is much simpler and involves only the physical and data-link layers (no higher layers are involved).

Figure 1.26

GSM Control Plane Protocol Architecture



As depicted in the figure, different protocols are used over the different interfaces defined within a GSM network. The following sections briefly discuss the protocols used on each interface of the GSM network.

1.8.1 GSM Air Interface Protocols

The physical layer of the GSM air interface provides a communication medium over the radio interface. It performs many vital functions including ciphering, channel coding/decoding, interleaving, burst formatting, modulation, and demodulation. These functions apply between the MS and the BTS. The link access protocol on the Dm channel (LAPDm) protocol also extends between the MS and BTS. It provides means to exchange message between layer 3 entities over the Dm channel of air interface. Here, the Dm channel is a term that refers to the collection of all the signaling channels required in the GSM system.

For the GSM air interface, the RR protocol is the most important protocol. This protocol allows the network (in particular, the BSC) to inform the MS in a cell about the availability of network connectivity. Apart from this, the RR is also used to maintain an RR connection that allows MS to exchange signaling and user information over the air interface. The RR protocol is specified in [3GPP TS 44.018].

1.8.2 GSM Abis Interface Protocols

The Abis physical layer typically has an E1 structure with a data rate of 2.048 Mbps. This structure comprises 32 channels with a capacity of 64 kbps each. This well-defined structure carries three different types of traffic over the Abis interface. This includes the BTS-BSC O&M signalling traffic, BTS-BSC call control signaling traffic, and the actual traffic channels carrying user data.

All signaling messages on the Abis interface are provided additional reliability through a link-layer protocol called the link access protocol on the D channel (LAPD). The main function of LAPD is flow and error control for reliable exchange of control messages over Abis interface.

The layer-3 control protocol on the Abis interface includes a BTS-BSC O&M signaling protocol and a separate protocol for BTS-BSC call control. For a BTS to work properly, the BSC must control the operations of the BTS. This control is done through functions like initialization, configuration, software download, and fault management among others. For this, an O&M signaling protocol is used between the BSC and BTS. This signaling is in turn controlled by the network administrator at the BSC through a network management system (NMS).

While BTS-BSC O&M signaling protocol is used for operational aspects, another layer-3 signaling protocol is defined for call control. This protocol enables the BSC to control the air interface by issuing appropriate commands to the BTS. Some examples of call control signalling procedures include commands for channel activation/release and for sending paging message for the paging procedure.

1.8.3 GSM A-Interface Protocols

The GSM A-interface exists between the core network and the access network (specifically between the MSC/VLR and BSS). The A-interface provides means for BSS management, call control and mobility management. The A-interface is based on the BSS application part (BSSAP) as specified in [3GPP TS 48.008]. The BSSAP is itself split into two subapplication parts. These are BSS management application part (BSSMAP) and direct transfer application part (DTAP).

The BSSMAP supports the procedures between MSC and BSS for call handling and resource management (e.g., paging procedures, reset, and handover procedures). This is the main signaling protocol between the MSC/VLR and the BSS. The DTAP essentially provides a relaying mechanism for transfer of MM and CC protocol messages across the BSS. In other words, BSS does not interpret these messages and transfers these messages transparently between MS and MSC/VLR.

The BSSAP protocol messages are carried over the SS7 protocol stack, which includes the SCCP and the MTP protocols. These protocols are not specific to mobile networks and are used in other Public Switched Telephone Networks (PSTN) as well.

1.8.4 GSM Core Network Interfaces

The main signaling protocol used within the GSM core network is the mobile application part (MAP). The MAP protocol defines signaling messages for all control procedures within the core network. This includes procedures for mobility management, call control and supplementary services management. The MAP protocol messages are themselves carried over the SS7 protocol stack.

1.9 GSM Addresses and Identities

In GSM, a number of identifiers are used for the purpose of addressing and identification. Each identifier serves a specific purpose. First comes the international mobile subscriber identity (IMSI) that uniquely identifies a subscriber. An IMSI may be associated with multiple mobile subscriber ISDN (MSISDN) numbers. The MSISDN can be viewed as the mobile phone number or the service identity. Apart from these two identifiers, there is a temporary mobile subscriber identifier (TMSI), which is used to hide the IMSI. To identify an ME, there is the international mobile station equipment identity (IMEI), which uniquely identifies an ME globally. There are other temporary identities as well whose need and function are detailed later in this section.

Then there are E.164 addresses used to identify network entities.

All these identifiers and addresses are explained in the following sections (see also Table 1.2). The reader is referred to [3GPP TS 23.003] for complete information on numbering, addressing, and identification schemes used in a GSM network.

1.9.1 Subscriber Identity

A subscriber is uniquely identified by its IMSI. The IMSI is stored in the SIM within the MS and kept hidden from ordinary access. As shown in Fig. 1.27, the IMSI is divided into three

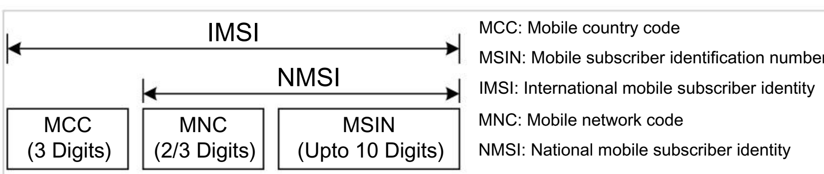
TABLE 1.2

GSM Addresses and Identifiers

Identity	*Description*	*Composition*
IMSI	Permanent identity that uniquely identifies a subscriber	MCC + MNC + MSIN
MSISDN	Service identity that is used for communication with a subscriber	CC + NDC+ SN
TMSI	Temporary identity that is used to hide the permanent identity IMSI of a subscriber	4 octets (chosen by operator)
LMSI	Temporary identity that is used by VLR to optimize database search	4 octets (allocated by VLR)
MSRN	Temporary identity that is allocated by VLR and is used to route calls directed to an MS	CC + NDC+ SN
IMEI	Permanent identity that uniquely identifies an MS	TAC + SNR
Location Number	Refers to the geographical position of the MS in terms of standardized coordinates	CC + NDC+ LSP
E.164 address	Used by MSC, GMSC, EIR, HLR, and VLR for the purpose of signaling	CC + NDC+ SN

FIGURE 1.27

Structure of IMSI



distinct parts. The first three digits of the IMSI is the MCC. The MCC identifies the country of domicile of the mobile subscriber. The next two or three digits is the MNC. The MNC identifies the home PLMN of the subscriber. The home PLMN of a subscriber is the mobile network to which the mobile subscriber is permanently associated. The last field of IMSI is the mobile subscriber identification number (MSIN). The MSIN uniquely identifies a subscriber within a PLMN. The combination of MNC and MSIN is called the national mobile subscriber identity (NMSI).

1.9.2 Service Identity

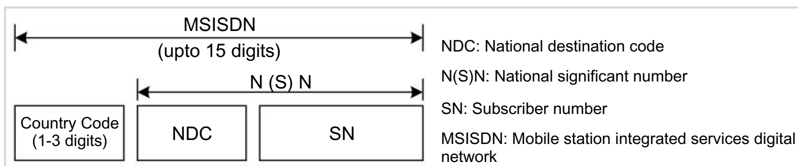
The mobile number used to contact a person is the MSISDN number and not the IMSI. Thus, an MSISDN can be viewed as a service identity because a subscriber may have multiple MSISDN, where each MSISDN identifies a particular service (voice call, fax, etc.). In other words, while the IMSI is a subscriber identity, the MSISDN is the service identity associated with the subscriber.

The MSISDN numbers are based on the ISDN numbering plan and allocated in such a manner that fixed-line ISDN or PSTN subscribers can call any mobile subscriber. The ISDN numbering plan is based on ITU-T specification E.164.

Figure 1.28 shows the structure of MSISDN. Like IMSI, an MSISDN number is composed of three distinct parts: a CC, a national destination code (NDC), and a subscriber number (SN). There is a one-to-one analogy between the elements of IMSI and MSISDN. The basic difference between the two is the number of digits allocated to individual elements. The country code is from 1 to 3 digits.

FIGURE 1.28

Structure of MSISDN



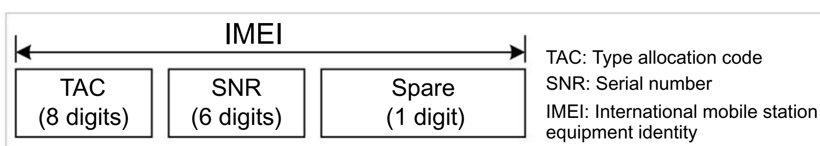
The MSISDN can be of a maximum of 15 digits. The size of national (significant) number depends upon the size of country code and can be of a maximum of 14 digits (when country code is of 1 digit).

1.9.3 Equipment Identity

An MS is identified uniquely by its IMEI. The IMEI is a 15-digit identifier and has a structure as shown in Fig. 1.29. The first eight digits form the type allocation code (TAC). The next six digits form the serial number (SNR). The last digit is spare and set to 0.

FIGURE 1.29

Structure of IMEI



The IMEI is used to uniquely identify an ME globally. It can be used to track a stolen handset. The IMEI of a handset can be known by typing the string *#06# (star hash 0 6 hash) on the MS.

1.9.4 Temporary Identities

Apart from IMSI and MSISDN, there are temporary identifiers used for specific purpose. These temporary identifiers are as follows:

- Temporary mobile subscriber identity (TMSI)

- Local mobile station identity (LMSI)
- Mobile station roaming number (MSRN)

Each of these identifiers is explained in the following sections.

1.9.4.1 Temporary Mobile Subscriber Identity

From security point of view, there is a requirement to hide the permanent identity IMSI of the subscriber. Thus, the IMSI is not frequently exchanged over the air interface. Instead, a temporary identity TMSI (and not IMSI) is used on the air interface. The TMSI is allocated by the VLR and has only local significance within the area controlled by VLR.

The TMSI consists of four octets. The exact encoding of TMSI is chosen by the agreement between the network operator and equipment manufacturer to suit local needs.

1.9.4.2 Local Mobile Station Identity

For the purpose of optimizing database search, a VLR may use a local identifier called the LMSI. The VLR sends the LMSI to the HLR during signaling message exchange along with IMSI/MSISDN. The HLR does not use the LMSI but keeps it along with IMSI/MSISDN in its database. In all further correspondence with the VLR, the HLR includes the LMSI sent earlier by the VLR. The VLR then uses the LMSI to optimize its database search.

The LMSI consists of four octets and is allocated by the VLR.

1.9.4.3 Mobile Station Roaming Number (MSRN)

In case of an incoming call for a subscriber, its associated VLR allocates a temporary number called the mobile station roaming number (MSRN). The MSRN is used to route calls directed to a MS. When a mobile-terminated call is received by GMSC, it queries the VLR (via HLR) for a number by using which it can route the call till the VLR. The VLR allocates a MSRN for the MS and passes it to HLR, which in turn forwards it to GMSC. The GMSC then uses the MSRN to route the call to the MS via MSC/VLR. Details of this procedure are covered in Section 1.10.2 when discussing call handling procedures.

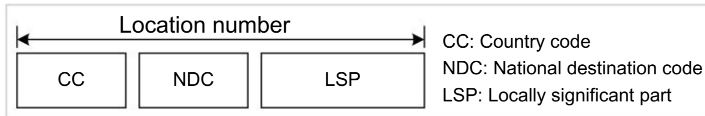
The MSRN is of the same format as the MSISDN, but is not the same as the MSISDN. The MSRN is allocated by the visited network according to the numbering plan of the visited PLMN.

1.9.5 Location Number

In Section 1.5, the location of a subscriber was defined in terms of LAs and cells. This location refers to the location of the MS in the hierarchical structure of the network. There is yet another notion of location (or position) of the MS. According to this notion, the location (or position) refers to the geographical position of the MS in terms of standardized coordinates (latitude and longitude). This information is used by application service providers to provide specialized services, also referred to as location services (LCS). An example of an LCS is a

FIGURE 1.30

Structure of Location Number



service that provides the list of restaurants around a given geographical location where the MS is located.

To provide LCS, a location number is required. The location number defines a specific location within a PLMN. The location number contains the country code, NDC, and a locally significant part (LSP). The structure of location number is depicted in Fig. 1.30. The exact structure of LSP is a matter for agreement between the PLMN operator and the national numbering authority in the country containing the PLMN.

1.9.6 Identifying Network Entities

The preceding sections elaborated upon the identifiers used for subscriber identification, service identification, and equipment identification. Apart from these identifiers, there are identifiers used for addressing network entities.

The core network entities including MSC, GMSC, EIR, HLR, and VLR are identified by using E.164 numbers. E.164 numbering format is the same format as used for MSISDN addressing (refer Section 1.9.2).

1.10 GSM PROCEDURES

Procedures in a GSM network are broadly classified into two categories, namely

- GSM BSS (or GSM access network) procedures
- GSM NSS (or GSM core network) procedures

Among the GSM BSS procedures, the RR control procedure forms the core part of the procedures. This is discussed briefly in the following section. Among the GSM NSS procedures, the MM and the CC form the core part of the procedures. These are also briefly discussed in the following sections.

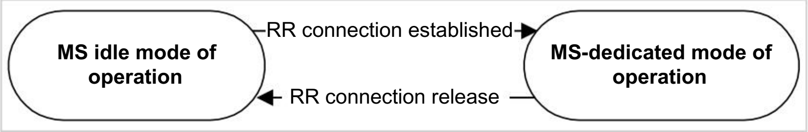
1.10.1 Radio Resource Control Procedures

The GSM BSS protocols are used to manage the radio resources between the MS and the access network. Of the GSM BSS protocols, the RR is the most important of the radio interface protocols. RR is used for signaling between the GSM network and the MS. As mentioned in Section 1.8.1, the RR protocol allows the network (in particular, the BSC) to inform the MS in a cell about the availability of network connectivity. Apart from this, the RR is also used to maintain an RR connection that allows MS to exchange signaling and user information over the air interface. RR also controls the other lower layer protocols on the radio interface.

An important aspect of the RR-protocol is the RR state machine (Fig. 1.31). The RR-state machine primarily consists of two states, namely idle and dedicated. The two states of the RR- state machine actually depict the current mode of operation of an MS. An MS operates in two basic modes, namely idle and dedicated. Idle mode, as the name suggests, is one in which the MS is not actively involved in any communication with the network and does not have resources assigned to it. In contrast, the MS has resources reserved for it in dedicated mode. For example, if an MS is switched on, then it first registers itself with the network. For this period during which it registers with the network, it stays in dedicated mode. As soon as the registration is over, it moves back to the idle mode and all resources assigned to it are released. The main procedures in idle mode include detecting the different BCCH carriers transmitting at that point of time, cell selection, and cell reselection.

Figure 1.31

RR State Machine



In dedicated mode, the MS has an RR connection with the BSC. To maintain this connection, various RR procedures are used. The most important among these includes measurement reporting. This enables the MS and BTS to report quality of connection to the BSC. Using these reports, the BSC can modify the transmission power levels of the MS or BTS (power control) or can perform handover where a new physical channel is assigned to the MS. This channel could belong to different entities resulting in different types of handover (intra-BSC, intra-MSC, or inter-MSC). The RR protocol is specified in detail in [3GPP TS 44.018].

1.10.2 Mobility Management Procedures

Mobility management procedures are used between the MS and the GSM NSS. The GSM BSS does not interpret these procedures and transparently relays messages meant for these procedures between the MS and the NSS. These procedures are implemented as part of an MM layer that its implemented at the MS and MSC. The MM procedures enable the mobility of user terminals, such as keeping track of the present location of the subscriber. It includes procedures whereby an MS can communicate its location within the GSM network to the GSM NSS. During MM, it is ensured that the identity of the user is confidential and only authenticated users can avail network services.

An important aspect of the MM layer procedures is the MM state model. In fact, the MM state model provides a summarized understanding of how the mobility information is managed by the MM layer. Figure 1.32 depicts a simplified three-state finite-state machine (FSM). In reality, however, the state machines are much more complicated than depicted here (refer to [3GPP TS 24.008] for details on FSM). However, for the sake of simplicity, a simplified and easy-to-understand picture of the state machine is presented in Fig. 1.32.

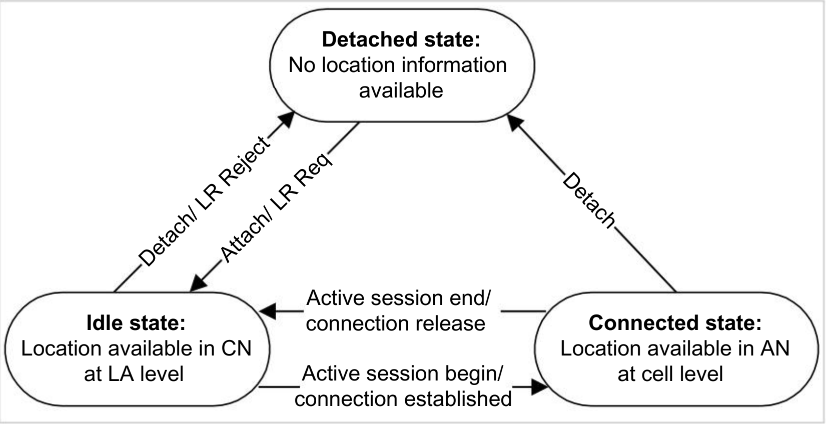
The following states are depicted in the state machine:

- **Detached state.** In this state, the network does not know the location of the MS. In other words, the MS is not reachable by the network. This may be the case when, for example, the MS is in a power-off state, or when the SIM is removed. The MS can move out from the detached state to one of the other two states by attaching itself to the network (e.g., on power-on), or by registering its location with the network through the location registration (LR) process.
- **Idle state.** In this state, the location of the MS is known to the network (the CN, to be precise, to which the MS is attached at the time). However, there is no active session for the MS (i.e., no voice call, etc., in progress). This MM-state machine corresponds to the idle mode of operation of the MS, as per the RR-state machine, and the MS is assumed to be idle in this state. From the idle state, the MS can move to the connected state when there is an active session in which it has to participate. Participation in a session is generally characterized by the establishment of a connection; hence the name. From the idle state, the MS can also move back to the detached state, for example, due to powering off, or removal of SIM from the equipment, or because of failure of the network to register the MS.
- **Connected state.** In this state, the location of the MS is known to the network (both the AN, and CN). Further, there is an ongoing session active for the MS (e.g., a voice call). Thus, in this state, the MS is not considered to be idle since it is involved in a session. This MM-state machine corresponds to the dedicated mode of operation of the MS, as per the RR-state machine. From the connected state, the MS can move to the idle state on completion of the session. It can also move back to the detached state, for example, as a result of powering off or removal of SIM from the equipment.

Besides the state transitions, Fig. 1.32 also depicts the granularity of the information available to the network as regards the location of the MS. In the detached state, no location information is available to the network since it does not know the MS in this state. Location

FIGURE 1.32

Mobility Management State Model



information is available to the network in both the connected and the idle state. However, the granularity of the location information available differs in both the states.

In the idle state, the location information of the MS is maintained by the CN (and not by the AN) only at the granularity of LA. However, in the connected state, the AN also knows the location of the MS, which is at a cell-level granularity. The logic behind this design is that the exact location of the MS is only required when there is an active session for it, and not otherwise. The MM-layer procedures provide means whereby the MS can communicate its current LA information to the GSM core network, even though it is in the idle state.

[3GPP TS 24.008] is a consolidated specification that discusses all procedures on the MS–MSC interface. The MM procedures on the MS–MSC interface can also be referred to from this specification. MM procedures in the core network are discussed in 3GPP TS [23.012]. The reader can refer to these specifications for more details on the MM procedures.

1.10.3 Call Control Procedures

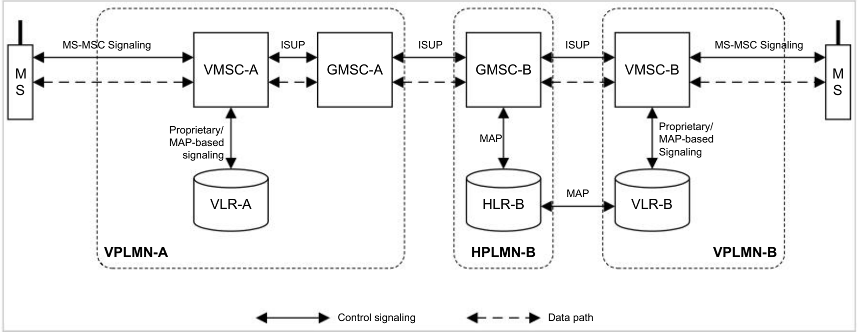
A number of protocols reside over the MM layer within the MS and the MSC. These are collectively referred to as the CM-layer protocols. This includes the CC, supplementary service (SS), and SMS protocols.

Among the CM-layer protocols, the CC protocol is the most important. The CC procedures involve handling of the mobile-originated (MO) calls and mobile-terminated (MT) calls for the subscriber. It includes verifications that need to be performed to ensure that a mobile subscriber only gets the services subscribed, and nothing more (e.g., if ISD calling is not enabled, subscriber should be barred from making ISD MO calls).

Figure 1.33 depicts the architecture for a basic mobile-to-mobile call from mobile subscriber A to mobile subscriber B. The basic mobile-to-mobile call basically contains both segments of a call, viz, MO call and MT call. Before proceeding to a discussion on this architecture, it is important to first clarify the terms HPLMN and VPLMN. The term HPLMN is used to denote the home PLMN of an MS, which maintains the subscription records and the current location information of the latter. The VPLMN is the visited PLMN where the MS is currently located/

FIGURE 1.33

Architecture for a Basic Mobile-to-mobile Call



registered. While all subscribers are allowed to roam within their HPLMN, only those with the roaming facility are allowed to visit networks (VPLMNs) outside of the HPLMN.

When discussing CC procedures, the convention followed in the GSM specifications is to use an "A" with the entities in the PLMN which initiate the call (MS-A, VMSC-A, and VLR-A) and a "B" with those that terminate the call (VMSC-B, VLR-B, and MS-B). Apart from this, the entities in the HPLMN are also suffixed with "B" (GMSC-B, HLR-B), especially since they are also participating in the MT call. This convention stems from the fact that the subscriber originating the call is called the "A" subscriber, while the subscriber terminating the call is called the "B" subscriber.

Figure 1.33 clubs together the MO and MT call segments to depict the architecture of a basic mobile-to-mobile call. In the figure, the call is shown to originate from subscriber A in VPLMN A to subscriber B, roaming in VPLMN B (the most exhaustive case)! As is evident from the figure, the call from the "A" subscriber in VPLMN-A to the "B" subscriber in VPLMN-B always goes through the GMSC of the HPLMN-B. This happens since the HPLMN-B is the only point of contact for VPLMN-A for reaching the "B" subscriber. HPLMN-B can thus be considered as the fixed part of the "B" subscriber that is known to VPLMN-A. HPLMN-B itself stores the actual/current location of the "B" subscriber. Once the call reaches GMSC-B in HPLMN-B, GMSC-B queries HLR-B as to the location of the "B" subscriber in order to determine the VPLMN of the latter. Once GMSC-B identifies the VMSC-B where the "B" subscriber is registered, it connects the call to VMSC-B. Thus, a basic mobile-to-mobile call takes the route as follows: from VMSC-A to GMSC-A, to GMSC-B, and then to VMSC-B.

A few important points that emerge from this description of the MO and MT call handling are as follows:

1. For an MO call, there is no need to involve the HPLMN (home network) of the calling subscriber. This is because all information required to handle the MO call (e.g., checks on whether or not the MS is allowed to initiate the call, etc.) is available at VLR-A (visited network). This information is transferred from HLR-A (in home network) to VLR-A (in visited network) at the time when the MS registers its presence in the visited network.
2. An MT call always involves interaction with the HPLMN (home network) of the called subscriber, since the current VPLMN (visited network) of the called subscriber is available and tracked only at its HPLMN. Hence, for GMSC-B, the HLR-B in the HPLMN is the point of contact to obtain the current location of the MS.
3. For the MT call, the HPLMN has to further interact with the VPLMN, since the call cannot be routed unless the MS is temporarily assigned some fixed ISDN number (within the VMSC-B area) for the duration of the incoming call. This temporary number is called the MSRN and is assigned to the MS by VLR-B. The call is connected from GMSC-B to VMSC-B using this MSRN (refer Section 1.9.4.3 for definition of MSRN).
4. While the exact location of the MS is known when it initiates an outgoing call, in case of an incoming call, the location of the MS is required to be exactly determined within the MSC area. For an MO call, the location of the MS is known since it is the MS that first

establishes contact with the network. However, in case of an MT call, the MS would normally be in the idle state (refer to MM-state model in Section 1.10.2), and hence its exact cell location would not be known. Paging procedures are used to determine the exact cell location of an MS for which an MT call arrives.

Paging is a mechanism whereby a request for location information is broadcasted within the LA where the mobile was last known to exist. Since the core network maintains the LA level information of the MS while in the idle state, it broadcasts a message (paging message) in this LA. On receipt of this message, the MS responds, giving its current cell location.

For more information on the CC procedures, the reader is referred to [3GPP TS 24.008], which is a consolidated specification that discusses all procedures on the MS–MSC interface. The call-handling procedures on the MS–MSC interface can also be referred to from this specification. Also, [3GPP TS 23.018] describes the basic call handling procedures within the core network. This covers core network procedures for both the MO and MT calls.

1.11 GSM Services

In GSM, the complete set of tele-services, applications, and supplementary services have been standardized. Services in GSM are classified into the following categories:

- GSM bearer services
- GSM tele-services
- GSM supplementary services
- Other bearer services

Each of these categories of services is discussed in further detail in the following sections.

1.11.1 Bearer Services

Bearer services refer to services that provide the capability of transmission of signals between two communicating entities. As the name suggests, the bearer service defines the lower layers capabilities as seen in the context of OSI reference model used in communication. Thus, bearer service provides a communication link between two entities for information transport. There is a freedom to use any higher layer protocol over the bearer services.

The GSM circuit bearer services are defined in [3GPP TS 22.002]. The services include synchronous bearer service and asynchronous bearer service. While synchronous mode supports transparent (T) data service, the asynchronous mode supports both T data service and nontransparent (NT) data service.

1.11.2 Tele-services

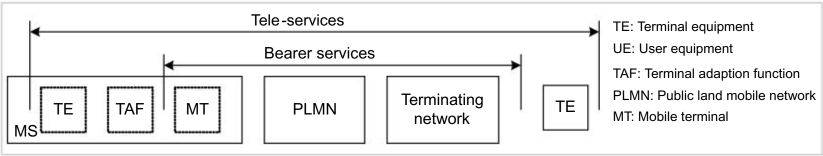
Tele-services are defined as services that provide the full capabilities for communication by means of terminal equipment and network functions. In simple terms, the scope of tele-service is not restricted to merely providing transport of user information like the bearer service. Tele-service provides complete services as seen from user's point of view.

A bearer capability is associated with every tele-service. The bearer capability defines the technical characteristics of a tele-service. Here, bearer capability merely refers to the lower layer capabilities (i.e., bearer capabilities) of the tele-service.

Figure 1.34 depicts the difference between bearer service and tele-service. As shown in the figure, while the former extends till the mobile terminal, the latter extends till the terminal equipment providing complete service to the user.

Figure 1.34

Basic Telecommunication Services



The different groups of Tele-services along with the individual tele-service as defined in [3GPP TS 22.003] are as follows:

- **Speech service.** This includes plain speech service and emergency calls.
- **Short message service.** This includes MO and MT SMS. This also includes the CBS.
- **Facsimile transmission.** This is the fax service defined for GSM. It includes the alternate speech and facsimile group-3 service and the automatic facsimile group-3 service
- **Voice group service.** This includes voice group call service (VGCS) and voice broadcast service (VBS). The VGCS enables a calling subscriber to establish a voice group call to a group of destination subscribers belonging to a predefined group call area and group identity. The group call area and group identity collectively identify a voice group call. On similar lines, VBS enables a calling subscriber to distribute speech into a predefined geographical area to all or a group of service subscribers located in this area.

1.11.3 Supplementary Services

Supplementary services are services that modify or supplement the basic services (i.e., bearer service and tele-service). Unlike the basic services that are independent in themselves, the supplementary services do not have any independent existence and are always associated with a basic service. Examples of supplementary services are call forwarding and call barring. A network operator may or may not offer supplementary services to the subscribers.

1.11.4 Other Bearer Services

GSM defines certain services that can themselves be used as bearer services. These include SMS, unstructured supplementary service data (USSD), and user-to-user signaling (UUS).

The SMS is actually a tele-service (as explained in Section 1.11.2). However, SMS can be used as a bearer over which applications can be built (e.g., a dating application). The USSD provides a mechanism whereby mobile users and PLMN operators can communicate with each other using means that are transparent to the MS and the intermediate network entities. This mechanism allows the development of services that are operator specific. The UUS is actually a supplementary services but can act as a bearer. UUS allows a subscriber to send (or receive) a limited amount of subscriber-generated information to (or from) another user in the call. This information is passed transparently through the network without any modification of the contents. The network does not try to interpret or act upon this information.

Conclusion

This chapter introduces the GSM network and the important aspects related to the GSM network, viz. GSM protocol architecture, GSM addresses and identities, and GSM procedures and services. Since the GPRS network is based on an existing GSM network, understanding of how a GSM network operates really helps interpret the deeper principles associated with a GPRS network. The overview of the GSM technology provided in this chapter should help in better understanding of the GPRS network fundamentals, which are covered in the remainder of this book.

Further Reading

While this chapter only provides an introduction to the GSM network principles, an in-depth understanding of the GSM network fundamentals would definitely help in better understanding of the GPRS principles. The reader is thus referred to the following other books, if more information on GSM is desired:

[GSM N.Narang] is a good introductory book for GSM. Readers who wish to gain more insight into the GSM technology must read this book to gain good understanding of GSM fundamentals. The book is well-organized to cover all aspects of the GSM domain, and is easy to read and understand.

[GSM M. Mouly] and [GSM G. Heine] are other good books on GSM, and can be referred for more details on GSM.

Chapter 2

GPRS Network Architecture

2.1 Introduction

Second Generation (2G) networks like GSM brought about a major change in the way mobile networks were built, but they are hampered by a number of limitations. First of all, the 2G networks provide very low data-transfer rates. This is because the 2G networks are designed primarily to offer voice services to the subscribers. These rates vary across technologies, but the average rate is of the order of tens of kilobits per second. Secondly, the 2G networks have low efficiency for packet-switched services. With the rising popularity of the Internet, there is a growing demand among the customers to access the Internet not just at home or at office but also while traveling. But wireless Internet access with the 2G networks is not efficiently implemented to serve this purpose. This is primarily because of circuit-switched nature of resource allocation, which is not suitable for bursty data transfers. In particular, the billing is based on airtime irrespective of whether the user has exchanged data or not, which is inferior to billing based on quantum of data exchanged. Thirdly, with a multitude of competing standards in place, a wireless user can roam to only those networks that support the same standard. This leads to limited roaming that can be offered to a customer and it hinders global roaming. Although 2G standards were an improvement over their 1G predecessors, they still lacked the ability to offer complete global roaming and were semiglobal in this aspect.

Thus, the growth of 2G networks was slowed down by low data-transfer rates, lack of support for packet-switched services, and semiglobal roaming facility, which could not reach global level. To overcome these drawbacks, various efforts were made to enhance the 2G networks. While proposing enhancements, the underlying objective was to make minimum changes in the existing network architecture. The outcome of these enhancements was the 2.5 Generation Networks (2.5G).

The first step toward this direction was High-Speed Circuit-Switched Data (HSCSD). To circumvent the drawback of low data-transfer rates, a simple solution is to use multiple time slots instead of one. Given that a GSM channel provides speeds of 9.6 or 14.4 Kbps, a speed of 57.6 Kbps can be obtained by using up to four channels. This is the principle of HSCSD. The advantage of this scheme is that it requires minimum changes in the network architecture. The flip side of this scheme is that it uses circuit switching, which is considered inefficient in terms of resource usage and costs. For better resource utilization, packet switching is used as exemplified in general packet radio services (GPRS).

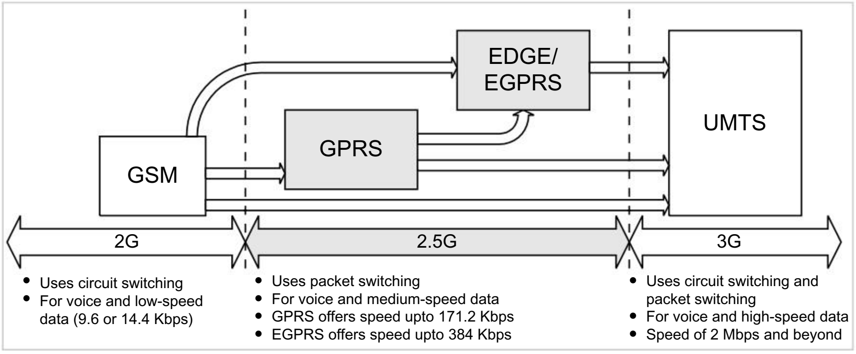
While HSCSD is viewed as an extension of GSM that was not very popular, GPRS emerged as the first genuine packet-based wireless technology. GPRS offers data services (e.g., Internet access) by using a packet-switching domain (i.e., GPRS reserves Radio Resources (RR) only when there is data to send; this provides efficient use of network resources). GPRS uses one to eight radio channels in the 200 kHz frequency to offer speeds up to 171.2 Kbps. An important advantage of GPRS is that it provides a migration step toward Third Generation (3G) networks. This is because the core network components of GPRS (e.g., Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN)), are an integral part of 3G Core network. Although there are software changes in network nodes present in 2.5G as compared to that present in 3G, the basic network architecture remains same.

The last pillar of the 2.5G networks is the Enhanced Data Rates for Global Evolution (EDGE). By using better modulation techniques, the data-transfer rates of GSM and GPRS could be increased up to three times. This improvement is called EDGE for GSM and Enhanced GPRS (EGPRS) for GPRS. Using EGPRS, a speed of up to 384 Kbps can be obtained. Chapter 9 of this book discusses some of the important aspects of EDGE.

The evolution of 2.5G networks, from 2G to 3G via 2.5G, is depicted in Fig. 2.1. A GSM operator may directly move from GSM to 3G, but this would not be the preferred route. Preferably, it may use GPRS as a stepping-stone toward 3G. Moving further from GPRS, EGPRS may also be used as an intermediate step. The exact path taken depends upon various factors including the existing infrastructure, availability of licenses, capital employed, and market demand among others.

FIGURE 2.1

Evolution Path of 2.5G Networks

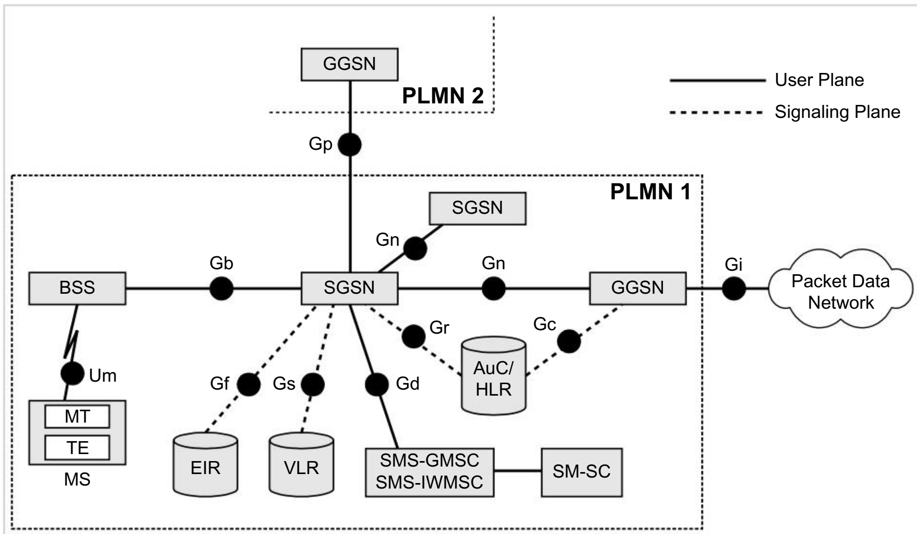


2.2 GPRS Network Architecture

A typical GPRS network can be modeled on its three basic parts or subsystems, namely Mobile Station (MS), Base Station Subsystem (BSS), and the Core Network (CN) (Fig. 2.2).

Figure 2.2

GPRS Network Architecture



The Mobile Station (MS) or the Mobile Equipment (ME) is used by a subscriber/user to access the services provided by the GPRS network. The ME can be further divided into two parts: the Mobile Termination (MT) and the Terminal Equipment (TE). While for voice applications the handset itself acts as both the MT and the TE for the user, the GPRS case may be different. For example, the handset could be connected to a laptop where the latter acts as TE while the handset provides MT.

To connect to the network, an MS interfaces with the BSS using the Um air interface. The BSS is an access network and it performs functions specific to the radio access technique used in GPRS. The BSS has two different types of entities: the Base Transceiver Station (BTS) that terminates the radio connection with the MS, and a Base Station Controller (BSC) that controls the resources of the BTS. The BTS is the central radio tower within a cell, while the BSC is an intelligent node that controls the radio resources of the BTS. BSC and one or more BTS collectively form the BSS. The BSC itself interfaces with the GSM Core Network over the A-interface. There is another logical entity called the Packet Control Unit (PCU) that typically resides along with the BSC. The PCU provides segmentation and reassembly of higher layer packets, and acknowledged or unacknowledged exchange of packets over the air interface.

GPRS CN has two types of entities: “databases” (or ”registers”) and ”routers”. While the registers like Home Location Register (HLR), Authentication Center (AuC), and Equipment

Identity Register (EIR) provide subscriber and equipment data handling functions, there are additional nodes for GPRS-specific functionality that include Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). These entities provide packet-routing and other important functions like mobility management.

2.3 Mobile Station

In GPRS, the handset (as a stand-alone entity or through a TE) provides GPRS services to the user. Based on capabilities of the MS (e.g., simultaneous use and number of time-slots), three GPRS MS Classes are defined that serve different needs of the user. These classes or GPRS modes of operation for the MS are as follows (note that the term simultaneous (attach, traffic, etc.) signifies the requirement to simultaneously support GPRS packet-switched services and GSM circuit-switched services including short message services (SMS)):

- **Class A.** In this case, the MS is attached to both GPRS and other GSM services. In other words, the MS supports simultaneous operations for GSM and GPRS including attach and voice/data transfer. Thus, a Class A MS can make calls and perform GPRS data transfer simultaneously. The heavy requirements of a Class A mobile makes it pretty expensive.
- **Class B.** In this case, the MS is attached to both GPRS and other GSM services, but it can avail any one set of services at a time. Furthermore, when a Class B MS is in both GSM idle mode and packet idle mode, it is able to monitor paging channels for both GSM and GPRS services.
- **Class C.** In this case, the MS is attached to either GPRS or other GSM services (i.e., alternate use only). If both GSM and GPRS services are supported, then a Class C MS can make and/or receive calls only from the manually or default selected service, i.e., either GPRS or GSM service. The status of the service that has not been selected is detached, i.e., not reachable.

2.4 Base Station Subsystem

The Access Network Architecture, also called the BSS, comprises one BSC and one or more BTS. The primary function of BSS is to communicate with the mobile stations in a certain area as viewed by the core network. The radio equipment of a BSS may cover one or more cells.

2.4.1 Base Transceiver Station

The BTS provides services in a cell. A BTS works on the instructions of BSC. The BTS communicates with the MS through the air interface and with the BSC through the Abis interface. The interface between BTS and BSC is based primarily on proprietary solutions. While BTS performs a number of functions for GSM, its role in GPRS is limited to hosting the Channel Codec Unit (CCU) that provides the following functions:

- Channel-coding functions, including Frame Error Checking (FEC) and interleaving

- Radio channel measurement functions, including measurements of received quality level, received signal level, and performing continuous timing advance
- For EDGE, in case of incremental redundancy mode of operation, the enhanced channel-coding functions

2.4.2 Base Station Controller

As the name suggests, the BSC controls one or more BTS. The primary task of BSC is to control the Radio Resources (RR) of GSM and GPRS Radio Access Network. For the GSM network, the important functions of BSC include RR management, control of BTS, intercell handovers, and power control. For the GPRS, the RR functionality resides in the Radio Link Control (RLC)/Medium Access Control (MAC) layer that is a part of the logical entity called the PCU.

2.4.3 Packet Control Unit

The PCU is a logical entity that manages the radio interface for the GPRS network. It is equivalent to the RR function that resides at the BSC in the GSM network. The PCU communicates with the CCU of the BTS. The PCU hosts the RLC/MAC layer and provides the following functionalities:

- Managing RR functions including broadcast channel handling, power control, and link adaptation among others
- Handling of channel requests on access channel and granting the requests on access grant channel
- Handling of uplink messages and sending Ack/Nack
- Handling of downlink messages including buffering and retransmission of RLC blocks
- Segmenting Logical Link Control (LLC) blocks into RLC blocks for downlink transmission
- Reassembly of RLC blocks into LLC blocks for uplink transmission

The PCU can reside at the BTS, BSC, or the SGSN. If not located at the BTS, it is called the "remote PCU". Relative advantages and disadvantages of PCU at various sites are discussed next.

2.4.3.1 PCU at the BTS

If PCU resides at the BTS, then the PCU and CCU interface would be an internal interface that would not require Abis, thereby saving Abis bandwidth. Further, the PCU (specifically the RLC/MAC layer) being closer to the MS, the round trip delay for the exchange of packets between the MS and the PCU would reduce considerably. However, this would imply deployment of PCU hardware at the numerous BTS sites in an existing network. In summary, while this scheme is better in terms of throughput and latency, it is costly in terms of network deployment.

To accomplish this scheme, a set of proprietary Layer 1, 2, and 3 protocols is needed on the Abis interface between the BSC and the BTS to carry the LLC payload between the PCU (at the BTS) and the BSC.

2.4.3.2 PCU at the BSC

The benefits of having PCU at the BSC are that there is an internal interface between the BSC and the PCU and the round-trip time is lesser as compared to hosting the PCU at the SGSN. The drawbacks include the higher round-trip time as compared to the PCU at the BTS, greater bandwidth requirement on the Abis, and a need for a proprietary protocol between in the PCU and the CCU that ensures synchronization. PCU at the BSC is one of the preferred options as it is in the midway between efficiency of hosting PCU at BTS and saving cost of deployment by keeping PCU at SGSN.

To make this scheme work, a set of proprietary Layer 1 and 2 protocols is needed on the Abis interface between the BSC and the BTS to carry the LLC payload between the CCU (at the BTS) and the PCU (at the BSC).

2.4.3.3 PCU at the SGSN

The benefit of having PCU at the SGSN is that there is least impact on the access network of the GSM, thereby leading to lower deployment cost and lesser number of BTS site upgrades. However, the drawbacks include the highest round-trip time as compared to other scenarios, greater bandwidth requirement on the Abis, and a need for a proprietary protocol between the PCU and the CCU that ensures synchronization and passes through an intermediate hop of BSC. Although PCU at the SGSN is a cost-efficient proposition, it is not used because of its inefficient location.

For realization of this scheme, a set of complex proprietary protocols is needed that may not be easy to accomplish.

2.5 GPRS Core Network

The GPRS core network has two types of entities: "databases" (or "registers") and "routers". The databases are used to manage and provide subscription information, authentication information, and equipment identity information. The routers are used for route packets MS to Packet Data Networks (PDN) and vice versa. The different core network entities under databases and routes are as follows:

- HLR
- AuC
- VLR
- EIR
- SGSN

- GGSN
- GPRS border gateway
- SMS entities
    - o Short Message Service Center (SM-SC)
    - o SMS Gateway MSC (SMS-GMSC)
    - o SMS Interworking MSC (SMS-IWMSC)

Each of these entities is discussed next.

2.5.1 Home Location Register

The HLR is the master database for a subscriber (i.e., it holds the subscriber data). The HLR maintains and provides subscriber data to other network entities on demand (i.e., the data is pulled by network entities). In certain cases, the subscriber data is sent to the network entities (e.g., when the data is modified by the operator, the HLR pushes it to network entities).

Majority of the subscriber data stored at HLR is provisioned by the network operator. Such data is referred to as permanent data (or static data). HLR also maintains temporary data (or dynamic data), which is obtained by HLR dynamically from other network entities. The temporary data is used for dynamic procedures (e.g., the SGSN address is a dynamic data maintained by HLR and is used for routing an incoming packet).

The following are some of the important information elements maintained by HLR for GPRS services:

- International Mobile Subscriber Identity (IMSI)
- Roaming restrictions
- SGSN address
- Barring status for various services
- Bearer service and Tele service information
- PDP Context information
- Supplementary service information
- CAMEL subscription information

[3GPP TS 23.008] provides a list of subscriber parameters along with their details. This list also specifies whether the parameter is managed by HLR or not. Based on the stored information, the HLR performs the following functions:

- **Access authorization.** This refers to allowing access to only authorized users. For example, it is checked whether the user is allowed to roam in a visited network or not.
- **GPRS mobility management.** This includes maintaining the location at which a user is registered. It also includes deleting registration information of an MS from SGSN if the MS has moved out of the area controlled by that SGSN.
- **Session management (SM) support.** The HLR supports the SM procedures. For example, during a Mobile-terminating (MT) session, HLR provides the address of SGSN where an MS is currently registered.

- **Facilitating a host of services.** HLR facilitates a host of services including SMS, supplementary service, and CAMEL service.

2.5.2 Authentication Center

The AuC holds authentication information. This information is used for authentication and other security-related functions (like ciphering). The AuC is often depicted as a part of HLR. Thus, the term AuC/HLR is used to represent the entity that performs the functions of HLR and AuC. Note that the interface between HLR and AuC is called the H-interface. This is a nonstandardized proprietary interface. Given this, the AuC is merged with HLR and it may also be referred to as HLR/AuC or simply HLR.

At the center of all AuC functions is a "secret key" that is shared between the AuC and the SIM of the UE. This secret key is associated with an IMSI and is used for authentication. It facilitates roaming for the subscriber because even in a visited network, the SGSN can authenticate the user by obtaining security information from AuC.

The AuC itself does not take part in the actual authentication. This function is performed by the SGSN. The role of the AuC is to provide SGSN with the necessary information, which can be used by the latter for authentication.

2.5.3 Visitor Location Register

Like HLR, VLR is a repository of information, which is obtained by the VLR from the HLR. This information is used by the VLR to handle GSM procedures like Mobile-originated (MO) MT calls. The VLR does not have any specific role to play in providing GPRS services. However, VLR may use the Gs interface with the SGSN to optimize some of the air interface activities. If the Gs interface is missing, there is no relevance of VLR for GPRS services. The Gs interface and importance of VLR in GPRS is discussed in Section 2.7.7.

2.5.4 Equipment Identity Register

In order to monitor the legitimacy of an MS in the GPRS network, an EIR is used. This EIR holds the list of IMEI used in the GPRS network. The IMEI is used for identifying an MS.

An MS can be classified under one of the following three categories:

- **White list.** This includes all the number series of IMEI that are permitted for use.
- **Black list.** This includes all the number series of IMEI that are barred from use.
- **Grey list.** Apart from the white and black lists, there is a grey list that is not barred. However, this is tracked by the network for various reasons (e.g., for evaluation).

The aforementioned categories determine whether an MS can be used in the network or not. While theoretically it is easy to trace stolen mobiles, practically it is difficult to accomplish. This is because a mobile stolen in one area may be used in another area or even in another country. Thus, the scheme can succeed only if there is one centralized database for all mobiles in the world.

2.5.5 Serving GPRS Support Node

The function of SGSN in the GPRS network is quite similar to that of MSC/VLR in GSM network. Thus, an SGSN maintains subscriber information obtained from the HLR and provides packet routing functions to/from GSSN. An SGSN controls one or more Routing Area(s) (RA).

The following are some of the important information elements maintained by the SGSN:

- IMSI
- One or more Packet-Temporary Mobile Subscriber Identity (P-TMSI)
- Zero or more Packet Data Protocol (PDP) addresses
- Routing Area (RA) where the MS is registered
- VLR number where the MS is registered (this is applicable only when there is a Gs interface between MSC/VLR and SGSN)
- GGSN address of each GGSN for which an active PDP context exists

SGSN performs the following functions using the information maintained by it:

- **GPRS mobility management.** The SGSN maintains the registration status of MS, performs RA updates, and obtains subscription information from HLR.
- **Session management.** SGSN helps in activation/deactivation of PDP context whereby the MS gets an address (e.g., Internet Protocol (IP) address) for communication with a Packet Data Network (PDN).
- **Security management.** SGSN also performs important security management functions including authentication and ciphering. While authentication is used to validate the identity of the MS, ciphering is performed during data transfer between the MS and SGSN.
- **Packet routing.** Apart from mobility management, the other important function of SGSN is to also delivers packet to MS from GGSN and vice versa.
- **Miscellaneous functions.** SGSN also performs other functions like charging data collection, data compression, and logical link management.

2.5.6 Gateway GPRS Support Node

The GGSN provides an interface with the PDN. It converts the GPRS packets received from SGSN into the appropriate format of the external network (typically Internet Protocol (IP) networks). In the reverse path, the GGSN converts the incoming packet to the GPRS packets and delivers it to the destined MS using the PDP context stored by it. The GGSN connects with the SGSN through an IP backbone over which the packets are tunneled, using the GPRS Tunneling Protocol (GTP). GGSN stores the current SGSN address of the user and subscriber profile in its database for various operations.

2.5.7 Border Gateway

In order to connect two PLMNs providing GPRS services, a Border Gateway (BG) is required. The BG is a gateway connecting a Public Land Mobile Network (PLMN) to an external inter-

PLMN backbone network (refer Fig. 2.4). The function of the BG is to provide the appropriate level of security for incoming and outgoing packets.

2.5.8 Short Message Service Entities

In order to provide SMS, the core network has three distinct entities. The first of these is the Short Message Service Center (SM-SC). The SM-SC is responsible for the relaying and store-and-forwarding of a short message between an MS and a Short Message Entity (SME). The SME is any entity that can send or receive SMS messages.

The SM-SC is considered to lie outside the PLMN. Thus, the functionality of SM-SC is not formally standardized by the 3GPP specifications.

Apart from SM-SC, there are two gateway entities that relay messages to and from SM-SC. These entities are SMS Gateway MSC (SMS-GMSC) and SMS Interworking MSC (SMS-IWMSC). The function of SMS-GMSC is to submit short messages to the MS from the SM-SC. While SMS-GMSC enables the delivery of a short message from SM-SC to MS, the SMS-IWMSC enables such a message to be delivered from MS to SM-SC. Collectively, SMS-GMSC and SMS-IWMSC provide the facility to deliver short messages from SM-SC to MS and from MS to SM-SC respectively.

2.6 GPRS Transmission Plane

The GPRS transmission plane consists of a layered protocol structure providing means for user information transfer, along with associated information transfer control procedures (e.g., flow control, error detection, error correction, and error recovery). As is the case with any other data network, IP has a central role to play in the GPRS network. However, given the demands of a wireless network, the GPRS does not have a simple IP routing network. In contrast, there is an intricate links of protocol that allows IP packets encapsulated in GPRS-specific protocols to travel from the MS to the GGSN from where they follow the IP routing paradigm. Within the GPRS PLMN, each GPRS hop has a set of protocol peers that allow data transfer.

The various protocol layers used in the GPRS transmission plane is depicted in Fig. 2.3. These are also summarized in Table 2.1 and explained in sections next.

2.6.1 Um Interface between MS-BSS

The first hop is over the air interface that uses the basic Time Division Multiple Access/Frequency Division Multiple Access (TDMA/FDMA) schemes as defined in GSM. While the basic scheme is retained, the logical channel structure in GPRS is completely redesigned to suit the bandwidth-on-demand requirements of GPRS. The RLC/MAC resides over the air interface. The role of RLC is to provide a reliable medium of data transfer over the air interface between the MS and the BSS. The MAC layer provides contention resolution for the air interface, and it also provides means whereby multiple users can access the shared

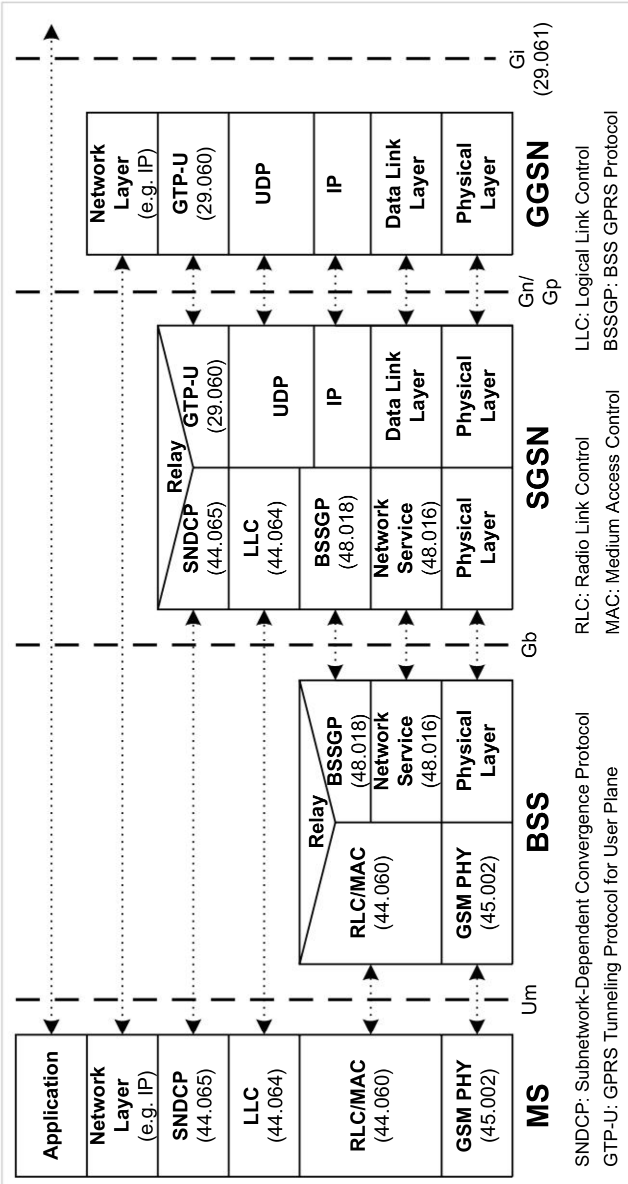


Figure 2.3 GPRS Transmission Plane

Table 2.1

GPRS Transmission Plane Interfaces

Interface	*Between*	*Description*	*Protocol*	*3GPP Specification*
Um	MS-BSS	Used between MS and BSS for exchange of user and signaling information	Physical RLC/MAC	43.064/ 45.002 44.060
Um	MS-SGSN	Used between MS and SGSN for exchange of user and signaling information	LLC SNDCP	44.064 44.065
Gb	SGSN-BSS	Used between SGSN and BSS for data transfer and mobility management	Physical Network Service BSSGP	48.014 48.016 48.018
Gn	SGSN-GGSN	Used to support mobility; applicable when GGSN and SGSN are located in same PLMN	GTP-U	29.060
Gp	SGSN-GGSN	Used to support mobility; applicable when GGSN and SGSN are located in different PLMNs	GTP-U	29.060
Gi	GGSN-PDN	Used to exchange data with external packet data network	IP	29.061

bandwidth. While the air interface necessarily terminates at the BTS, the RLC/MAC resides at the PCU that may be located at the BTS, BSC, or even SGSN.

2.6.2 Um Interface Between MS and SGSN

While the air interface and RLC/MAC terminate within the BSS, the Logical Link Control (LLC) provides a highly reliable link between MS and SGSN, which is independent of the underlying radio interface. The ciphering is applicable at the LLC layer. The LLC provides services to the Subnetwork Dependent Convergence Protocol (SNDCP). This layer maps network-level characteristics onto the underlying network characteristics. The SNDCP layer also performs functions like compression and decompression of user data, and segmentation and reassembly of network PDUs.

The user payload carried as an IP packet is encapsulated within the GPRS network. In other words, the source and destination address of the user IP packet is not relevant for packet routing. The IP packet is just like an application payload and is carried transparently between the MS and GGSN.

2.6.3 Gb Interface Between BSS and SGSN

On the Gb interface, frame relay is used as network service layer to carry RLC/MAC between the PCU and the SGSN. Over the network service layer is the BSS GPRS Protocol (BSSGP) that uses the frame relay to provide virtual connections between the BSS and the SGSN. The BSSGP also conveys routing and QoS-related information between the BSS and the SGSN.

2.6.4 Gn/Gp Interface Between GSNs

The next hop is the Gn interface between the SGSN and the GGSN. The main protocol used on this interface is the GPRS tunneling Protocol for User Plane (GTP-U). The user payload (typically IP packets) is encapsulated as GTP packets and exchanged using UDP/IP routing. Finally at GGSN, the IP packets are decapsulated and exchanged as normal IP packets.

2.6.5 Gi Interface Between GGSN and PDN

The GGSN is the gateway of the PLMN. It interfaces with the PDN over the Gi interface. For external network, the GGSN is a router that routes incoming packets. The important difference between a normal IP router and the routing functionality of the GGSN is that the IP packets are tunneled to the MS and not routed as normal IP datagrams. This is required because the ease of fixed-line routing is not possible in the mobility supported by the PS domain.

Figure 2.4 shows the connectivity of a GGSN with external PDN. Typically, the PDN could be the Internet so that the GPRS subscriber could access it using the GPRS services. In the 3GPP specifications, support is provided for both IPv4 and IPv6 based services.

[3GPP TS 29.061] defines various scenarios wherein packet services could be accessed using the IPv4, IPv6, and PPP over the Gi interface.

2.7 GPRS Control Plane

While the transmission plane provides means for exchange of user data, the control plane is used to ensure the availability of the transmission plane. Thus, the control plane facilitates signaling between the MS and the GPRS network and between the different entities of the GPRS network.

The GPRS control plane is depicted in Fig. 2.5. As is evident in the figure, the physical layer and RLC/MAC protocol running between the MS and the BSS, and the Gb interface between the BSS and the SGSN, are common for both the GPRS transmission plane and the control plane. What this implies is that the BSS provides a bearer for the carriage of the GPRS information (user or signaling information).

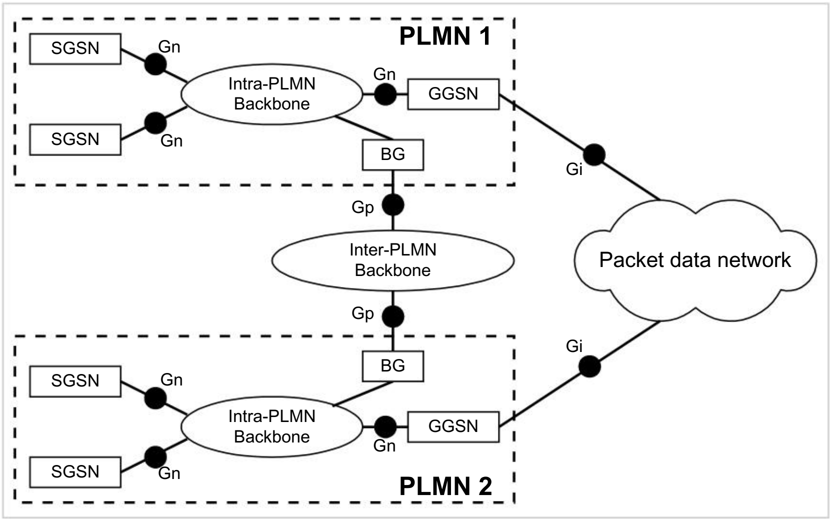


FIGURE 2.4

Intra-PLMN and Inter-PLMN Backbone in GPRS Network

The main protocol in the signaling plane is the GPRS Mobility Management (GMM) and Session Management (SM) protocols running between the MS and the SGSN. Then there is the control plane version of the GPRS Tunnelling Protocol (GTP-C) running between the SGSN and the GGSN.

While Fig. 2.5 shows control plane protocols between MS and GGSN, there are a number of interfaces that are defined between different core network entities like SGSN, GGSN, EIR, AuC/HLR, and VLR. These interfaces are summarized in Table 2.2 and further explained in sections next. The Um interface between MS and BSS and the Gb interface between BSS and SGSN are common for transmission plane and control plane and are not explained any more.

2.7.1 Um Interface Between MS and SGSN

The control plane procedure on the Um interface between MS-SGSN allows the MS to access the service of the GPRS network. Two main sets of procedures applicable are GMM and SM procedures (refer Fig. 2.5). These procedures are defined in [3GPP TS 24.008].

GMM procedures allow an MS to "attach" itself (i.e., register) with the GPRS network whereby its presence is then known to the network. The counterpart of "attach" is "detach", which allows the MS to leave the network gracefully. Then there are other mobility management procedures related to authentication and RA update. The different GMM procedures are summarized next:

- **GPRS attach.** The GPRS attach procedure is used to indicate that the MS is active in the network.
- **GPRS detach.** This procedure can be invoked by the MS if the MS is deactivated or if the SIM is detached from it. Once the GPRS is detached, the MS becomes inactive in the network.

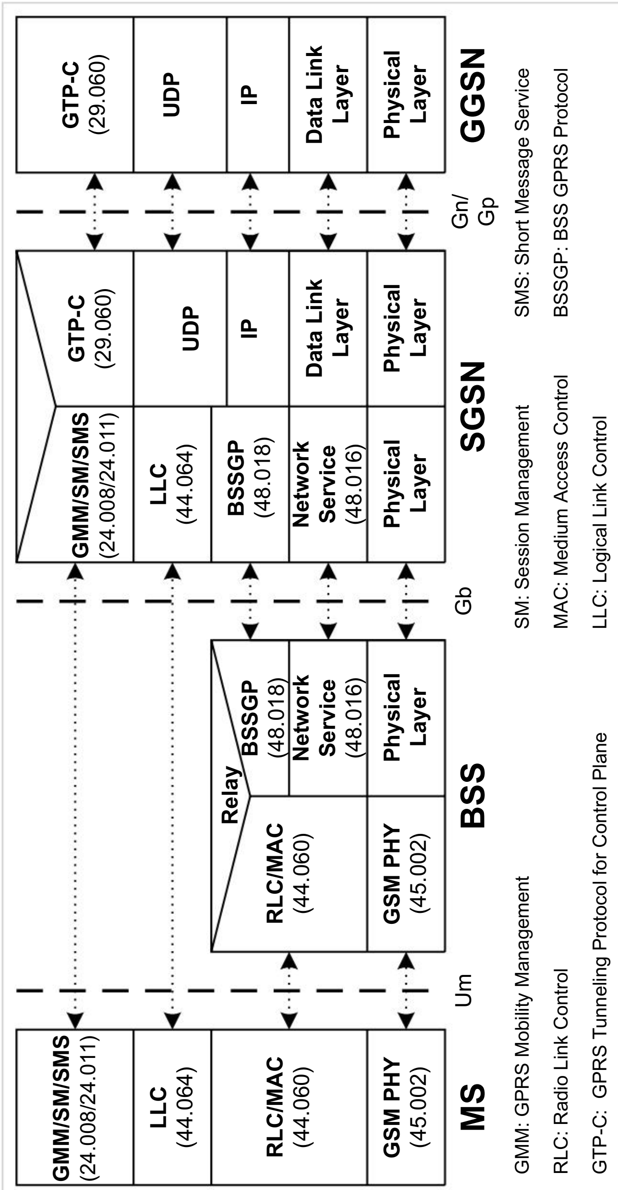


Figure 2.5
GPRS Control Plane

TABLE 2.2 GPRS Control Plane Interfaces

Interface	*Between*	*Description*	*Protocol*	*3GPP Specification*
Um	MS-BSS	Used between MS and BSS for exchange of user and signaling information	Physical RLC/MAC	43.064/45.002 44.060
Um	MS-SGSN	Used between MS and SGSN for exchange of user and signaling information	LLC GMM SM SMS	44.064 24.008 24.008 24.011
Gb	SGSN-BSS	Used between SGSN and BSS for data transfer and mobility management	Physical Network Service BSSGP	48.014 48.016 48.018
Gn	SGSN-GGSN	Used to support mobility; applicable when GGSN and SGSN are located in same PLMN	GTP-C	29.060
Gp	SGSN-GGSN	Used to support mobility; applicable when GGSN and SGSN are located in different PLMNs	GTP-C	29.060
Gr	SGSN-HLR	Used by the SGSN to obtain subscriber information from HLR	MAP	29.002
Gc	GGSN-HLR	Used by the GGSN to retrieve information about the location and supported services for the MS, to be able to activate a packet data network address; this is optional interface	MAP (Note 1)	29.002
Gf	SGSN-EIR	Used by SGSN to enable EIR to verify the IMEI retrieved from MS; this is optional interface	MAP	29.002
Gd	SGSN-SMS-IWMSC SGSN-SMS-GMSC	Used to deliver to and receive short message from MS to SM service center	MAP	29.002
Gs	SGSN-VLR	Used for coordinating the functions of SGSN and VLR when an MS has both GSM and GPRS services; this is optional interface	BSSAP+	29.016/ 29.018

Note 1: If there is no SS7 interface in the GGSN, any GSN in the same PLMN that has an SS7 interface can be used as a "GTP to MAP protocol" converter, thereby providing a signaling path between the GGSN and the HLR.

- **Normal routing area update.** This procedure is used to update the RA of an MS in the network after a change in RA served by the MS.
- **Periodic routing area update.** This is used to periodically notify the presence of the MS in the GPRS network.
- **P-TMSI reallocation procedure.** This procedure is used to allocate a new P-TMSI. The P-TMSI is the GPRS equivalent of TMSI in GSM network. Thus, the relevance and applicability of P-TMSI is similar to that of TMSI. The only difference is that TMSI applies to a Location Area (LA), while P-TMSI applies to an RA.
- **GPRS identification procedure.** This procedure is similar to the MM identification procedure. It is used to obtain the IMSI or IMEI from the MS when the network does not have it.
- **GPRS authentication procedure.** The GPRS authentication procedure is used to corroborate the identity of the user.
- **Combined routing area update.** Apart from the above, a few GMM procedures are defined that optimize the operations when both GSM and GPRS services are used. One of these procedures is combined RA Update. This procedure is used to collectively perform the LA and RA update. In this procedure, the SGSN conveys the received information to the VLR so that the MS does not have to separately communicate with VLR.
- **Combined GPRS attach/detach.** Like combined RA update, the combined GPRS attach/detach enables an MS to indicate its activity in both GSM and GPRS network.

While mobility management tackles the mobility of the MS for the purpose of registration and deregistration, the SM procedures help in the assignment of the PDP address and the management of the PDP context. The PDP context is necessary for exchange of packets with entities of PDN. Note that unlike GSM network, there is no concept of calls in GPRS networks. Thus, the call handling procedure of GSM is not applicable to this domain. An analogous concept, termed as sessions, is applicable in this case. A session can be viewed as a PDP context maintained by the MS and the GSN for information exchange in the GPRS network.

A PDP context contains the PDP type, address, and access point name, among others. This PDP context is used to communicate with external PDN. The SM function of core network enables the activation, deactivation, and modification of a PDP context. The SM procedures can be performed only when a GMM context exists between the MS and the network. If a context does not exist, it must be created first.

The SM procedures include the following:

- **PDP context activation.** This procedure is used to create a PDP context between the MS and the network. The creation of PDP context is initiated by the MS. The network may also ask the MS to initiate the activation of a PDP context (this may happen when a PDU is received by GGSN for MS and no PDP context exists in the GGSN database).
- **PDP context modification.** This procedure is used by the network, or by the MS, to modify the attributes associated with a PDP context. The modifications may be for the QoS or for the priority, or for other parameters negotiated during the activation of PDP context.

- **PDP context deactivation.** This procedure is used by the network or by the MS to deactivate an existing PDP context.

Apart from mobility management and SM, the MS-SGSN interface also allows the exchange of SMS. Note that the SMS can be exchanged using both the GSM as well as GPRS network.

2.7.2 Gn/Gp Interface Between GSNs

The Gn/Gp interface between GSNs is based on the GTP-C (refer Fig. 2.5). This interface is primarily used for creation, modification, and deletion of PDP context. It is also used to obtain SGSN context information maintained at an SGSN during inter-SGSN RA update.

The GTP-C along with GTP-U is defined in [3GPP TS 29.060].

2.7.3 Gr Interface Between SGSN and HLR

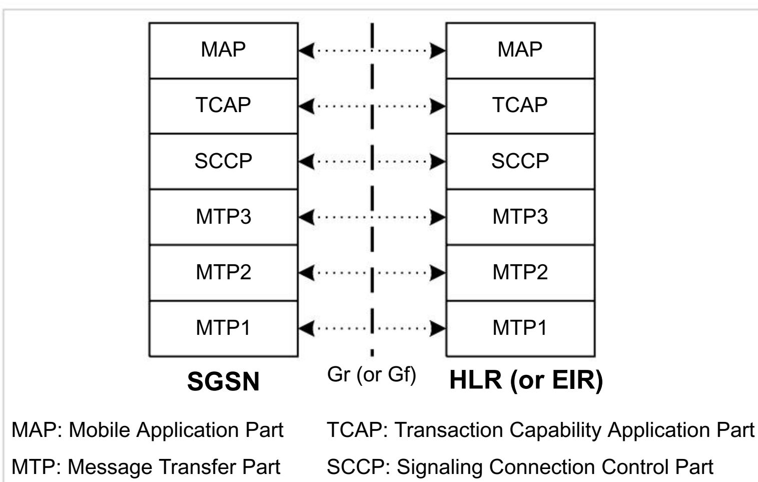
The Gr interface is used by SGSN to update data related to the location of an MS at the HLR. This interface is also used by SGSN to obtain subscriber data from HLR. Thus, the Gr interface between SGSN and HLR is equivalent to the D interface between VLR and HLR.

When an MS roams into a new SGSN area managed by a different SGSN, the new SGSN updates the HLR about this movement. The HLR then sends to the new SGSN subscriber the required data to provide various GPRS services. The HLR also asks the old SGSN to delete any record maintained for the said subscriber. In case only the RA changes but the SGSN remains the same, the HLR is still informed about this change. However, as the SGSN is not changed, there is no need to provide subscriber information or to delete information from the old SGSN.

Apart from the crucial function of GMM, SGSN also uses this interface to obtain authentication information from the HLR. This information is then used to authenticate the subscriber.

Figure 2.6

Protocol Architecture of Control Plane for Gr Interface



The Gr interface is also used by SGSN to purge an MS. Once an MS is marked "Purged" by the HLR, the subscriber is unreachable for MS-terminated requests received by HLR (e.g., network-requested PDP-context activation).

The MAP protocol, specified in [3GPP TS 29.002], is used for the Gr interface.

2.7.4 Gc Interface Between GGSN and HLR

To carry out the network-requested PDP-context activation procedure, the GGSN needs to know the SGSN with which an MS is registered. For this, the GGSN uses the Gc interface to request the HLR to provide the SGSN address. Further, if the network-requested PDP-context activation fails, GGSN sends a failure report to HLR over the Gc interface.

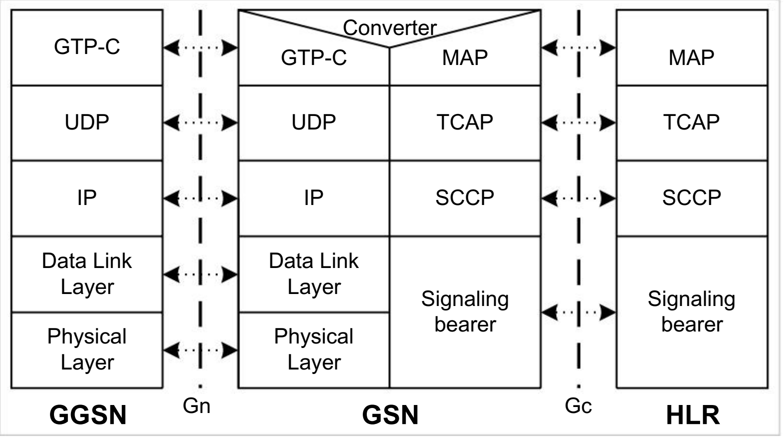
The MAP protocol, defined in [GPP TS 29.002], is used for the Gc interface. The implementation of this interface is optional. In case there is no SS7 interface in the GGSN, any GSN in the same PLMN that has this interface can be used as a "GTP to MAP protocol" converter, thereby providing a signaling path between the GGSN and the HLR.

Note that the MAP is a computationally complex protocol and it increases the software and hardware cost of a GGSN. Thus to simplify the implementation of GGSN and to remove the SS7 interface from GGSN, specifications make the implementation of Gc interface for a GGSN optional. However, this option does not imply that the GGSN does not need to communicate with HLR, which is unavoidable. What this means is that GGSN must have alternate means to obtain information from HLR without actually implementing the SS7-based Gc interface. To provide such means, the GTP-C has additional messages that are used by GGSN to communicate with HLR via a GTP-MAP protocol converter. The GTP-MAP protocol converter is shown in Fig. 2.7.

The benefit of such a GTP-MAP protocol converter is that many GGSN can then communicate with a single GSN that supports such protocol conversion. In such a scenario,

FIGURE 2.7

GTP-MAP Protocol Converter



all such GGSNs only need to support an IP interface with the GSN. The GSN can then do the protocol conversion for all GGSNs.

A question then may arise that why such an interface is not specified for SGSN, which too interfaces with HLR. The answer is that SGSN uses MAP to communicate with HLR, EIR, and SMS-GMSC/SMS-IWMSC. Now, if SGSN has to do away with MAP interface, it needs to have protocol converter for all these interfaces. This leads to standardization complexity and thus has not been specified by the standards.

2.7.5 Gf Interface Between SGSN and EIR

The EIR is the central database for maintaining the list of IMEI used in the GPRS network. In order to ascertain the legitimacy of a UE, SGSN may use the Gf interface requesting EIR to check the IMEI retrieved from the MS.

The MAP protocol, defined in [GPP TS 29.002], is used for the Gf interface.

2.7.6 Gd Interface Between SGSN and SMS-GMSC/SMS-IWMSC

To provide SMS, the SGSN interfaces with the SMS-MSC (SMS-GMSC and SMS-IWMSC), using the Gd interface.

The Gd interface is based on MAP protocol, which is specified in [3GPP TS 29.002].

2.7.7 Gs Interface Between SGSN and MSC/VLR

The functions of MSC/VLR in the GSM network and SGSN in the GPRS network are somewhat similar. For example, both MSC/VLR and SGSN handle IMSI attach/detach functions. This is just one example, as there are many other commonalities as well. Given this scenario, if the functions of the two entities could somehow be combined, there could be vital savings in terms of radio resources.

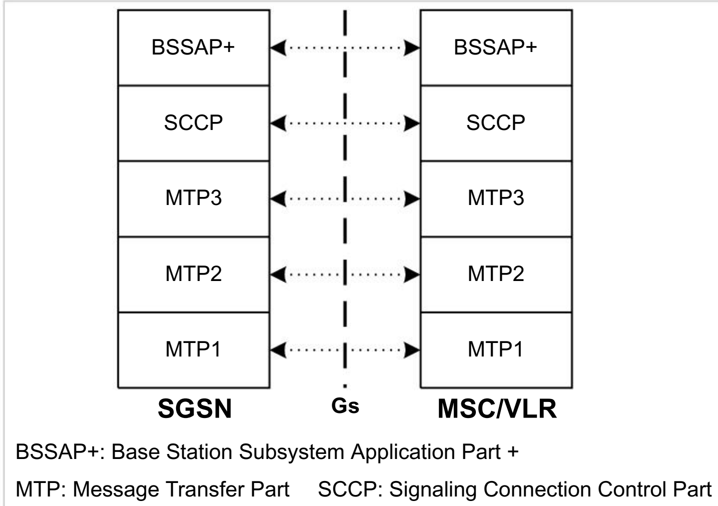
In order to do this, the optional Gs interface is defined between SGSN and MSC/VLR. The Gs interface allows various procedures like IMSI attach/detach via SGSN, paging for GSM via SGSN, and coordination of Routing-Area/Location-Area (RA/LA) Update.

The SGSN and MSC/VLR communicating over the Gs interface maintain an association with each other. This association is created/updated during various message exchanges between the two. The exchanges take place using the Base Station Subsystem Application Part + (BSSAP+) protocol. The protocol architecture for Gs interface is shown in Fig. 2.8. This protocol architecture is similar to any of the MAP-based interfaces (e.g., Gr interface between SGSN and HLR). The only key difference is the absence of TCAP protocol, which is specifically used along with MAP and is not required for BSSAP+. Other protocol layers including SCCP and MTP are same.

The BSSAP+ protocol is specified in [3GPP TS 29.018]. The BSSAP+ layer functions are summarized as follows:

FIGURE 2.8

Protocol Architecture of Control Plane for Gs Interface



- **Paging for GSM.** VLR uses the SGSN for paging for the GSM via the BSSAP+ protocol.
- **Location update.** SGSN carries out the normal location update or IMSI attach via the BSSAP+ protocol.
- **Alert.** BSSAP+ provides the means for VLR to detect any activity by MS in the GPRS.
- **IMSI attach.** BSSAP+ is used by SGSN to indicate the VLR of an explicit or implicit IMSI attach for GSM services.
- **Information exchange.** SGSN acts as a relay to transfer information from VLR to MS and vice versa.

Another specification, [3GPP TS 29.016], provides the lower layer requirements for the Gs interface.

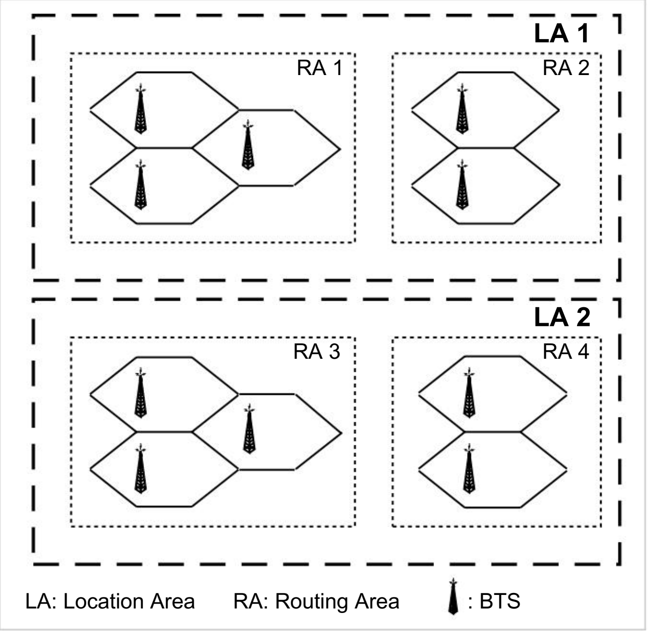
2.8 Hierarchical Organization of GSM/GPRS Network

To support mobility of a subscriber from one location to another, the GSM/GPRS network is organized as a multitier hierarchical structure. This hierarchical structure enables a particular network entity to have only that much information which is required for its functioning. For example, a VLR has information only of the LA of an MS but it does not know the exact cell location. Despite this, through the use of paging, the exact cell location is determined when required. The hierarchical division is done to reduce the storage/processing load on the network entities (like the VLR), save radio resources, and also save battery consumption of the MS.

The hierarchical structure of GSM/GPRS network is depicted in Fig. 2.9. As shown in the figure, at the lowest level of hierarchy is the cell. LA and RA come at the next level. For the

FIGURE 2.9

Hierarchical Structure of GSM/GPRS Network



GSM network, a collection of one or more cells forms an LA. For the GPRS network, a collection of one or more cells forms an RA. Further, an LA contains one or more RA. At the highest level of the hierarchy is a PLMN, which is not depicted in the figure. Note that an LA does not have much significance in a GPRS network, but it is discussed as it forms an integral part of the GSM/GPRS network hierarchy. Moreover, when the Gs interface is present, the VLR uses the services of SGSN for some of the mobility management operations including attach/detach and paging.

The hierarchical structure of GSM/GPRS network is explained in greater detail in the following sections. While the GSM structure was discussed in Chapter 1, they are repeated here for clarity.

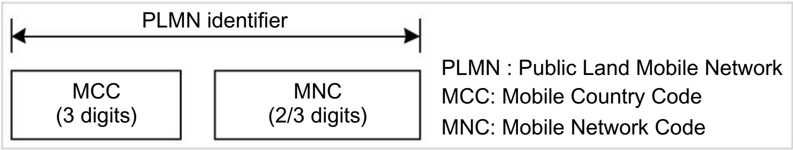
2.8.1 Public Land Mobile Network

In GSM/GPRS network, at the highest level of the hierarchy is a PLMN. A PLMN is defined as a telecommunications network providing mobile cellular services. A PLMN is uniquely identified by its PLMN identifier.

The PLMN identifier comprises Mobile Country Code (MCC) and Mobile Network Code (MNC) as shown in Fig. 2.10. The MCC is of three digits and it identifies the country to which the PLMN belongs. The next two or three digits of the PLMN identifier is the MNC. It identifies a particular PLMN within a country. It is recommended that within a country identified by the MCC, all PLMNs either use only two digits for MNC or only three digits. A combination of the two schemes is not recommended.

Figure 2.10

Structure of PLMN Identifier



2.8.2 Location Area

LA is defined as an area in which an MS may move freely without updating its current location at the VLR. In case an MS moves outside its LA, it informs the VLR of its current location through the location update procedure.

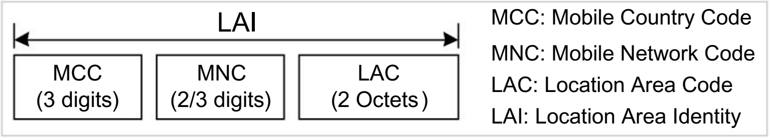
An LA includes one or more cells. The reason for grouping of cells into LA is to facilitate efficient location management. To understand this, note that location management requires tracking the current location of the MS so that a terminating call can be delivered to the MS. Since the MS updates its location only at the change of LA, the VLR has accurate information of the same. When a terminating call for an MS arrives, the VLR pages the MS to seek the exact location of the MS (in terms of its current cell location). Upon receiving the paging request, the MS responds back with its current cell location. This information is used to setup a connection with the MS.

Given this, it is evident that if LA is as small as a cell, there is no need to page the MS. However, this would require constant activity between MS and VLR whenever MS moves to a new cell resulting in consumption of RR and battery power. If the LA is very large, the paging area has to be performed in a very large area, which is undesirable. Thus, grouping of cells into LA allows us to arrive at a balance between the accuracy of information maintained at the VLR versus the uplink radio capacity and battery power consumed in the process.

Each LA is uniquely identified by an LAI. The structure of LAI is shown in Fig. 2.11. The MCC and MNC are same as the Mobile Country Code and the Mobile Network Code of that PLMN to which the LA belongs. The last two octets of the LAI is the LAC that identifies a location with a PLMN. Collectively, the LAI forms a unique identifier for an LA across all PLMNs.

Figure 2.11

Structure of Location Area Identity



2.8.3 Routing Area

The RA for GPRS is analogous to the LA for the GSM. RA is defined as an area in which an MS may move freely without updating its current location at the SGSN. In case an MS moves outside its RA, it informs the SGSN its current location through the RA update procedure.

Like an LA, an RA too may include one or more cells. Grouping of cells into an RA facilitates efficient location management. This is for reasons similar to that mentioned for LA, where a balance is achieved between the frequency of location updates and the area in which paging is done for mobile-terminated sessions.

One important difference between RA and LA is that an RA is always contained within an LA. In other words, LA may contain one or more RA.

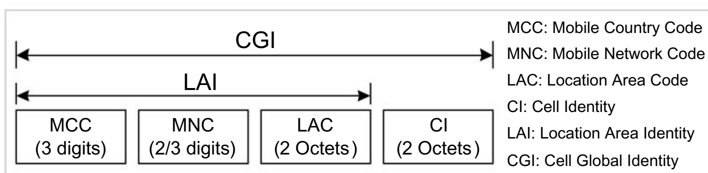
Each RA is uniquely identified by an Routing Area Identity (RAI). Since an RA is a subset of an LA, the RAI is derived from LAI. In fact, the RAI is LAI plus a Routing Area Code (RAC) of 1 octet. The RAC uniquely identifies an RA in an LA. Simplifying it, we get RAI = LAI + RAC.

2.8.4 Cell Global Identity

At the lowest level of GSM/GPRS hierarchy is the cell, which is identified by the Cell Identity (CI). A CI is unique within an LA. To identify a cell uniquely across PLMNs, an identity called the Cell Global Identity (CGI) is defined. CGI is obtained by the concatenation of LAI and the CI. The structure of CGI is depicted in Fig. 2.12.

Figure 2.12

Structure of Cell Global Identity



Conclusion

This chapter introduces the GPRS network architecture and its important interfaces and protocols. It is only an introductory chapter that forms the base for all further discussions in this book. Concepts discussed in this chapter are further elaborated in the remaining chapters of this book.

Further Reading

The mobile wireless area is a vast field and plenty of text is available for different generations of the technology. In the GSM space, [GSM M. Mouly] is an excellent reference. Other good references on GSM include [GSM N. Narang] and [GSM G. Heine]. For GPRS, [GPRS E. Seurre] is a very good reference. For a summarized view, [GPRS C. Bettstetter] is a very good white paper for quick reading and understanding of the subject. Some of the performance aspects of GSM and GPRS are covered in [GSM T. Halonen]. For 3G, the important references include [3G H. Holma], [3G J. Korhonen], and [3G S. Kasera].

For GPRS details, the reader can refer [3GPP TS 23.060], which is the single and most comprehensive specification on GPRS. The stage 1 details (i.e., requirement aspects) are available in [3GPP TS 22.060]. For GPRS network architecture and interfaces, reader is referred to [3GPP TS 23.002].

For the transmission protocols, the physical layer aspects are covered in [3GPP TS 43.064] and [3GPP TS 45.002]. For the RLC/MAC protocol, [3GPP TS 44.060] is referred. Other protocol references on Um side include [3GPP TS 44.064] for LLC and [3GPP TS 44.065] for SNDCP. On the Gb interface, [3GPP TS 48.014], [3GPP TS 48.016], and [3GPP TS 48.018] can be referred for physical layer, network service layer, and BSSGP layer, respectively. On the Gn/Gp interface, the GTP-U protocol is detailed in [3GPP TS 29.060]. The Gi interface has details in [3GPP TS 29.061].

For the control plane, the key protocols include GMM/SM specified in [3GPP TS 24.008] and SMS specified in [3GPP TS 24.011]. The GTP-C protocol is detailed in [3GPP TS 29.060]. Apart from this, the very important core network protocol MAP is specified in [3GPP TS 29.002]. MAP uses the services of TCAP specified in ITU-T Q.77x series. Then other SS7 protocols include SCCP specified in ITU-T Q.71x series and MTP specified in ITU-T Q.70x series. Then there is the BSSAP+ protocol specified in [3GPP TS 29.016] and [3GPP TS 29.018].

Chapter 3

GPRS Air Interface

3.1 Introduction

The GPRS air interface is a layered structure comprising of two distinct layers, namely Radio Frequency Layer (RFL) and Physical Link Layer (PLL). The RFL performs the modulation of the physical waveforms based on the sequence of bits received from the higher layer (i.e., PLL). On the receiving side, this layer demodulates received waveforms into a sequence of bits, which are transferred to the physical link layer for further processing. Thus, the key functions included in the RFL include the modulation/demodulation, and transmission/reception of radio blocks. The modulation/demodulation is similar to GSM and is detailed in [3GPP TS 45.004]. Another important RFL specification includes [3GPP TS 45.005] that describes transmitter and receiver characteristics and associated performance requirements.

The PLL operates above the RFL and provides higher physical layer functions including channel coding/decoding, interleaving, synchronization and timing advance, power control, measurements handling, cell selection/reselection, and Discontinuous Reception (DRX). Among various functions listed above, the channel coding/decoding and interleaving concepts are discussed in Section 3.8. This section talks about various coding schemes from CS1 to CS4 and also mentions the use of Forward Error Correction (FEC) technique for error detection and correction. Section 3.9 talks about the synchronization aspects and explains techniques like continuous timing advance and timing advance management based on network polling. At the end of this chapter, Section 3.10 elaborates upon power control techniques including uplink and downlink power control methods. Other RFL functions including measurement handling, cell selection/reselection, and DRX are explained in next chapter as part of Medium Access Control (MAC) procedures.

In order to understand these fundamental GPRS air interface functions, the block and channel structure of GPRS is explained first that forms the basis of all subsequent discussions. It may be noted that this chapter should be read in conjunction with Chapter 4 on MAC where some of the concepts related to air interface handling including Temporary Block Flow (TBF), Temporary Flow Identity (TFI), Network Mode of Operation (NOM) and Uplink State Flag (USF) are discussed in detail. These concepts are directly used in this chapter without explaining them.

3.2 GSM versus GPRS Air Interface

To understand GPRS air interface, one can start with GSM air interface and analyze how various aspects have been altered and evolved to satisfy the requirements of GPRS. To start with, GPRS reuses the basic air interface structure defined for GSM (as discussed in Chapter 1). This includes the Time Division Multiple Access/Frequency Division Multiple Access (TDMA/FDMA) multiplexing scheme using which the BTS supports multiple frequencies, where each frequency has a recurring pattern of eight timeslots. The TDMA frame structure, the symbol period, and the TDMA repetition period in GPRS are as they are in GSM. Furthermore, no new burst structure is defined for GPRS and the basic burst structure for normal and access burst is reused. However, the frequency correction burst and synchronization burst are not used in GPRS because GSM and GPRS both use a common Broadcast Control Channel (BCCH) and there is no separate BCCH for GPRS that carries frequency and time synchronization information.

If one looks at the frame structure, there are two types of frames in GSM, a 26-frame multiframe exclusively for traffic and a 51-frame multiframe for signaling. This dual structure concept is done away with in GPRS and replaced by 52-frame multiframe that carries both signaling and traffic. In this 52-frame multiframe, there are 12 blocks of four slots each, two idle slots and two slots used for handling timing advance.

The logical channels in GPRS, although on similar lines, are different as compared to GSM. For "broadcast channels", the Frequency Correction Channel (FCCH) and Synchronization Channel (SCH) are done away with because the GPRS synchronization takes place using the BCCH of GSM. For BCCH, there is an optional Packet Broadcast Control Channel (PBCCH) that carries GPRS-specific System Information (SI) called the Packet System Information (PSI).

The "common control channels" in GSM and GPRS have great similarity. Thus, one finds the Random Access Channel (RACH), Access Grant Channel (AGCH) and Paging Channel (PCH) in GSM having corresponding channels in GPRS including Packet Random Access Channel (PRACH) for carrying uplink access requests, Packet Access Grant Channel (PAGCH) for carrying downlink channel assignment messages, and Packet Paging Channel (PPCH) for downlink paging messages.

Because of the fundamental difference in channel allocation (circuit-switching versus packet switching), the GSM and GPRS "dedicated channels" for signaling and traffic are quite different. Thus, Packet Data Traffic Channel (PDTCH) is allocated on demand as against the static nature of the Traffic Channel (TCH) in GSM. There is no direct equivalence of Slow Associated Control Channel/Fast Associated Control Channel (SACCH/FACCH) in

GPRS. For timing advance purposes, the Packet Timing Advance Control Channel (PTCCH) is used. For other signaling purposes, the Packet Associated Control Channel (PACCH) is used. The PACCH, like PDTCH, does not have static allocation for the users.

3.3 GPRS Multiframe Structure

The GPRS air interface is a recurring TDMA frame that comprises of eight timeslots (numbered modulo 8). The physical content of a timeslot is called a "burst". GPRS uses the access burst and the normal burst among the various burst types supported by GSM. At the next level is a collection of 52 TDMA frames resulting in a 52-frame multiframe. This structure, unlike the GSM's dual structure of 26-frame multiframe for traffic and 51-frame multiframe for signaling, is used for both traffic and signaling.

The basic structure of this 52-frame multiframe is depicted in Fig. 3.1. As shown in the figure, the multiframe comprises of 12 radio blocks numbered as B0 to B11. A radio block is carried over four consecutive TDMA frames of a timeslot. These blocks carry signaling or user data message in normal burst format. The only exception is the Mobile Station (MS) response to network polling message that can be sent through four access bursts in a block. Apart from the 12 radio blocks, there are four other timeslots in a GPRS multiframe, two of which are idle and remaining two are used for continuous timing advance.

3.4 GPRS Logical Channel

In GPRS, the basic physical channel is referred as Packet Data Channel (PDCH) identified by its Absolute Radio Frequency Channel Number (ARFCN) and the timeslot. As a cell site may be used for both GSM and GPRS operations, the pool of timeslots allocated for GPRS is used for PDCH. How the pool of channels is divided between GSM and GPRS, statically or dynamically, is an internal matter of the Base Station Subsystem (BSS). In the simplest scheme, some part of the pool could be fixed for GSM, some for GPRS, and there could be a common pool that could be shared between GSM and GPRS. Nonetheless, at any given time, a timeslot can either be used for GSM or GPRS, but not for both of them simultaneously.

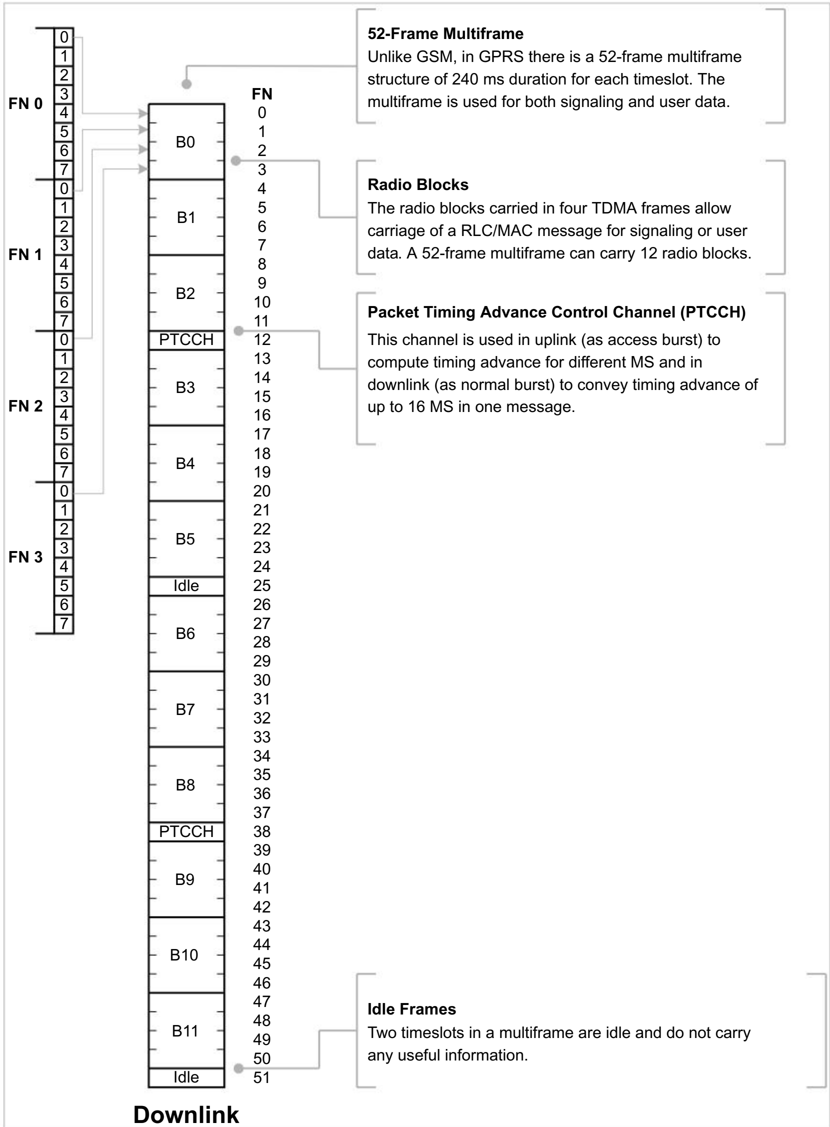
Depending upon how a channel is used, a number of logical channels are defined that map to the PDCH. At the top level, a logical channel is divided into two categories, namely traffic and signaling channels (Fig. 3.2). These are further explained next.

3.5 GPRS Traffic Channel

The GPRS traffic channel, called the Packet Data Traffic Channel (PDTCH), is used to carry user data in uplink and downlink. Unlike GSM where the uplink and downlink allocation is simultaneous, in GPRS the uplink and downlink allocations are independent and have no relation with each other. An MS can be allocated one or more uplink and/or one or more downlink PDTCH at any given time. The maximum number of PDTCH that an MS can

FIGURE 3.1

GPRS Multiframe Structure

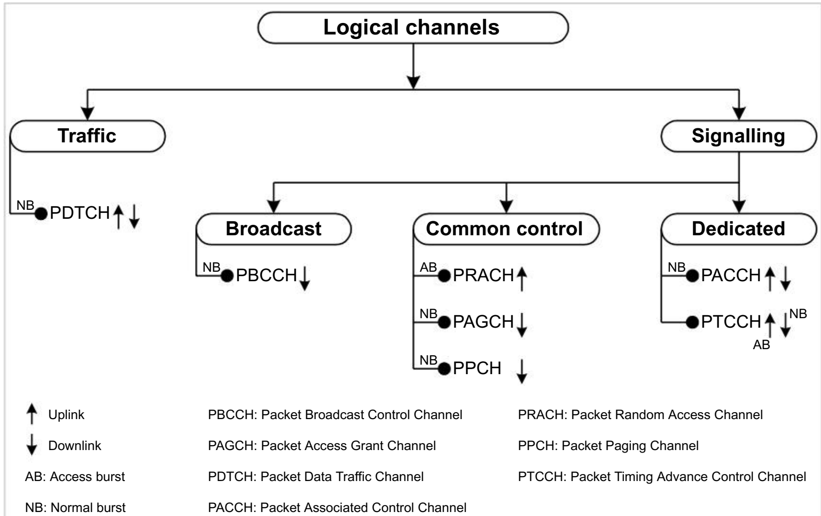


support is linked to the multislot capability of the MS. Furthermore, depending upon the Coding Scheme (CS1 to CS4) used, the PDTCH provides varying data rates from 9.05 to 21.4 Kbps. Thus, theoretically, if an MS is allocated eight timeslots, then it can support a maximum data rate of 171.2 Kbps (8 × 21.4). However, practical limitations mean that the maximum rates are much lower, say 85 Kbps.

The PDTCH uses normal burst structure.

FIGURE 3.2

GPRS Logical Channels



3.6 GPRS Control Channel

While the PDTCH is used for user data, nothing is possible without control channels that carry the necessary signaling information for the configuration of the traffic channels. Depending upon the purpose they serve, the control channels are of different categories: "Packet Broadcast Control Channel," "Packet Common Control Channel," and "Packet Dedicated Control Channel" (see Fig. 3.2).

3.6.1 GPRS Broadcast Control Channel

Broadcast channels, as the name suggests, are broadcast on the air interface and are used by the mobile population at large. In GSM, these broadcast channels allow the mobile to latch on the BCCH frequency. In particular, these channels carry frequency (FCCH), timing (SCH), and system information (BCCH). GPRS does not have its separate FCCH and SCH, but there is an optional provision to carry Packet System Information (PSI) specific to GPRS on a PDCH. This logical channel is called the Packet Broadcast Control Channel (PBCCH) and carries the PSI. As the PBCCH is optional, if it is not allocated, the system information specific to GPRS is broadcast on the BCCH (on System Information 13). The support of GPRS itself (and hence the presence of System Information 13) is indicated in the mandatory system information of GSM (i.e., on System Information 3).

The PBCCH uses normal burst structure.

3.6.2 GPRS Common Control Channel

The next set of channels are the common control channels that can be viewed as a pool of channels that are used by the MS or the network as and when the need arises. The common channels are used in the uplink for random access (to initiate dedicated channel allocation) and in the downlink to respond to the random access by granting a dedicated channel or to urge the mobile to initiate random access by paging it. The different types of common control channels defined by the specifications are explained next.

3.6.2.1 Packet Random Access Channel

The Packet Random Access Channel (PRACH) is used by the MS to initiate uplink transfer for sending data or signaling information. Packet access burst is used on PRACH. Two formats are defined for the access burst, namely the 8-bit format and the 11-bit format. The exact format to be chosen is specified by the network in the system information through the ACCESS_BURST_TYPE parameter in PSI 1 message. As compared to the 8-bit format, the 11-bit format allows MS to convey additional information to the network in the beginning of connection request. The information include the Type (1-bit), Multislot class (5-bit), Priority (2-bit), and Random bits (3-bit). While this structure is used for one-phase access, there are structures defined for short access request, two-phase access request, and page response among others.

3.6.2.2 Packet Access Grant Channel

The Packet Access Grant Channel (PAGCH) is used in the packet transfer establishment phase to send resource assignment to an MS prior to packet transfer. This is used to respond to the packet channel request on the PRACH. In the uplink when a channel is requested on PRACH, the PAGCH is used in downlink to carry the packet uplink assignment message. For establishment of a downlink channel also, the PAGCH is used to carry the packet downlink assignment message.

3.6.2.3 Packet Paging Channel

The third type of packet common control channel is the Packet Paging Channel (PPCH). The PPCH is used to page an MS prior to downlink packet transfer. The PPCH can be used for paging of both circuit-switched and packet data services. The paging for circuit-switched services on PPCH is applicable for Class A and B GPRS MSs in network operation mode I.

The scheduling of messages on PPCH is based on Dis-continuous Reception (DRX) mode for efficient transfer of Paging messages. The DRX is based on the paging groups and is similar to the technique used in GSM. The difference between GSM DRX and GPRS DRX is that one applies for PCH while the other applies for PPCH. Furthermore, the GPRS DRX is slightly more complex as it uses the notion of SPLIT_PG_CYCLE, which allows delivery of more than one paging messages in a paging group. The details of SPLIT_PG_CYCLE are beyond the scope of this book.

3.6.3 GPRS Dedicated Control Channel

The last category of control channels is the dedicated control channels. The nature of these channels in GPRS is quite different as compared to GSM. In GSM, there was a SACCH (for measurement reporting, power control, and timing advance), FACCH (for handover and call establishment/release), and Stand-Alone Dedicated Control Channel (SDCCH) (for SMS, location update, and initial call control). For various reasons, there are no equivalent channels for any of these in GPRS.

In GPRS, there are two types of dedicated control channels as described next.

3.6.3.1 Packet Associated Control Channel

The Packet Associated Control Channel (PACCH) conveys signaling information related to a given MS. The signaling information includes, e.g., acknowledgments and power control information. PACCH also carries resource assignment and re-assignment messages. The PACCH shares resources with PDTCHs that are currently assigned to one MS. Additionally, an MS that is currently involved in packet transfer, can be paged for circuit-switched services on PACCH.

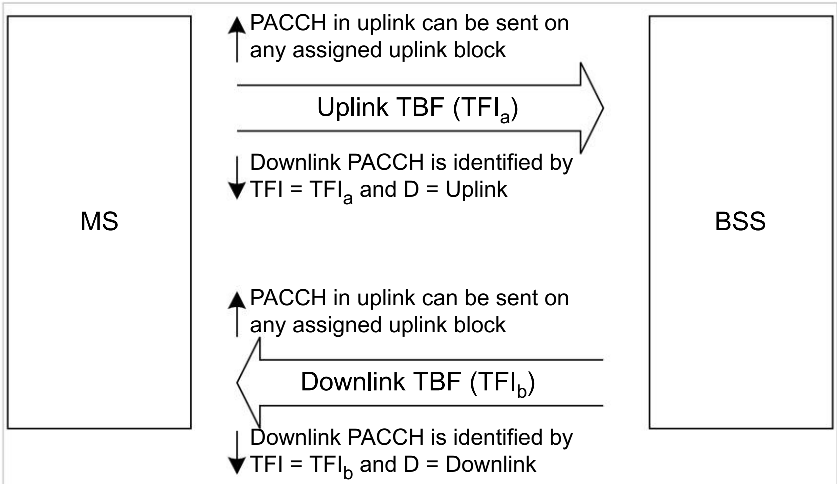
The PACCH is dynamically allocated on a block-by-block basis on the same physical channel that carries the PDTCH. Furthermore, the PACCH is a bidirectional channel that is dynamically allocated both in uplink and downlink irrespective of whether the corresponding PDTCH is for uplink or downlink. However, because of the dynamic nature, its usage is controlled by the network. For this purpose, the "D" (direction) bit is defined in the Downlink Radio Link Control/Medium Access Control (RLC/MAC) blocks. The D bit along with the Temporary Flow Identity (TFI) identifies the destination MS of the control message. Thus, even though the message flows in the downlink direction, the message can be for the MS with the PDTCH on the downlink or for the MS with the PDTCH on the uplink. Because of the unidirectional allocation of PDTCH, the D bit is necessary for knowing the direction of the Temporary Block Flow (TBF) and the TFI is needed to identify the target MS among many MS that share the PDTCH with the target MS. The PACCH allocation for uplink and downlink is explained in Fig. 3.3.

3.6.3.2 Packet Timing Advance Control Channel

Given that there is no permanent channel like SACCH in GPRS, there is a need to ensure that even when packet transmission is not continuous, the MS always stays synchronized with BTS. Thus, the Packet Timing advance Control Channel (PTCCH) is defined exclusively to maintain Timing Advance (TA) of the MS. Given that the TA value is computed using access burst in the uplink and informed to the MS using normal burst in the downlink, two types of PTCCH are defined, namely PTCCH/U that uses access burst and PTCCH/D that uses normal burst. The PTCCH/U is used to transmit access burst to allow estimation of the timing advance for one MS in packet transfer mode. A structure spanning four 52-frame multiframes is used for PTCCH where up to 16 MS can send their access bursts. In the downlink, the PTCCH/D is used to transmit timing advance information updates to several MS (up to 16). The use of uplink and downlink PTCCH provides continuous timing advance functionality that is further discussed in Section 3.9.2.

FIGURE 3.3

Uplink and Downlink PACCH



3.7 GPRS Channel Organization

The structure of a GPRS channel (i.e., PDCH) is based on the presence or absence of Packet Common Control Channel (PCCCH) on that channel. A PDCH that carries PCCCH for common control signaling is called the "master channel," while the other channels that do not carry the PCCCH are called "slave channels."

A master channel, apart from the PCCCH, carries the data channels (PDTCH) and the dedicated signaling channels (PACCH). Note that all PDCHs carry PTCCH irrespective of whether they are master channels or slave channels. The PBCCH is carried on one of the master channels. Furthermore, it is optional for the network to have the master channel. Under such circumstances, the PBCCH and PCCCH functions are carried over the Broadcast Control Channels (BCCH) and Common Control Channels (CCCH).

The slave channels only carry the data channels (PDTCH) and the dedicated signaling channels (PACCH). These do not carry the PBCCH or the PCCCH.

3.7.1 GPRS Downlink Channel Organization

The downlink structure of GPRS PDCH channels depends upon whether it is a "master channel" or a "slave channel". The slave configuration is very simple and comprises of PDTCH, PACCH, and the PTCCH. The master configuration is discussed next.

The master channel is of two types, one that carries PBCCH and one that does not. The first category of master channel includes PBCCH, PCCCH (including the PAGCH and PPCH), PDTCH, PACCH, and PTCCH. On one carrier frequency (i.e., ARFCN), there can be at most one master channel that carries the PBCCH. If PBCCH is not supported, then there is no master channel present for the ARFCN.

The other category of master channel is one that carries PCCCH (including the PAGCH and PPCH), PDTCH, PACCH, and PTCCH. There can be zero or more such channels. Such a channel can exist only if the PBCCH exists in the network; in other words, a master channel without PBCCH cannot exist if the master channel with PBCCH does not exist on any PDCH of the carrier. The PSI 2 carries information of which all timeslots of an ARFCN carry the PCCCH along with other parameters like the training sequence and the frequency-hopping information.

If the above concepts are compared with GSM, then in GSM too, there can be one and only one BCCH channel per cell, and optionally some additional CCCH channels timeslots (apart from the CCCH carried on timeslot 0 of the BCCH). In GPRS too, there can be a maximum of one PBCCH (master channel with PBCCH) and optionally one or more channels with PCCCH (master channels without PBCCH). The only difference between the two is that in GSM every cell must have a BCCH channel, and in GPRS the PBCCH itself is optional and its functionality can be served using BCCH (implying that master channel itself becomes optional).

The information whether PBCCH is supported or not is communicated to the MS on SI 13 in the "SI 13 rest octets". This tells the MS whether the PBCCH is supported or not. If it is supported, this element also carries other details like the timeslot number, training sequence, and ARFCN on which the PBCCH is supported. The information of the PCCCH is carried on the PBCCH (recall that the PCCCH does not exist without the PBCCH).

In order to identify the logical channels carried on GPRS blocks, the notion of ordered list is defined. This ordered list defines the sequences of blocks as B0, B6, B3, B9, B1, B7, B4, B8, B2, B8, B5, and B11. The use of this list as against the sequential list of blocks is helpful in spreading the logical channels evenly in a multiframe. As an example, consider the PBCCH. The number of PBCCH in a multiframe is defined by the parameter BS_PBCCH_BLKS, which indicates the number of blocks (between 1 and 4) reserved for the PBCCH within the 52-multiframe. While this 2-bit BS_PBCCH_BLKS field has a value of 0–3, the number of PBCCH blocks are 1 greater than the value and hence 1–4. Now, if there were only two PBCCH, then these would be spread evenly on B0 and B6. Similarly, if there were four PBCCH, then also these would be spread evenly on B0, B6, B3, and B9. If one compares GSM with GPRS, there is no equivalent of this parameter in GSM. However, GSM does have the option of using extended BCCH thus occupying two blocks.

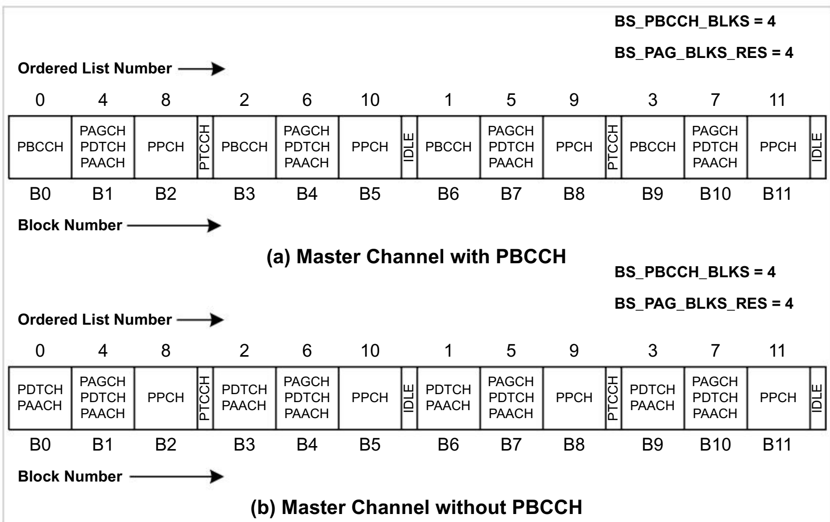
Like PBCCH, the number of PAGCH blocks in a multiframe is determined by the parameter BS_PAG_BLKS_RES that ranges from 0 to 12. This parameter essentially indicates the number of blocks that can be either used by the common control channel PAGCH (but not PBCCH or PPCH) or by PDTCH and PACCH. The latter option provides added flexibility to the network of using blocks in master channels for data transfer when there are not enough access grant channels to be sent on the air. If one compares GSM with GPRS, then there is exactly similar parameter BS_AG_BLKS_RES that defines the number of CCCH blocks used exclusively for AGCH (and not PCH). However, as GSM is circuit switched in nature, there is no option of sharing AGCH channels with traffic channels as seen in GPRS.

Figure 3.4 shows the allocation of logical channels on the master channel where BS_PBCCH_BLKS = 4 and BS_PAG_BLKS_RES = 4. Figure 3.4(a) shows allocation of PBCCH on B0, B6, B3, and B9. It also shows the allocation of PAGCH or PDTCH/PACCH when BS_PAG_BLKS_RES = 4. As shown, the PAGCH uses the four blocks that follow PBCCH (i.e., B1, B7, B4, and B10). The remaining blocks are then used for PPCH.

Figure 3.4(b) shows the structure of master channels without the PBCCH. As shown in the figure, the blocks that are reserved for PBCCH are used for PDTCH. Other part is similar to Figure 3.4(a).

FIGURE 3.4

Logical Channel Allocation for Master Channel in Downlink



3.7.2 GPRS Uplink Channel Organization

The uplink channel organization, like downlink structure, depends upon whether it is a master or a slave configuration. The slave configuration in uplink is like downlink and comprises the PDTCH, PACCH, and the PTCCH. For the master configuration, while the downlink structure has two flavors, one with PBCCH and one without PBCCH, the uplink has only one structure that comprises the PRACH, PDTCH, PACCH, and PTCCH.

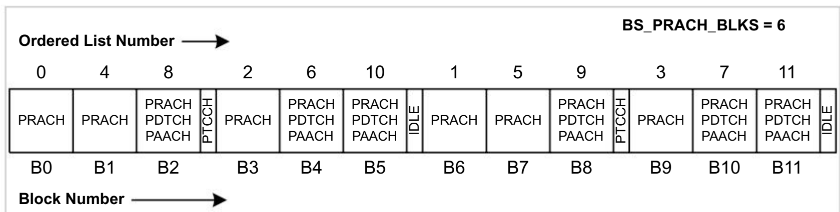
For allocation of PRACH, two types of pools are defined. First pool of PRACH includes blocks that are exclusively reserved for the PRACH and are not shared with the PDTCH/PACCH. The size of this pool is determined by the parameter BS_PRACH_BLKS (having values of 0–12). The parameter BS_PRACH_BLKS, along with other parameters BS_PBCCH_BLKS and BS_PAG_BLKS_RES that govern the allocation of common control channels, is broadcast using the information element PCCCH Organization Parameters over PSI 1.

The remaining PRACHs coexist with PDTCH and PACCH and are allocated on a dynamic basis. The network uses the specific Uplink State Flag (USF) of the previous block to convey the occurrence of PRACH on a given block. In particular, the binary value of "111" (USF = FREE) indicates that the corresponding uplink radio block contains PRACH. Note that for the PRACH pool that is a part of fixed allocation, an MS does not have to monitor the USF that is broadcast on the downlink.

Figure 3.5 depicts the channel allocation for a master channel in uplink when BS_PRACH_BLKS is equal to six. Thus, six blocks are reserved exclusively for PRACH. The remaining six blocks are shared between PRACH and PDTCH/PACCH.

FIGURE 3.5

Logical Channel Allocation for Master Channel in Uplink



3.8 Channel Coding

In GPRS, various coding schemes are defined to handle the noise, interference, and fading of the air interface. These coding schemes, called CS1 to CS4, have an inverse relationship between the "data rates" supported and the "robustness" against errors. Thus, CS1 provides the lowest rate but has the highest robustness. In contrast, the CS4 scheme has the highest data rate but the lowest redundancy for error handling. For an MS, it is mandatory to support all coding schemes, while for the network, only CS1 is mandatory. Table 3.1 summarizes various parameters impacting the GPRS coding schemes.

The basic idea is to take a payload, and add a Block Check Sequence (BCS) to it. Thereafter, the USF either in its original form or with extra redundancy for added protection (referred to as "precoded USF") is prefixed. The USF thus forms the first 3 bits in all coding schemes. Incorrect coding of USF would cause considerable damage for subsequent uplink transmission.

A tail is also added and the resultant bit stream is convolutionally coded. The output bits are then punctured to give an output of 456 bits. (The puncturing allows different coding schemes with same coding rate to give different output rates. Thus, even though CS1, CS2, and CS3 all use 1/2 coding, they have different puncturing schemes thereby giving different output rates. At the receiving side, the punctured positions, also called holes, are filled with 0s before they are decoded.) The bits left after puncturing is then sent over four bursts. Among various steps listed above, the exact steps depend upon the coding scheme used. These are explained later in this section.

TABLE 3.1

Coding Parameters for GPRS Coding Schemes

Scheme	*Code rate*	*USF*	*Precoded USF*	*Radio Block excluding USF and BCS*	*BCS*	*Tail*	*Coded Bits*	*Punctured Bits*	*Data rate kb/s*
CS1	1/2	3	3	181	40	4	456	0	9.05
CS2	2/3	3	6	268	16	4	588	132	13.4
CS3	3/4	3	6	312	16	4	676	220	15.6
CS4	1	3	12	428	16	—	456	—	21.4

Legend

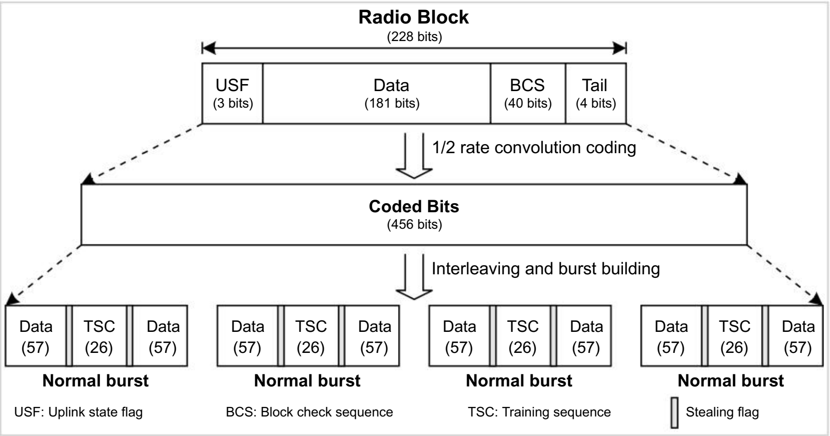
- **Code Rate.** This rate, k/n, is the ratio of the number of bits into the convolutional encoder (k) to the number of channel symbols output by the convolutional encoder (n) in a given encoder cycle. The ratio falls with higher coding scheme indicating lesser redundancy.
- **Uplink State Flag (USF).** This is the flag sent in downlink for indicating uplink allocation.
- **Precoded USF.** Redundancy added to the USF for higher codecs whose code rate is low in order to provide greater robustness to the USF bits.
- **Radio Blocks (excluding USF/BCS).** These are the data bits carried in the radio blocks.
- **Block Check Sequence (BCS).** These are added for the error detection of the data part.
- **Tail.** These four tail bits are used to reset the convolutional coder so that each frame has the same initial status.
- **Coded Bits.** This is the output of the convolutional encoder before puncturing.
- **Punctured Bits.** These are the bits punctured (or removed) before the four bursts are built.
- **Data Rate.** This refers to the data rate of the actual payload (excluding all signalling and error handling bits). For example, for CS1, it is 12 (blocks) × 181 bits (per block)/240 ms (per multiframe).

If one considers CS1 (see Fig. 3.6), then it does not use precoding of the USF because it already offers a very high level of error protection. The 181 bits, along with 3-bit USF, 40 USF bits, and 4 tail bits are convolutionally coded to give 456 bits of data that are sent over four bursts. The CS1 coding scheme is very similar to the coding scheme used for control channels like SACCH in GSM. In SACCH too, 184 bits are coded to give 456 bits.

In CS2, the number of encoded bits is higher (268 bits). Furthermore, as the effective coding rate is less than 1/2, the USF is precoded into a 6-bit USF. Along with the 268 data bits and precoded USF, there is 16-bit BCS (as against the 40-bit BCS for CS1) and 4 tail bits. These 294 bits are then convolutionally coded into 588 bits using 1/2 coder. The 588 bits are more than what can be accommodated in four normal bursts, and therefore some of these bits are punctured according to a specified rule. The bits left after puncturing are sent in a block over four normal bursts (see Fig. 3.7). The puncturing process does not touch the USF, which goes as 12 bits of coded information.

Figure 3.6

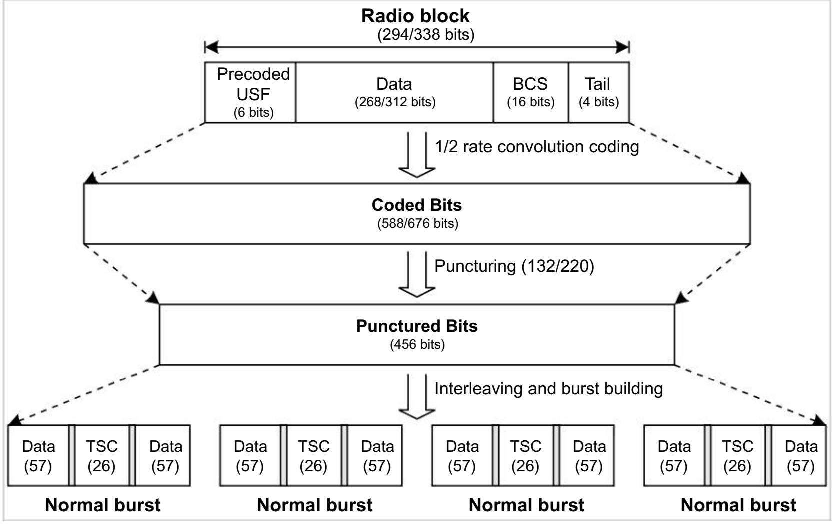
Channel Coding in CS1



The channel coding for CS3 is similar to that used for CS2. The key difference is in the number of input bits (294 versus 338). Consequently, the number of punctured bits is also higher (see Fig. 3.7).

Figure 3.7

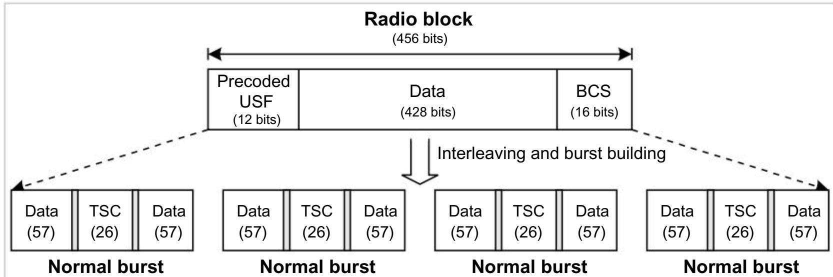
Channel Coding in CS2/CS3



The CS4 offers the highest data rates and thus uses slightly different form of coding. In fact, to offer this data rate, it totally does away with convolution coding (see Fig. 3.8). Thus, it takes 428 bits of data along with the 16-bit BCS and the 12-bit USF to form the 456 bits. Some points here may be noted. As the convolution coding is not used, the USF is offered higher protection by having a 12-bit precoding instead of the 6-bit precoding used in CS3 and CS4. Thus, for the

Figure 3.8

Channel Coding in CS4



USF, the net effect is same without the convolution coding. Consequently, the tail bit is also not used as it is linked to the resetting of the convolution coding.

In order to allow the receiver to know which coding scheme is used by the sender, the stealing flags are used. Recall that in GSM, stealing flags are used to differentiate between the traffic channel (TCH) and the signaling channel (FACCH). As there is no notion of FACCH in GPRS, the stealing flags have no relevance. Thus, these flags in GPRS are used to identify the coding scheme. While four different coding schemes can be identified by 2 bits, a total of eight stealing flags are used instead. These stealing flags correspond to the four bursts carried in a block where each burst carries a pair of stealing flags. Table 3.2 shows the mapping between the stealing flags and the coding schemes. If one compares the different flags, it can be observed that any two combinations differ in 5 or 6 bits. This allows to handle some errors in decoding the stealing flags (as against the use of just 2 bits).

Table 3.2

Mapping Between Coding Scheme and Stealing Flag

Coding Scheme	*Stealing Flag Pattern*
CS1	11111111
CS2	11001000
CS3	00100001
CS4	00010110

3.9 Timing Advance

The GSM and GPRS networks are synchronized such that the transmission and reception must take place within well-defined time boundaries. Due to different nature of channel allocation, these wireless technologies pose different challenges with respect to timing advance. In GSM, there is a dedicated signaling channel SACCH, both for SDCCH and TCH, which allows the MS to regularly send and receive Timing Advance (TA) information from the BTS. Such a dedicated channel is not available in GPRS because of its bandwidth-on-demand characteristics. Moreover, the channel allocation in GPRS is unidirectional and thus it cannot be assumed that the bidirectional channel is available for exchange of timing advance.

To handle this problem, two techniques are defined. The most common one that is solely controlled by the BTS is the "continuous timing advance" procedure. Another option is through PCU where packet polling is used requesting the MS to send access bursts so that the network can compute the timing advance value and inform it to the MS. Apart from these two techniques, the method for calculating the initial timing advance is explained next.

3.9.1 Initial Timing Advance

As in GSM, the initial timing advance is calculated through the channel request/packet channel request message sent as access burst on RACH or PRACH. The immediate assignment (on AGCH) or packet uplink/downlink assignment (on PAGCH) then carries the estimated timing advance value to the MS. This MS uses this value for the uplink transmissions until the continuous timing advance update provides a new value.

The above procedure has two exceptions. The first exception is when the packet queuing notification is used. In this case, because of buffering delays, the initial estimated timing advance might become too old to be sent in the packet uplink/downlink assignment. The other exception is that when the packet uplink/downlink assignment is to be sent without prior paging (i.e., in the ready state), no valid timing advance value may be available.

Under the above exception conditions, the network can either force the MS to use the polling procedure to trigger the calculation of timing advance (see Section 3.9.3), or the packet uplink/downlink assignment can be sent without the timing advance value, which forces the MS to use the continuous timing advance to calculate the first timing advance value to be used.

Note that if the timing advance information is not provided in the assignment message, the mobile is not allowed to send normal bursts on the uplink until it receives a valid timing advance either in packet timing advance/power control message or through the continuous timing advance procedure.

3.9.2 Continuous Timing Advance

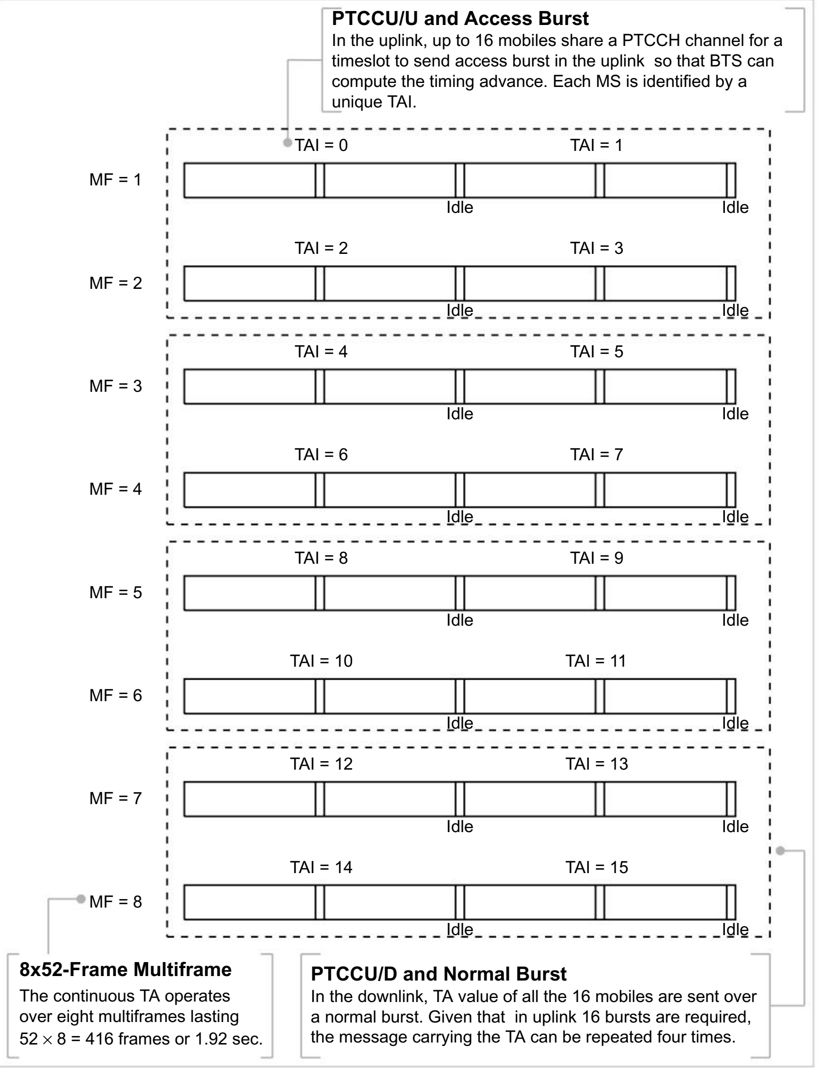
In order to provide SACCH-like facility for exchange of timing advance without incurring the bandwidth overheads of a dedicated SACCH, GPRS uses a very innovative mechanism called the "continuous timing advance." In this technique, up to 16 users can know their timing advance using a very small fraction of the timeslot. For this, the PTCCH is used which has different structures in uplink and downlink. In uplink, an MS sends access burst over PTCCH/U, which enables the BTS to calculate the timing advance value for that MS. In the downlink, the BTS broadcasts timing advance value of up to 16 MS using a normal burst.

Every timeslot supports continuous timing advance procedures for up to 16 MS irrespective of whether they have been allocated an uplink or downlink TBF. To identify the position for an MS, the notion of Timing Advance Index (TAI) is defined. TAI can be viewed as an index, which in uplink identifies on which PTCCH/U the access burst is to be sent, and in downlink identifies the position where an MS has to look for its TA value. The TAI is informed to the MS through the packet uplink assignment or the packet downlink assignment message. In case the network does not send the TAI value, the MS does not use the continuous timing advance procedure. In such a case, the MS uses packet-polling procedure as discussed in the next section.

The whole structure for continuous timing advance procedure spans four 52-frames multiframes (416 TDMA frames or 1.92 seconds) that has 16 PTCCH slots. Figure 3.9 shows the relevance of TAI and the structure used for continuous timing advance. As shown, the TAI value ranges from 0 to 15. Thus, as per this structure, each MS gets one chance to send access

FIGURE 3.9

PTCCH/U, PTCCH/D, and Continuous Timing Advance



burst in uplink. However in downlink, the case is different. Each MS gets its TA value in every normal burst that carries timing advance value for that PTCCH timeslot. The normal burst is a structure having 23 octets (184 bits) with 1 octet for each of the 16 MS and 7 octets with "fill bits". For a TAI that is not allocated, a value of $0 \times 7F$ is used, or else the actual TA value is used.

If an MS is allocated different TAI values for simultaneous uplink and downlink packet transfer, the MS may choose to use any one or both PTCCH subchannels. If two subchannels are used, the MS shall always use the received TA value corresponding to the last transmitted PTCCH uplink burst.

3.9.3 Timing Advance Calculation Through Polling

Apart from continuous timing advance technique, the network can compute the timing advance through the polling mechanism. In this technique, the network sends PACKET_POLLING_REQUEST to the MS requesting the MS to send PACKET_CONTROL_ACK as four access bursts in the uplink. Through these access bursts, the timing advance is calculated and the same informed to the MS using the packet timing advance/power control message. The network can also poll the MS by setting the poll bit in the packet uplink/downlink assignment message. In this case too, the MS responds with PACKET_CONTROL_ACK by sending four access bursts. Multiple access bursts allow the scheme to be more robust. Moreover, when a block is allocated to an MS, it cannot be further shared between multiple MS.

In general, there are two options for sending PACKET_CONTROL_ACK in the uplink, either as four access bursts or through normal RLC/MAC control block. This is communicated to the MS through the parameter CONTROL_ACK_TYPE sent in GPRS cell options via PSI 1.

3.10 Power Control

Power control refers to controlling the transmission power of the MS and BTS on the air interface. Power control is important because it improves the spectral efficiency by reducing unnecessary interference in the air interface. In the uplink, it increases the battery life of the MS by allowing it to transmit at lower level when a higher power level is not needed. In the downlink, a lower transmission power allows the BTS to reduce its power consumption (and thereby reducing the power and operational costs). The principle of power control essentially entails sending at a lower power when the transmitting entity is closer to the receiving entity and vice versa.

In the uplink, the transmission power level of MS is controlled using the Uplink Power Control procedures. In the downlink, the transmission power level of BTS is controlled using the Downlink Power Control procedures. The following subsections describe the uplink and downlink power control procedures. For details of these procedures, reader is referred to [3GPP 45.008], [3GPP TS 43.064] and [3GPP 44.060].

1.10.1 Uplink Power Control

In the uplink, the transmission power level of MS is controlled using the Uplink Power Control procedures. The power control procedure is used to calculate the RF output power value, P_{CH}, to be used on each individual uplink PDCH assigned to the MS. The RF output power, P_{CH}, used by the MS on each individual uplink PDCH is calculated by the following formula:

$$P_{CH} = \min(\Gamma_0 - \Gamma_{CH} - \alpha * (C + 48), PMAX)$$

where the different parameters are as explained below.

- **Gamma constant (Γ_0):** This is the constant parameter that is 36 dBm (for GSM 900) or 39 dBm (for GSM 1800 or 1900 band).
- **Gamma for a channel (Γ_{CH}):** This is an MS and channel specific power control parameter, sent to the MS in an RLC control message (e.g., Packet Uplink Assignment, Packet Timeslot Reconfigure or Packet Timing Packet Power Control/Timing Advance messages). This parameter is represented as a 5-bit value that ranges from from 0 to 62 dB in steps of 2 dB. For those uplink PDCHs, for which Γ_{CH} has not been defined, value 0 is used.
- **Alpha (α):** This is a system parameter that varies from 0 to 1. It is broadcast on BCCH/PBCCH or optionally sent to MS in an RLC control message (Packet Uplink Assignment, Packet Timeslot Reconfigure or Packet Timing Packet Power Control/Timing Advance messages).
- **Average Signal Level (C):** This is the normalised received signal level at the MS.
- **PMAX:** This is the maximum allowed output power in the cell GPRS_MS_TXPWR_MAX_CCH if PBCCH exists otherwise it is MS_TXPWR_MAX_CCH. Note that the channel request on RACH/PRACH is sent using PMAX.

In essence, the power level is capped by PMAX implying that in no case the power can exceed the maximum allowed transmission power for the cell. In other cases, the transmission power in the uplink is calculated as a constant (Γ_0) minus the necessary factors ($\Gamma_{CH} + \alpha * (C + 48)$).

The exact usage of these parameters alpha and gamma results in two variants of uplink power control techniques, namely Open Loop Power Control and Closed Loop Power Control. In power control parlance, a 'loop' refers to the feedback, which originates from B to A due to transmission from A to B. Thus, a power control technique is 'closed' when network provides explicit feedback on the uplink received power level to be used. In contrast, if the network does not provide any explicit feedback about the uplink received power level, then the loop is said to be open.

Specifically, if power control is totally controlled by MS, it is called Open Loop Uplink Power Control. In contrast, in Closed Loop Uplink Power Control the network provides feedback to the MS such that MS can alter its transmission power. These techniques are further explained below.

3.10.1.1 Open Loop Uplink Power Control

In open loop power control technique, the uplink power control is based on the signal loss in the downlink direction with the assumption that the uplink loss is same as downlink power loss. Thus, the parameter α is set to a value 1 and the parameter Γ_{CH} is kept constant (i.e., network does not update the value of Γ_{CH} during data transfer). Thus, the uplink power reduces to Γ_0 - Γ_{CH} - α * C - α *48 = Constant – C. Thus, as the received signal level C decreases, the MS output power increases (implying that the MS is moving further away from the BTS and needs to increase the uplink power to compensate for the path loss. As the MS moves closer to the BTS, the received signal level increases and the MS reduces the uplink power. Note that the received signal level C has inverse relation with the output power (as represented by negative sign). Thus a decrease in C results in increase of output power and vice-versa. Further note that when the received signal level decreases and the ouput power increases, then ultimately a limit is reached where the output power cannot be increased any further. This limit is the PMAX power specified by GPRS_MS_TXPWR_MAX_CCH or MS_TXPWR_MAX_CCH.

3.10.1.2 Closed Loop Uplink Power Control

In closed loop power control technique, the uplink power control is based on feedback received from the network. In pure closed loop technique, α is set to a value 0. This renders the expression α * (C + 48) useless implying that the uplink power is based on the Γ_{CH}. The network uses the uplink measurements RXLEV and RXQUAL to explicitly send the updated power in downlink messages (e.g., Packet Timing Packet Power Control/Timing Advance) to signal the updated power value to the MS. Note that the value of α can vary between 0 to 1, where if one moves towards 0, the downlink path loss becomes insignificant while if one moves towards 1, the uplink measurement by the network is rendered useless. A value of α between 0 and 1 implies that a combination of open loop (based on received signal level) and closed loop (feedback from the network) power control techniques are used.

3.10.2 Downlink Power Control

In the downlink, the transmission power level of BTS is controlled using the Downlink Power Control procedures. While some of the techniques described in the previous section also apply to the BTS, there are additional considerations related to whether the PDCH is over BCCH/PBCCH frequency or not and on the type of the channel. These considerations are further explained below.

- **Downlink Transmission of PBCCH/PCCCH:** Since these channels are broadcast in nature and are applicable for cells anywhere in a cell, the power control techniques are not relevant for them. Thus, the BTS transmits at a constant power level (so that the power level is sufficient for all MS in the cell). This power level is the BCCH power level, which can be reduced by a power reduction fact Pb.
- **Downlink Transmission of PTCCH:** The PTCCH carries the timing advance information of up to 16 MS and it is unviable to perform power control as the MS may be

at various distances from the BTS. Thus, the PTCCH irrespective of whether it is carried on BCCH/PBCCH frequency or some other frequency is transmitted at BCCH power level. For PBCCH, the power reduction factor can be applied.

- **Downlink Transmission of PDTCH:** The case of PDTCH is slightly complicated because the power control algorithm has to not only take care of the power requirements of the MS for which the packet is intended but has to also consider the MS for which the USF is targeted. In other words, while applying power control the network has to ensure that the output power is sufficient for the MS for which the RLC block is intended as well as the MS for which the USF is intended under dynamic or extended allocation. If the power control is not used, then on BCCH frequency, the BCCH output is used; on other frequency, the transmission is on BCCH power with the optional power reduction or there can be no transmission when there is no data.

For more details on the power control technique used in the downlink, especially on the PDTCH, the reader is referred to [3GPP 45.008] and [3GPP 44.060].

Conclusion

This chapter provides an overview of GPRS air interface. GPRS reuses the basic air interface structure defined for GSM. This includes the TDMA/FDMA multiplexing scheme using which the BTS supports multiple frequencies, where each frequency has a recurring pattern of eight timeslots. The TDMA frame structure, the symbol period, and the TDMA repetition period are also there in GPRS as in GSM. Furthermore, no new burst structure is defined for GPRS and basic burst structure for normal and access burst is reused. If one looks at the frame structure, the dual structure concept is done away with GPRS and is replaced by a 52-frame multiframe that carries both signaling and traffic. In this 52-frame multiframe, there are 12 blocks of four slots, two idle slots and two used for handling timing advance. The logical channels in GPRS, although on similar lines, are different as compared to GSM. Various GPRS logical channels have been discussed in this book. Other aspects like channel coding and timing advance also have been covered. The related concepts are carried further in the next two chapters on RLC/MAC.

Further Reading

The GPRS air interface shares some of the literature with the GSM air interface and is described in the 45.abc series of GSM specifications. However, [3GPP TS 43.064] is the most important specification for the air interface. The specification first describes the packet data logical channels including broadcast, common control, and traffic channels. It then specifies how the logical channels map to the physical channels. Thereafter, the radio resource procedures are described. This section covers various aspects of air interface including master-slave concepts, radio resource operating modes, physical radio frequency layer and physical link layer, cell reselection, timing advance, power control, measurement reporting, and high-level aspects of RLC/MAC. This specification is actually a good starting point for the understanding of the GPRS air interface and also serves as a tutorial material.

[3GPP TS 45.001] is an introductory specification that describes the TDMA/FDMA structure used in GSM and provides general information on frame hierarchy (including hyperframe, superframe, and multiframe), bursts, and frequency hopping. Additionally, it describes the block structure of GPRS. The details, however, have not been covered in this specification.

[3GPP TS 45.002] is a more elaborate specification that describes various types of logical channels, and the permitted channel combinations. It also describes how the logical channels map to physical channels. It also specifies the paging group handling for GPRS.

[3GPP TS 45.003] specifies the channel coding rules for different logical channels and coding schemes (CS1 to CS4) including rules for USF precoding, parity and tail bits computation, the interleaving rules, and rules of how channel-coded information is mapped to a burst.

[3GPP TS 45.004] is a very small specification that describes the GMSK modulation technique used in GSM and GPRS.

[3GPP TS 45.005] describes the transmission and reception rules for both BTS and MS over the air interface. In particular, it deals with the topics like output power of MS/BTS for different GSM/GPRS systems, output RF spectrum, spurious emissions, radio frequency tolerance, and modulation tolerance.

[3GPP TS 45.008] is another important specification that deals with radio subsystem link control. The key topic covered in this specification includes the GPRS idle mode handling.

[3GPP TS 45.010] deals with synchronization aspects in GSM/GPRS and covers topics like timing advance.

Chapter 4

Medium Access Control Layer

4.1 Introduction

In Chapter 3, the air interface used for GPRS (i.e., the physical layer) has been discussed. This chapter takes the next step and describes the link layer used to manage the physical layer. This link layer in GPRS comprises of twin protocols, Radio Link Control (RLC) and Medium Access Control (MAC). The MAC layer is predominantly responsible for the management of the shared transmission resources. It provides the mechanism to establish the Temporary Block Flow (TBF) that allows point-to-point transfer of the signaling information and data between the MS and the network. Additionally, the MAC layer also controls the reception of the broadcast information from the network. On the other hand, the RLC layer acts as an interface between the upper layers and the MAC layer for the transmission of the upper layer Protocol Data Unit (PDU). It provides flow control and error control during the data transfer. It also performs the segmentation and the reassembly of the upper layer PDUs. The RLC/MAC layer is completely responsible for radio resource management. This chapter provides a detailed description of the MAC procedures. The RLC layer procedures and the block formats used for RLC/MAC are discussed in Chapter 5.

To understand MAC, the chapter is organized as follows. It first discusses the MAC layer procedures. The "modes of operations" and the procedures performed as part of these modes of operation are discussed. The concepts that allow the multiplexing of the radio resources are discussed in detail. The reception of the broadcast information has been elaborated in the next section. The subsequent sections introduce the concepts of paging and the cell selection.

The later sections of the chapter are dedicated to the signaling procedures associated with the TBF. The procedures for the establishment and release of the uplink and the downlink TBF are discussed in detail.

4.2 Multiplexing Principles

The most important aspect of the GPRS technology is the ability to share the same set of resources among multiple users. This feature, while increasing the bandwidth usage, increases the implementation complexity considerably. This section deals with the concepts that allow this multiplexing of resources. These concepts lay the foundation as to how MAC layer operates.

4.2.1 Temporary Block Flow

In the circuit-switched technologies, the timeslots are allocated to a user dedicatedly for the entire duration of the call, irrespective of whether these resources are being used or not. This leads to inefficient utilization of resources especially when the radio resources are not needed for the entire duration. Typically, the data transfer between the MS and the data networks is bursty in nature due to the concept of push and pull of data flow between them. As a result, a dedicated allocation of the radio resources would lead to highly inefficient utilization of the radio interface. In order to solve this problem, the timeslots in GPRS are allocated to an MS only when needed (i.e., the resources are allocated on a temporary basis). Moreover, the timeslots can be shared between multiple users at the same time. Furthermore, the transfer of data takes place in the form of data blocks (where one block is equal to four bursts). Such a transfer of the upper layer PDUs between two entities on the packet data physical channels is called a TBF. Formally, a TBF is a physical connection used by two radio resource entities to support the unidirectional transfer of Logical Link Control (LLC) PDU on packet data physical channels. The TBF is allocated to radio resource on one or more PDCHs and comprises a number of RLC/MAC blocks carrying one or more LLC PDUs. A TBF is temporary and maintained only for the duration of the data transfer. Managing the TBF is the central function of RLC/MAC protocol.

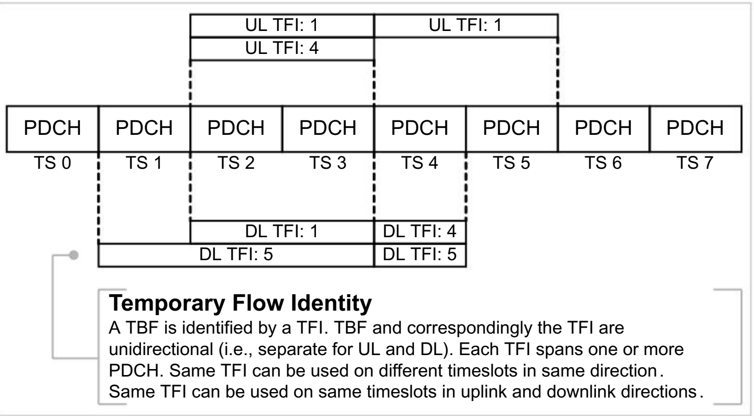
The flow of data may be from the MS to the network or from the network to the MS. The data flow from the network to the MS constitutes the “downlink TBF” whereas the data flow from the MS to the data network constitutes the “uplink TBF”. An MS may be involved in an uplink and a downlink TBF at the same time. In this case, the MS is said to be involved in a “concurrent TBF”.

TBF is unidirectional in nature and is allocated for each direction separately. One or more packet data channels may be used for a TBF (i.e., more than one timeslot may be allocated for the same TBF). This depends on the allocation made by the network for the TBF. The number of timeslots that are allocated to an MS for a TBF further depends on the multislot class of the MS. Also, different TBFs may be active on the same channel. In other words, the same physical channel may be allocated to different MS and hence different TBFs. In order to differentiate between the TBFs, a unique identity is allocated to each TBF, known as the Temporary Flow Identifier (TFI). The network allocates the TFI when resources are allocated for the TBF through resource assignment message that precedes LLC transfer. The same TFI is included in every RLC header belonging to a particular TBF as well as in the control messages associated to the LLC frame transfer (e.g., acknowledgments) in order to address the peer RLC entities. A downlink TBF is identified by a "Downlink TFI" and an Uplink TBF is identified by an "Uplink TFI". The TFI values are unique on a channel in a particular direction (in the uplink or downlink direction). In other words, the same value of TFI can be used for TBFs on different timeslots in the same direction or for TBFs on the same timeslot in different directions. The following example would elaborate the concept.

Figure 4.1 depicts the allowed usage of the TFI values for different TBFs established on same or different PCDH(s). In the example presented, three uplink and four downlink TBFs are active. The UL TFI denotes the TFI allocated to the uplink TBF, whereas the DL TFI denotes the TFI allocated to the downlink TBF. The PDCH 2 and 3 have one downlink TBF (identified by the TFI value 1) and two uplink TBFs (identified by the uplink TFI value 1 and 4) established on them. Hence, the downlink and uplink TBF on the same PDCH can share the same value of the TFI.

Figure 4.1

Example for the Possible TFI Values for TBF on Same/Different PDCH



Also, the PDCH 1, 2, and 3 are allocated to a downlink TBF identified by TFI value 5. Similarly a downlink TBF on the PDCH 4 is also identified by TFI value 5. Hence the TBF in the same direction but on different PDCH can share the same value of the TFI. A similar example is the uplink TBF established on PDCH 2 and 3 and PDCH 4 and 5, which are allotted the same TFI value of 1.

The TFI values allow the sharing of the same radio resources among different MS. In the downlink direction, the MS monitors the allocated PDCH for data blocks containing the allocated values of the downlink TFI. In the above example, let the MS sharing the PDCH 4 for the downlink TBF be MS 1 (TFI = 4) and MS 2 (TFI = 5). Both the MS 1 and MS 2 would monitor the PDCH 4 for the downlink data blocks and decode the downlink TFI encoded in these blocks. If the decoded TFI value is 4, the MS 1 would continue to decode the data contents of the block and MS 2 would discard this data block. Similarly, if the decoded TFI value is 5, the MS 1 would discard this block and the MS 2 would decode the data block.

Allocating different TFI values hence allows the network to share the same PDCH between multiple downlink TBFs.

In the uplink direction, the TFI values are encoded in the uplink blocks being transmitted by the MS. The sharing of resources is made possible by another parameter, called the Uplink State Flag (USF). This concept is introduced in the next section.

4.2.2 Uplink State Flag

The dynamic allocation of resources is the most important feature that allows the multiplexing of resources in GPRS. The network controls the transmission of the data blocks in the downlink direction, and the MS identifies the blocks by the encoded TFI values. However, in order to allow the sharing of resources in the uplink direction, a mechanism is devised to control the transmission of data from MS sharing the same PDCH. This mechanism is based on the use of the parameter USF that allows dynamic allocation of resources in GPRS and sharing of the uplink resources.

During the uplink resource allocation, the network may distribute resources dynamically to the MS for the requested uplink TBF. As part of the allocation details, the MS is informed about the allocated timeslot(s), the uplink TFI that identifies this TBF, and another unique identifier used for dynamic allocation, the USF. The details of this allocation are discussed in the subsequent sections of this chapter.

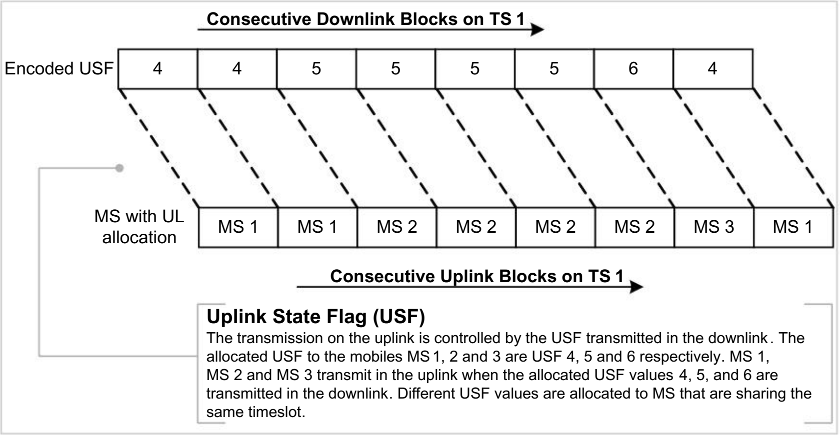
After receiving the uplink allocation, the MS continuously monitors the allocated timeslots and decodes the downlink blocks being transmitted on these timeslots for the USF value. If the decoded value is the same as the USF value allocated to the MS, the MS transmits on the next radio block in the uplink. If, however, the decoded value is not the same as the allocated USF value, the MS does not transmit any data and keeps monitoring the downlink blocks for its allocated USF value. In this case, another MS that has been allocated to this USF value transmits on the next uplink radio block. The USF hence allows the network to share the same resource between multiple MSs. The concept of resource sharing in the uplink direction is elaborated in the following example.

Consider a scenario in which a network receives requests from three MSs to initiate packet data transfer. The network allocates the same timeslot to the MS, say TS 1, and different USF values, 4, 5, and 6 to the MS 1, MS 2, and MS 3 respectively. The transfer of the data would be controlled by the USF values transmitted by the network as depicted in the Fig. 4.2.

In the above example, the transmission of the uplink radio blocks as controlled by the USF values in the downlink direction is depicted. Here, the USF values 4, 5, and 6 are allocated to the MS 1, MS 2, and MS 3, respectively. As shown in Fig. 4.2, the USF encoded in first radio block is 4. This triggers the uplink transmission from MS 1 in the next uplink radio block. Similarly, USF value 4 is encoded in the next radio block triggering the uplink transmission from MS 1. However, the third downlink radio block contains the USF value 5, which results in the transmission from MS 2 on the next uplink radio block. The subsequent downlink radio blocks contain the USF values 5, 5, 5, 6, and 4 that result in the uplink transmission from MS 2, MS 2, MS 2, MS 3, and MS 1, respectively.

FIGURE 4.2

Example for the Uplink Transmission Controlled by USF Values



Another concept that deserves attention is the USF granularity. The USF granularity is communicated to the MS at the time of the resource assignment (i.e., packet uplink assignment message). Depending upon USF granularity, when an MS decodes the assigned USF in the downlink block, it transmits on one or four consecutive radio blocks.

4.3 Modes of Operation

There are different modes of operation for the MS depending on whether the MS is involved in a PS data transfer or any radio resource connection.

- Packet idle mode. If an MS is not a part of any ongoing TBF, it is said to be in the Packet Idle (PI) mode. In this mode, the MS keeps monitoring the relevant paging subchannels for any TBF being initiated for this MS. The paging subchannels may be on the Packet Common Control Channel (PCCCH) or the Common Control Channel (CCCH), depending on the network configuration. If a data transfer is initiated, the TBF establishment is triggered and the MS moves into the Packet Transfer (PT) state.
- Packet transfer mode. If an MS is a part of an ongoing TBF on one or more packet data physical channels, it is said to be in PT mode. The TBF may be uplink or downlink or the MS may be involved in a concurrent TBF. When the data transfer of the upper layers terminates, the TBF is released. In case all the TBFs have been released, the MS moves into the PI state.
- Dual transfer mode. If an MS is involved in an radio resource connection and ongoing TBF simultaneously, it is said to be in Dual Transfer Mode (DTM). In this case, the dedicated traffic channels are allocated to the MS for the radio resource connection and one or more PDCH(s) are allocated to the MS for the ongoing TBFs. The network allocates the resources to the MS depending upon the MS DTM capabilities. When the upper layer PDU has been successfully transferred, the TBF is released. In case all the ongoing TBFs have been released, the MS moves from the DTM to the dedicated mode.

However, if in the DTM mode the radio resource connection of the MS is terminated, all the ongoing TBFs are released and the MS moves into the PI mode.

4.4 Broadcast Information Management

The network continuously broadcasts the information about the serving cell and the neighbor cells on the broadcast channel. This information is important for the MS to determine how to access the network. The information is broadcast in the System Information (SI) messages on the Broadcast Channel (BCCH) and the Packet System Information messages (PSI) on the Packet Broadcast Channel (PBCCH). The first and the second subsections detail the information broadcast in the messages on the BCCH and the PBCCH, respectively. The third subsection deals with the scheduling of these messages on the BCCH and the PBCCH. The last subsection provides details of the acquisition mechanisms used by the MS to read these messages.

4.4.1 System Information Broadcasting on BCCH

The network continuously broadcasts information on the BCCH. When the MS camps into a cell, it reads the broadcast information in order to determine if it can camp on the cell. Additionally, the MS uses these parameters to access the network. As part of the broadcast information, a set of SI messages is broadcast. Some of these messages are mandatory (SI 2 – 4), while the others are broadcast optionally (SI 1, 2 bis, 2 ter, 7, 8, 13, 16, and 17).

The BCCH is used for broadcasting the GSM and the GPRS network parameters. The following messages provide all the necessary information:

- **SI 1 message.** This optional message provides information about the Random Access Channel (RACH) control parameters and the frequencies that can be used in a particular cell.
- **SI 2 message.** This mandatory message provides information about the RACH control parameters and the BCCH allocations of the neighboring cells.
- **SI 2 bis and 2 ter message.** These are optional messages that provide information on the extension of the BCCH allocation of the neighboring cells.
- **SI 2 quat message.** This optional message provides information about the additional measurement and reporting parameters and also the UTRAN neighbor cells.
- **SI 3 message.** This mandatory message provides information about the RACH control parameters, the cell identity, the location area identifier, and the cell selection parameters.
- **SI 4 message.** This mandatory message provides information about the Cell Broadcast Channel (CBCH) parameters in addition to the cell identity, the location area identification, and the cell selection parameters.
- **SI 7 and 8 message.** These optional messages provide information about the cell reselection parameters to be used in the cell.

- **SI 9 message.** This optional message provides information about the BCCH scheduling.
- **SI 13 message.** This message is transmitted on the BCCH if GPRS is supported in the cell. The details about the PBCCH are provided in this message.
- **SI 16 and 17 message.** These messages provide information about the cell selection and the reselection parameters.

There are other SI messages transmitted on the BCCH that provide non-GSM information to the MS. Additionally, there are other SI messages that are transmitted on the Slow Associated Control Channel (SACCH). The GSM SI is briefly discussed here as the focus is on the SI for GPRS. For details of GSM SI, the reader is referred to [GSM N. Narang].

4.4.2 System Information Broadcasting on PBCCH

As discussed in the previous section, the GPRS capability of the cell is indicated by the network in the SI message, SI 13. This SI message provides information about the GPRS broadcast channel, the PBCCH. The PBCCH may or may not be present in the cell. If the PBCCH is not present, the BCCH is used to broadcast the GPRS parameters. However, if the PBCCH is present, the PSI messages are used to broadcast the GPRS parameters. In this case, the MS reads the PSI messages.

The PSI messages provide the following information:

- Packet System Information 1. This message provides information about the control of PRACH, the description of the PCCCH, the power control parameters, and the cell selection parameters.
- Packet System Information 2. This message provides information about the Cell Allocations (CA) and the mobile allocations that are applicable in the cell. This message also provides information about the other PSI messages that would be broadcast on the PBCCH.
- Packet System Information 3. This message provides information about the BCCH allocation in the neighbor cells and the cell selection parameters for serving cell and the nonserving cells.
- Packet System Information 3 bis. This message provides information about the BCCH allocations in the neighbor cells and the cell selection parameters for nonserving cells.
- Packet System Information 3 ter. This message provides information on additional measurement and reporting parameters.
- Packet System Information 3 quat. This message provides information about the 3G neighbors and the additional reporting and measurement parameters.
- Packet System Information 5. This message provides information about the network-controlled cell reselection and the measurement reporting.
- Packet System Information 13. This message contains the same information that is contained in SI 13.

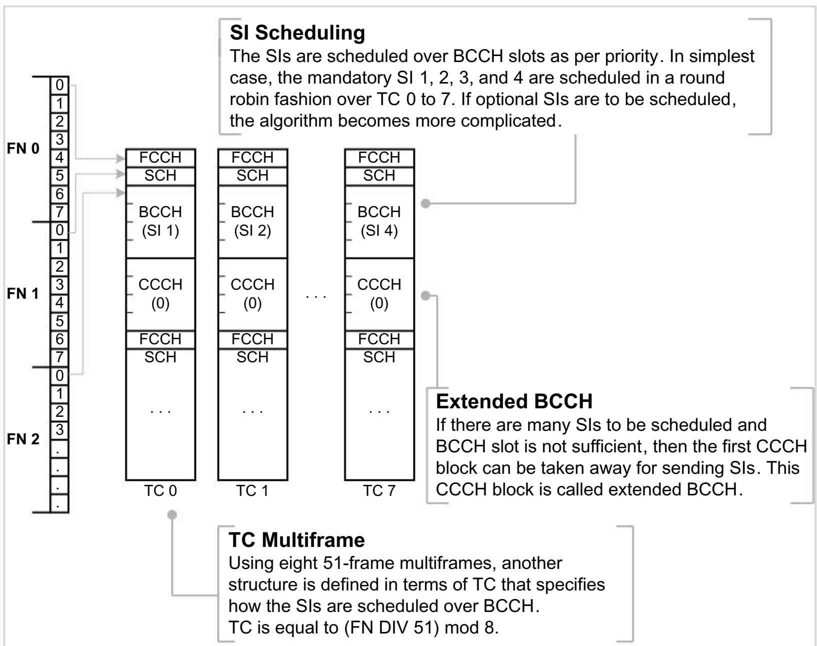
There are other PSI messages that are used for broadcasting information about the non-GSM networks.

4.4.3 Scheduling of the SI and PSI Messages

The information broadcast in the SI messages (on BCCH) and the PSI messages (on PBCCH) provides MS the information needed by it to access the network. If one considers the BCCH, then there are many types of SI messages, but the 51-frame structure has space for only one block of four bursts (2–5) to carry BCCH messages. Thus, some form of scheduling algorithm is required that manages transmission of SI messages on different BCCH channels. [3GPP TS 45.002] provides some rules for transmission of SI messages over BCCH. While the list of rules is quite long, the summarized view is explained next (see Fig. 4.3).

FIGURE 4.3

Scheduling of System Information on BCCH



Typically, the SI messages are scheduled over the BCCH channel of the BCCH frequency on the timeslot 0. Every cell has only one BCCH frequency, and therefore there is only one BCCH available for the transmission of SI. However, as a network option, the SI messages can also be scheduled over the first CCCH block that is reserved for Access Grant Channel (AGCH). Under such circumstances, the parameter BS_AG_BLKS_RES is set to a value more than 0 because the first CCCH block is taken away for SI and the paging messages cannot be transmitted over it. The first CCCH block taken for the purpose of SI transmission is called "Extended BCCH".

Furthermore, to impart some order in scheduling of SI, a variable is introduced called the TC, which is equal to {(FN DIV 51) mod 8}. In simple terms, TC divides the multiframes in groups of eight. Based on the TC value, the scheduling rules are as given in Table 4.1. The

Table 4.1

Scheduling Rules for Different SI

SI Type	*TC*	*Allocation*
SI 1	0	BCCH Normal
SI 2	1	BCCH Normal
SI 2 bis	5	BCCH Normal
SI 2 ter	5 or 4	BCCH Normal
SI 2 quater	(5 or 4) or	BCCH Normal or
	5	BCCH Extended
SI 3	2 and 6	BCCH Normal
SI 4	3 and 7	BCCH Normal
SI 13	4 or	BCCH Normal or
	0	BCCH Extended

important SIs include SI 1, 2, 3, and 4. Thus, if they are to be scheduled only, then the scheduling policy would be {1, 2, 3, 4, 1, 2, 3, 4} as per Table 4.1. In this example, note that on TC 4, in absence of any other SI (i.e., SI 2 ter, SI 2 quater and SI 13), the SI 1 can be sent instead. SI 2 quater and SI 13 can be sent on BCCH normal or BCCH extended. For detailed scheduling rules, refer [3GPP TS 45.002].

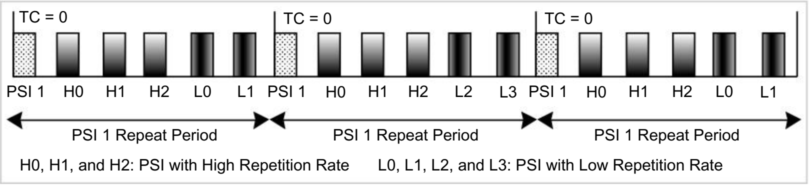
Similarly, the rule for scheduling on the PBCCH is governed by the following formula: TC = (FN div 52) mod PSI1_REPEAT_PERIOD, where PSI1_REPEAT_PERIOD is between 1 and 16 and is indicated in the SI 13 message on BCCH. As the name suggests, the PSI1_REPEAT_PERIOD indicates the repetition rate of the PSI 1 which is transmitted every time the TC is 0. Apart from PSI 1, there are two other categories of PSI, the messages with high repetition (indicated by PSI_COUNT_HR with a value of 0–16) and the ones with low repetition (indicated by PSI_COUNT_LR with a value of 0–63). Then the scheduling abides by the following four rules (see Fig. 4.4):

1. PSI 1 shall be sent in block B0 when TC = 0.
2. If the value of the parameter BS_PBCCH_BLKS is greater than 1, the PSI 1 shall also be sent in block B6 when TC = 0.
3. The PSI messages in the group sent with high repetition rate shall be sent in a sequence determined by the network and starting at TC = 0, using the PBCCH blocks within each multiframe in the order of occurrence, which are not occupied according to rule 1 or 2. The sequence of these PSI messages shall be repeated starting at each occurrence of TC = 0.
4. The PSI messages in the group sent with low repetition rate shall be sent in a sequence determined by the network and continuously repeated using the PBCCH blocks within each multiframe in the order of occurrence, which are not occupied according to rules 1–3. The sequence of these PSI messages shall be restarting at FN = 0.

The above rules are explained through the example shown in Fig. 4.4. In this example, the PSI1_REPEAT_PERIOD is 6 indicating that the PSI 1 message is repeated after every 6 multiframes. The parameter PSI_COUNT_HR is 3 indicating that there are three messages

FIGURE 4.4

Scheduling of System Information on PBCCH



sent at high repetition rate in every cycle of six frames. Lastly comes the low-priority messages that have the PSI_COUNT_LR of 4 indicating that the L0 and L1 messages are sent in one frame and L2 and L3 are sent in the next frame.

4.4.4 Acquisition of System Information by MS

The MS reads the information broadcast on the BCCH to find the network details and various parameters discussed previously. These parameters are important for the MS to access the network. The SI 13 message on the BCCH indicates whether PBCCH is defined in the serving cell or not. In case the PBCCH is defined and the MS wants to access the GPRS services, the MS reads the PBCCH information. On the other hand, if the PBCCH is not defined in the cell, the MS reads the information broadcast on the BCCH. Here, the monitoring of the BCCH information is discussed first and then the monitoring of the PBCCH information.

If the PBCCH is not present in the cell, the MS reads the SI messages. If the MS has selected a new cell that does not have a PBCCH, it performs "complete acquisition of the SI" messages. The MS cannot perform packet access unless it has acquired SI 3 and, if present, SI 1 messages. Also, the MS must make at least one attempt to acquire all the SI messages that may be transmitted within a single TC cycle on the BCCH.

However, if the PBCCH is present in the cell, the MS shall receive the PSI messages. When a new cell has been selected that has PBCCH defined in it, the MS performs a "complete acquisition of the PBCCH messages". The MS cannot perform packet access until

- it has acquired PSI 1 message
- it has acquired a consistent set of PSI 2 messages
- it has made at least one attempt to receive the complete set of PSI messages on the PBCCH

The MS attempts to receive the PSI 1 message once in every 30 seconds. The information about the scheduling of the other PSI messages is contained in the PSI 1 message, which allows the MS to acquire these messages. It is possible that some of the network parameters may be modified hence resulting in a change in the SI being broadcast on the channel. It may happen that some of the PSI messages get modified while the other messages remain the same. In such a case, the PSI 1 message provides information about the exact messages that have been modified so that only these messages are acquired by the MS. This procedure is called the "partial acquisition of PBCCH information". It is controlled by the PSI 1 message, which contains PBCCH_CHANGE_MARK and the PSI_CHANGE_FIELD. If the value of the

PBCCH_CHANGE_MARK is incremented by one from its previous value, then the MS performs the partial acquisition of the PSI messages. The PSI_CHANGE_FIELD provides the information about the exact message(s) the MS must read:

- If the parameter indicates change in a specified PSI message(s), then the MS acquires at least one instance of these message(s).
- If the parameter indicates change in an unspecified PSI message(s), then the MS acquires at least one instance of each message belonging to a consistent set of PSI messages on the PBCCH.
- If the parameter indicates change in an unknown type of PSI message, the MS does not acquire any PSI message.

If the value of the PBCCH_CHANGE_MARK is incremented by more than one from its previous value, then the MS performs a complete acquisition of the PSI messages.

4.4.5 Frequency Parameters

When the MS is involved in an uplink or a downlink data transfer, the network allocates radio resources to the MS. These resources are allocated in the assignment messages. In addition to the radio resource allocation, the assignment message also provides information on the frequencies to be used for the TBF.

The set of all the frequencies that can be used in the serving cell constitute the Cell Allocation, (CA). On the other hand, the frequencies that can be used by the MS for a particular TBF are known as the Mobile Allocation (MA). The information about the CA is broadcast in the PSI. The information about the MA is provided in the assignment messages, the Packet Uplink Assignment and the Packet Downlink Assignment depending on the type of TBF that is being initiated. The frequency parameters in the assignment message may contain the definition of the nonhopping channels or the hopping channels. In case of nonhopping frequency parameters, a single Absolute Radio Frequency Channel Number (ARFCN) is defined, whereas in case of hopping channel, a set of frequencies is defined to be used for the TBF.

4.5 Paging Procedures

If the network has any data to be transferred to an MS, a downlink TBF has to be established. If the MS is in PI mode, the network pages the MS in order to initiate the cell update procedure by sending an LLC frame. This is because the network does not know the cell-level details of the MS (i.e., it only knows the Routing Area information). Thus, the network pages the MS in all cells of the Routing Area in which the MS is currently located. The downlink data transfer is then started for the MS. In case the MS is in PT mode, the network does not need to page the MS. In this case, the downlink TBF is initiated by downlink TBF establishment procedures that are discussed in the later sections of this chapter.

4.5.1 Network Operating Modes

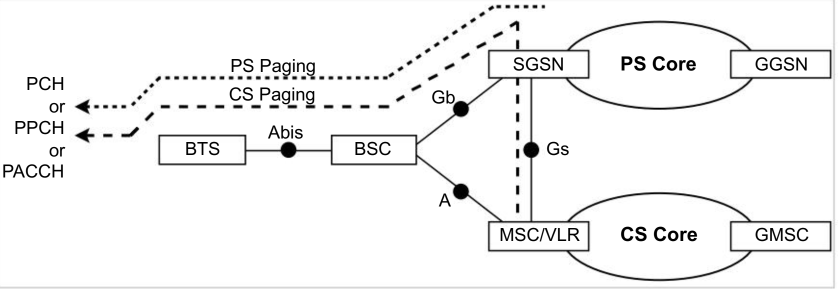
The MS can be paged for the circuit- or the packet-switched services using a CS or a PS paging message, respectively. The CS paging is initiated by the MSC/VLR and the PS paging is initiated by the Serving GPRS Support Node (SGSN). The paging messages for these procedures may be sent on the paging subchannel of the PCCCH or the CCCH. In the absence of any coordination, the MS may have to look for paging messages on different paging channels. To simplify the work of MS, the concept of paging coordination is defined. Paging is said to be coordinated when the Mobile-Service Switching Center/Visitor Location Register (MSC/VLR) and SGSN have a Gs interface that is used to send paging messages from MSC/VLR to MS via SGSN.

The concept of paging coordination and the channel to be used for the paging message is linked to three factors. First of all, this depends upon whether PCCCH is supported in the network or not; the second factor is the state of the MS (i.e., whether the MS is in PT mode or PI mode); and lastly, it depends upon the Network Mode of Operation (NMO) being used by the network. The MS listens to the SI type 13 (or PSI 1 if PBCCH is supported) to know the value of NMO carried in the GPRS Cell Options information element. Based on the value of NMO, the MS listens to the appropriate paging subchannel. There are three modes of operation for the network that are as follows:

- **NMO I.** In this mode, the CS paging messages for a GPRS-attached MS are transmitted on the same channel that is used for the PS paging messages (i.e., paging subchannel on the PCCCH if it is supported or the paging subchannel on the CCCH). The paging message may also be transmitted on the GPRS data channel (i.e., PACCH) if the MS is in PT mode. Thus, in this case, the MS needs to monitor only one channel for the paging messages. For this mode, it is mandatory to support the Gs interface between MSC/VLR and the SGSN that can be used for paging coordination. Figure 4.5 shows the CS and PS paging messages in this mode.
- **NMO II.** In this mode, the paging subchannel on the CCCH is used for both the PS as well as the CS paging messages. In this case, the MS needs to monitor only the CCCH paging subchannel. The Gs interface is not applicable in this mode.
- **NMO III.** In this mode, the CS paging messages are transmitted on the paging subchannel on the CCCH. The PS paging messages can be transmitted on the CCCH or

FIGURE 4.5

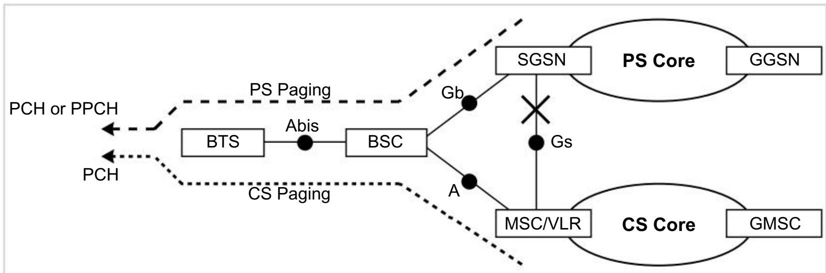
Paging Message Transfer in NMO I



the PCCCH, if the packet channels are defined in the cell. The MS, hence, needs to monitor the paging subchannels on the CCCH and the PCCCH if it needs to receive both the CS and the PS paging messages. The Gs interface and paging coordination is not applicable in this mode. Figure 4.6 shows the CS and PS paging messages in this mode. The scenario for NMO II is similar to the one depicted in Fig. 4.6 except that the channels are only PCH (as PPCH are not applicable in NMO II).

FIGURE 4.6

Paging Message Transfer in NMO III



4.5.2 Discontinuous Reception (DRX) Mode on CCCH

An MS in idle mode is paged in order to initiate a packet data transfer or to establish a radio resource connection. The MS thus needs to monitor the downlink paging channels on the CCCH or the PCCCH, whichever is applicable. This involves a lot of power consumption for the MS. In order to reduce the power consumption, paging groups are defined and an MS listens to the paging subchannels corresponding to its paging group. The MS is then said to work in DRX mode. To operate in DRX mode, the MS needs to calculate the CCCH_GROUP (or PCCCH_GROUP for GPRS) and the paging group for CCCH and the PCCCH. This section elaborates upon DRX on CCCH, while the next section discusses DRX on PCCCH.

In order to operate in DRX mode on CCCH, the mobile needs to calculate the CCCH_GROUP and the paging group for CCCH. The CCCH is defined on the BCCH frequency of the serving cell. The CCCH channels are generally carried with the BCCH on the timeslot 0 (TS0). However, there can be more than one CCCH in the same cell. In this case, the first CCCH is carried on TS0, the second on TS2, the third on TS4, and the fourth on TS6. Each of the CCCH carries the paging messages for the mobiles belonging to its CCCH_GROUP. Thus mobiles belonging to a particular CCCH group listen to the CCCH belonging to that CCCH_GROUP.

In order to calculate the CCCH_GROUP and the paging group, the mobile requires information about the control channel defined in the cell. As discussed earlier, the description of the control channels is provided by the network in the SI 3, which is broadcast on the BCCH. As part of the broadcast information, the following parameters are provided to the mobile:

- **CCCH_CONF.** This parameter has five different values. Value of 0 indicates that the Stand-Alone Dedicated Control Channel (SDCCH) is not combined with CCCH, while value of 1 indicates that SDCCH is combined with CCCH. Then three other possible values (4, 6, and 7) indicate that there are two, three or four CCCH on the BCCH frequency. In summary, the parameter defines whether there is one or more CCCH; and if there is one CCCH, then whether it is combined with SDCCH or separate. The number of CCCH timeslots on the BCCH frequency is referred to as BS_CC_CHANS.

- **BS_AG_BLKS_RES.** This variable indicates the number of blocks reserved for the AGCH. In noncombined configuration where SDCCH is separate, this variable has values of 0–7, while in combined configuration, this variable has a value of 0–2. Note that a value of 0 does not imply that there is no message sent on AGCH. It merely implies that there are no exclusive blocks for AGCH. The messages for AGCH can be sent over any CCCH (including those that carry paging). However, paging messages can never be carried over blocks reserved for AGCH.
- **BS_PA_MFRMS.** Apart from parameters defined earlier, another important parameter in the computation of paging group is the BS_PA_MFRMS. This parameter defines the number of 51-frame multiframe between the transmission of paging messages to mobiles belonging to the same CCCH_GROUP. Thus, if the value of this variable is 3, then after three 51-frame multiframes, an MS of the same paging group gets a chance. This variable ranges from 2–9 resulting in a paging cycle of 0.47 seconds when a paging group is handled after every two 51-frame multiframes. The paging cycle is 2.2 seconds in case the cycle is maximum and a paging group is handled after every nine 51-frame multiframes. A higher value results in higher response time, while a lower value implies higher battery consumption. Thus, a balance is attained between the two.

In summary, the first two parameters (CCCH_CONF and BS_AG_BLKS_RES) allow the mobile to determine the number of paging blocks available in one 51-frame multiframe. The third parameter (BS_PA_MFRMS) determines the periodicity of the paging messages. The total number of available paging blocks on one CCCH is hence the number of paging blocks available in one 51-multiframe on one CCCH multiplied by BS_PA_MFRMS.

The mobile calculates the CCCH_GROUP using the following formula:

CCCH_GROUP (0 .. (BS_CC_CHANS-1)) = ((IMSI mod 1000) mod (BS_CC_CHANS x N)) div N

where,

IMSI = International Mobile Subscriber Identity

BS_CC_CHANS = Number of CCCH timeslots on the BCCH frequency

N = Number of paging blocks "available" in one 51-multiframe on one CCCH × BS_PA_MFRMS

In simple terms, the above formula for calculating CCCH_GROUP allows the mobile to determine the exact physical channel on which the paging for it would occur. The value of CCCH_GROUP can vary between 0 to (BS_CC_CHANS –1), implying that the MS gets its paging on TS = 0, or TS = 2 and so on depending upon how many CCCH timeslots are defined in the network.

Besides the CCCH_GROUP, the MS has to determine the paging subchannel to be monitored. This subchannel can be viewed as the exact block within a paging cycle on which an MS that belongs to a paging group can be paged. The paging group is calculated using the following formula:

Paging Group = ((IMSI mod 1000) mod (BS_CC_CHANS × N)) mod N, where all parameters are defined earlier.

In simple terms, mod 1000 is applied to IMSI and again mod N operation is performed to the result (assuming that there is only one CCCH timeslot). In other words, the last three digits of the IMSI are used and divided into *N* groups, where *N* is the number of paging blocks available in a paging cycle.

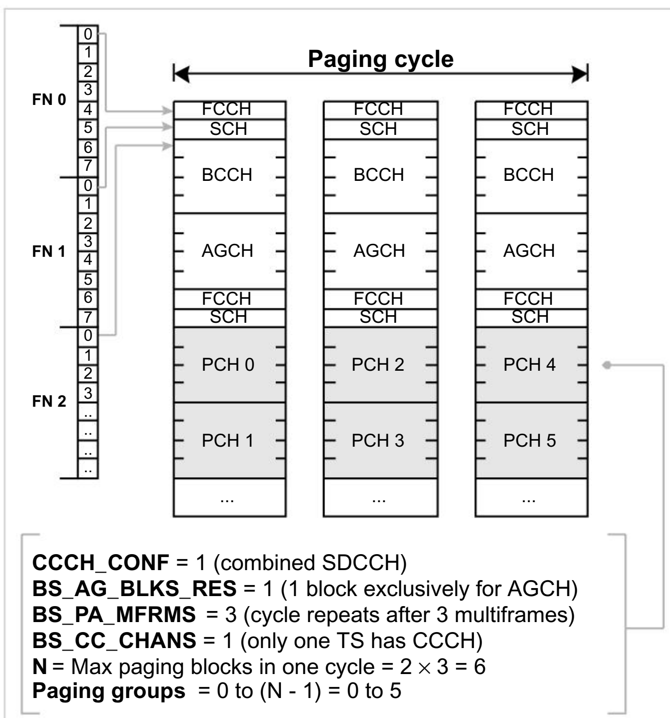
Figure 4.7 depicts a simple example of the use of various parameters that are a part of the DRX technique. Sample values are also provided in this figure. As shown in the example, the MS belongs to any of the six paging groups (0–5), calculated using the simplified formula ((IMSI mod 1000) mod N).

4.5.3 Discontinuous Reception Mode on PCCCH

If the packet channels are defined in the serving cell, the network pages the MS on the PCCCH in order to initiate the downlink data transfers. The paging channels are divided into paging subchannels, and the messages for a particular MS are always sent on a particular subchannel. As discussed in the previous section, to reduce the battery consumption, an MS

FIGURE 4.7

Example Illustrating the Calculation of Paging Group and Paging Cycle



can listen to only its own paging subchannel for the paging messages. This mode is known as the DRX.

Just like the CCCH, DRX mode on PCCCH requires the MS to know two parameters, namely the PCCCH group and the paging group. Thus, the broad frameworks of DRX technique on CCCH and PCCCH are similar. However, one important difference is that while CCCH can operate on multiple timeslots of only one carrier (i.e., the BCCH carrier), the PCCCH in contrast can exist on multiple timeslots of multiple carriers. The knowledge of a particular timeslot (i.e., the PDCH in GPRS parlance) of a particular carrier identifies the PCCCH group of the MS. Further, the paging channels on a paging PDCH are divided into subchannels and the paging messages for a particular MS are sent on a subchannel depending on the IMSI of the MS. These subchannels are also known as paging groups. Hence, in order to work in the DRX mode, the MS first needs to identify the carrier (i.e., ARFCN) and the PDCH that is carrying its paging messages, denoted by the PCCCH_GROUP, and then identify the paging subchannel, denoted by the PAGING_GROUP.

As discussed earlier, the PCCCH may occur on multiple carriers and multiple PDCHs on each carrier. The network broadcasts the PCCCH description in the cell in the PSI 2 message. The description provides information about the training sequence code to be used on the PCCCH carriers. It also provides the details of the nonhopping PCCCH carriers that include the ARFCN and the timeslot or information of the hopping PCCCH carriers that include the mobile allocation number, mobile allocation index offset, and the timeslot.

To identify the PCCCH_GROUP, a parameter BS_PCC_CHANS is used that denotes the number of PDCH that carry the PCCCH, similar to the BS_PCC_CHANS that identifies the number of CCCH timeslots on the BCCH frequency that carry the CCCH. The maximum value for this parameter BS_PCC_CHANS is 16 and so the PCCCH_GROUPs can range from 0 to BS_PCC_CHANS-1. The PCCCH_GROUPs are mapped to the different PDCH on the basis of the PCCCH description that is broadcast by the network. The lowest-numbered PCCCH_GROUP is mapped to the lowest PDCH carrying the PCCCH on the first carrier (hopping or nonhopping) mentioned in the PCCCH description, the next higher numbered PCCH_GROUP is mapped onto the next higher numbered PDCH of the carrier, and so on. Once all the PDCHs of the first carrier have been used, the next PCCCH_GROUP is mapped to the lowest PDCH carrying PCCCH on the next carrier in the PCCCH description. Hence, the last PCCCH_GROUP gets mapped on the highest-numbered PDCH carrying PCCCH on the last carrier defined in the PCCCH description.

The MS identifies the PCCCH_GROUP using the formula

$$\text{PCCCH\_GROUP} = ((\text{IMSI mod } 1000) \text{ mod BS\_PCC\_CHANS})$$

Thus, the PCCCH_GROUP of an MS is based on the IMSI module 1000 modulo number of PDCH that carry the PCCCH.

Once the MS has identified the PDCH carrying the PCCCH, the next step is to identify its PAGING_GROUP on this PDCH. The occurrence of the paging blocks belonging to the MS in the DRX mode is controlled by the parameter SPLIT_PG_CYCLE. This parameter is set by the MS and sent to the network during the GPRS attach procedure. The SPLIT_PG_CYCLE may be applicable on the CCCH and PCCCH, but this is optional for both the MS and the network. The MS indicates support for the SPLIT_PG_CYCLE on the CCCH in the DRX parameters sent in the attach request, whereas the network indicates the support in the SI 13 or the PSI 1 message.

The PAGING_GROUP is defined as follows:

$$\text{PAGING\_GROUP}(0 \ldots M-1) = \left\{ \begin{array}{l} (\text{IMSI mod } 1000) \text{ div BS\_PCC\_CHANS} \\ + (\text{IMSI mod } 1000) \\ + \text{Max}\,((m * M) \text{ div SPLIT\_PG\_CYCLE}, m) \end{array} \right\} \text{mod } M$$

for $m = (0 \ldots \text{Min}(M, \text{SPLIT\_PG\_CYCLE}) - 1)$,

where IMSI refers to the International Mobile Subscriber Identity of the MS.

In the above formula, the parameter M indicates the number of paging blocks available on one PCCCH. In the case of PCCCH channels, the number of blocks is calculated across 64 multiframes. Hence the value M can be given as (12 – BS_PAG_BLKS_RES – BS_PBCCH_BLKS) × 64. This is derived as the number of blocks in one multiframe (12) minus the blocks reserved for PAGCH and PBCCH, which is then multiplied by 64 (analogous to the parameter BS_PA_MFRMS of CCCH).

4.5.4 Non-DRX Mode

When an MS has left the packet transfer mode or the dual transfer mode and enters into the packet idle mode, the MS also enters the non-DRX mode. Additionally, an MS operating in the NC 2 mode of operation enters the NC 2 non-DRX mode after sending the measurement report. This mode of operation is explained in the next section.

The time for which the MS is in the non-DRX mode in the first case is determined by the minimum value of two timer values. The first value is the NON-DRX_TIMER that is set by the MS during the GPRS attach procedure.

The second parameter is the DRX_TIMER_MAX that is broadcast by the network in the PSI1 and the PSI13 messages. In the second case, however, the duration for the non-DRX mode is determined by the NC_NON_DRX_PERIOD broadcast by the network in the PSI 5 message. During this mode, the MS listens to all the paging blocks in the downlink direction. In other words, there are no constraints on the paging subchannels being monitored by the MS. Hence, the initiation of the downlink data transfer from the network is faster in this case since the network does not need to wait for a specific paging subchannel before paging the MS.

4.6 Cell Reselection

When the MS is camped on a cell or a data transfer is ongoing on a particular cell, the MS keeps monitoring the neighboring cells in addition to the serving cell in order to determine if the neighboring cell is in a better position to cater to the MS. The MS keeps monitoring the signals from the serving cell and the neighbors, and analyses them to find the cell that provides the signal with the highest signal strength. In case the cell with highest signal strength is not the serving cell, then the MS selects the neighbor and makes it the serving cell. This process is known as the "cell reselection". The neighbor cell which provides the higher signal strength can also be referred as the target cell.

When the decision about the cell reselection is taken, the MS continues its operation in the serving cell and attempts to acquire the SI messages being broadcast on the target cell. The behavior of the MS depends on whether the serving cell supports the Cell Change Notification (CCN). In case the serving cell does not support CCN, the MS aborts the procedures in the serving cell and shifts to the target cell as soon as it acquires the broadcast information. In case the CCN is supported in the serving cell, the MS follows the CCN procedures. However, if the cell reselection decision has been taken, then an MS cannot continue its operation in the serving cell for more than 5 seconds whether the serving cell supports the CCN or not.

The support of GPRS is indicated in the SI. If the MS, however, had received a PBCCH description of the target cell in its previous serving cell, then it assumes that the GPRS service is available in the target cell.

Some criteria have been defined for the cell reselection. In order to perform the cell reselection, the MS needs some special parameters related to the serving cell and the neighboring cells. These parameters are broadcast by the network on the BCCH or the PBCCH, whichever is defined on the cell. In case the BCCH is defined, the information is broadcast on the SI 3 or SI 4 message. The neighbor cell parameters are obtained from the broadcast channel of the neighboring cell. In case the PBCCH is defined on the serving cell, the entire information about the present cell and its neighbors can be obtained from the PSI 3 message.

4.6.1 Modes of Cell Reselection

The MS continuously needs to monitor the serving cell and its neighbors in order to determine if cell selection should be performed or not. This decision may be taken by the MS or by the network on the basis of the measurements made by the MS. There are different modes of operation for the cell reselection that are as follows:

- **Normal MS control, NC0.** In this mode of operation, the MS performs autonomous cell reselection. The measurements are made by the MS, but these are not sent to the network.
- **MS control with measurement reports, NC1.** In this mode of operation, the MS performs the autonomous cell reselection. However, the measurement reports are sent by the MS to the network.
- **Network control, NC2.** In this mode of operation, the network controls the cell selection. The MS sends the measurement reports to the network. In the NC2 mode, there are some special cases when the MS can perform the cell reselection. These cases occur during downlink signaling failure or the random access failure.

The mode that would be used by an MS is determined by the NETWORK_CONTROL_ORDER parameter that is transmitted on the broadcast channel. The network uses the SI 13 and the SI 2 quat on the BCCH and the PSI 5 message on the PBCCH to broadcast this information. The possible values of this parameter and their significance are listed in Table 4.2.

Table 4.2

Significance of the Possible Values of NETWORK_CONTROL_ORDER

Value	*Abbreviation*	*Significance*
00	NC0	Autonomous cell reselection by the MS. No measurement reporting to the network.
01	NC1	Autonomous cell reselection by the MS. Measurement reporting to the network.
10	NC2	Network controlled cell reselection. Measurement reporting to the network.
11	Reset	A reset is sent from the network to the MS to force it to return to the broadcast parameters. The measurement-reporting parameters are sent on the broadcast channel. However, parameters can be sent to a specific MS on the PACCH or the PCCCH. On the reset, the MS returns to the parameters that were broadcast on the BCCH or the PBCCH.

In addition to the network control mode, the GPRS Mobility Management (GMM) state of the MS also determines the behavior of the MS. An MS that is in GMM STANDBY state always uses the autonomous cell reselection procedure and does not send the measurement reports to the network. The MS behavior for various modes and GMM state is listed in Table 4.3.

4.6.2 Criteria for Cell-Reselection

As mentioned in the introduction of this section, there are some criteria that are used for the cell reselection. This information is broadcast by the network on the BCCH or the PBCCH, whichever is present in the cell. The main aim of these criteria is to make sure that the MS camps on the cell that has the maximum probability of successful communication between the MS and the network in the uplink and downlink directions. There are two criteria that can be used:

- **(C1, C2) criteria.** This criterion is used in case the PBCCH is not defined in the serving cell.
- **(C′1, C31, C32) criteria.** This criterion is used in case the PBCCH is defined in the serving cell. It has been introduced for GPRS.

Table 4.3 MS Behavior Depending on the GMM State

NCO Value	*GMM State*	*Cell-Reselection Mode*	*Measurement Reporting to Network*
NC0	STANDBY	Autonomous cell reselection	No
	GMM READY	Autonomous cell reselection	No
NC1	STANDBY	Autonomous cell reselection	No
	GMM READY	Autonomous cell reselection	Yes
NC2	STANDBY	Autonomous cell reselection	No
	GMM READY	Network-controlled reselection	Yes

Even when the PBCCH is not defined in the serving cell, the (C'1, C31, C32) criteria can be used by the MS. This is possible if all the parameters to be used for the cell reselection are sent to the MS in an RLC/MAC control message.

4.6.2.1 C1 Criterion

C 1 is the path-loss criterion parameter used for the cell selection as well as for the cell reselection in GSM. In GPRS, this criterion is used for cell reselection when PBCCH is not defined in the cell. This parameter is defined as

$$C1 = (A - \text{Max}\,[B, 0])$$

where,

$$A = \text{RLA\_C} - \text{RXLEV\_ACCESS\_MIN}$$

$$B = \text{MS\_TXPWR\_MAX\_CCH} - \text{P}$$

Here, the RLA_C is the average value of the received signal level at the MS. The RXLEV_ACCESS_MIN is the minimum signal level that the MS should receive in order to access the system. MS_TXPWR_MAX_CCH is the maximum transmit power that the MS can use in order to access the system and the parameter P is the maximum RF output power of the MS.

The value $A > 0$ indicates that the received signal level is higher than the minimum signal strength that is needed at the MS to access the system. In other words, the value $A > 0$ indicates that the MS is in the cell coverage area and the downlink is good. On the other hand, a value $A < 0$ indicates that the MS is outside the cell coverage area and the reception is bad, hence indicating that the MS cannot access the system.

The value B similarly indicates whether the transmission capabilities of the MS are sufficient to access the system. Value $B < 0$ indicates that the transmission capabilities are sufficient.

The path-loss criterion is satisfied if C1 > 0. In other words, an MS can select a cell if $A > 0$ and $B < 0$. Also, if $B > 0$ and $A - B > 0$, the MS transmission capabilities are compensated by the reception margin and the cell can be selected. However, if $B > 0$ and $A - B < 0$, the transmission capabilities are not compensated and the cell is not selected.

4.6.2.2 C′1 Criterion

The path-loss criterion C′1 is used for the cell reselection when PBCCH is defined in the cell. It is the same as C1 except that the parameters used are specific for GPRS. The parameter is defined as follows:

$$C'1 = (A - \text{Max}\,[B, 0])$$

where,

A = RLA_C – GPRS_RXLEV_ACCESS_MIN

B = GPRS_MS_TXPWR_MAX_CCH – P

This criterion is satisfied if C′1 > 0. Thus, one can observe that C1 and C′1 are same except that the latter uses the GPRS-specific parameters for cell reselection.

4.6.2.3 C2 Criterion

This criterion is used for the cell reselection procedure in GSM. In GPRS, this criterion is used for cell reselection when PBCCH is not defined in the cell. The parameter is defined as follows:

C2 = C1 + CELL_RESELECT_OFFSET – TEMPORARY OFFSET * H (PENALTY_TIME - T)

where,

for the nonserving cell,

$H(x) = 0$ for $x < 0$

1 for $x \geq 0$

For the serving cell,

$$H(x) = 0$$

In other words, if the penalty time is higher than T,

C2 = C1 + CELL_RESELECT_OFFSET – TEMPORARY OFFSET

and if the penalty time is less than T,

C2 = C1 + CELL_RESELECT_OFFSET.

Here, the timer T is applied to every cell that is present in the list of the strongest carriers. This timer is started as soon as the MS places a cell in the strongest carriers except when the previous serving cell is placed on the list of strongest carriers at cell reselection. In this case, T shall be set to the value of PENALTY_TIME (i.e., expired). Temporary offset is used to penalize the cell during the penalty time. In other words, a cell newly entering the list of strongest cells pays a penalty in terms of a TEMPORARY OFFSET till the penalty time expires. The CELL_RESELECT_OFFSET is applied to the C2 criterion of the serving cell and is used to prioritize one cell over the other, especially in the case of multiband operations.

4.6.2.4 C31 Criterion

This is a new cell reselection criterion that is used in GPRS in addition to the existing GSM cell reselection criterion. This parameter is a signal strength criterion, which is used to decide whether prioritized cell reselection can be used or not. The cell with the highest priority is used from all the cells that satisfy the C31 criterion. If more than one cell has the highest priority, then the cell with the highest C32 value is used for the cell reselection. It may happen that no cell satisfies the C31 criterion, in which case the cell with the highest C32 value is selected. In order to belong to the highest priority class, a threshold RXLEV is required, known as the Hierarchical Cell Structures (HCS) threshold or HCS_THR. This parameter has to be satisfied in order to apply the HCS GPRS reselection.

The C31 parameter for the serving cells is defined as follows:

$$\text{C31 (Serv)} = \text{RXLEV (Serv)} - \text{HCS\_THR (Serv)}.$$

If the neighbor cell and the serving cell belong to the same priority class, the C31 parameter for the neighboring cell is defined as follows:

$$\text{C31 (Neigh)} = \text{RXLEV (Neigh)} - \text{HCS\_THR (Neigh)}.$$

In case the neighbor cell and the serving cell belong to different priority class, the C31 depends on the GPRS_PENALTY_TIME, the duration for which the offset GPRS_TEMPORARY_OFFSET is applied, and the timer T, which is started when a cell is included in the list of the strongest carriers by the MS.

If T < GPRS_PENALTY_TIME,

$$\text{C31 (Neigh)} = \text{RXLEV (Neigh)} - \text{HCS\_THR (Neigh)} - \text{GPRS\_TEMPORARY\_OFFSET}$$

If T = GPRS_PENALTY_TIME,

$$\text{C31 (Neigh)} = \text{RXLEV (Neigh)} - \text{HCS\_THR (Neigh)}.$$

Note that the concepts of penalty time and offset are similar to that used in C2 criterion, which penalize a cell till penalty time as soon as it is placed in the list of strongest carriers.

4.6.2.5 C32 Criterion

As discussed earlier, the cell with the highest C32 criterion value is selected in case more than one cell satisfying the C31 criterion has the same priority. Also, if no cell satisfies the C31 criterion, then the cell with the highest C32 value is selected. The C32 criterion applies an individual offset to penalize a cell belonging to the same priority level during a certain time. The GPRS_RESELECT_OFFSET parameter allows a neighbor to be penalized.

The C32 criterion for the serving cell is defined as follows:

$$\text{C32 (Serv)} = \text{C'1 (Serv)}.$$

If the neighbor cell and the serving cell belong to the same priority class, then the C32 criterion for the neighboring cell is defined as follows:

C32 (Neigh) = C1 (Neigh) + GPRS_RESELECT_OFFSET (Neigh) – GPRS_TEMPORARY_OFFSET, if T < GPRS_PENALTY_TIME; and

C32 (Neigh) = C1 (Neigh) + GPRS_RESELECT_OFFSET (Neigh), if T ≥ GPRS_PENALTY_TIME.

If the neighbor cell and the serving cell do not belong to the same priority class, C32 criterion for the neighboring cell is defined as follows:

C32 (Neigh) = C1 (Neigh) + GPRS_RESELECT_OFFSET (Neigh).

The C32 criterion also applies individual hysteresis values to the neighboring cells till 15 seconds after cell reselection in order to avoid unnecessary cell reselections. This value is subtracted from the C32 criterion values of the neighbor cells. Also, in cases where the neighbor cell does not belong to the same routing area, the hysteresis value is applied to the neighbor cell in order to avoid the reselection and the resulting RA update. This is done to avoid the additional signaling. In summary, the hysteresis provides a bias toward current cell and current routing area to avoid cell reselection oscillations till the difference is so significant that warrants a cell reselection.

4.6.3 Autonomous Cell Reselection Process

This procedure is used by the MS to decide on the cell reselection. The decision on the reselection is taken on the basis of the values of the cell reselection criteria parameters that were discussed in the previous section.

4.6.3.1 (C1, C2) Criteria

In case the (C1, C2) criteria are used for the cell reselection, the following should be satisfied for the reselection to occur:

- Path loss on the serving cell should be very high, i.e., C1 < 0.
- Serving cell is barred.
- Random access has failed even after maximum defined attempts for access.
- A neighboring cell with higher C2 criterion is detected. However, the C2 value of the neighbor cell must exceed the C2 value of the serving cell by at least the CELL_RESELECT_HYSTERESIS value in order to avoid any unnecessary cell reselections.
- Downlink signaling failure.

4.6.3.2 (C'1, C31, C32) Criteria

In case the (C'1, C31, C32) criteria are used for the cell reselection, the following should be satisfied for the reselection to occur:

- Path loss on the serving cell should be very high, i.e., C'1 < 0.

- Serving cell is barred.
- Random access has failed even after maximum defined attempts for access.
- A neighboring cell is detected to be better than the serving cell. The best cell is the cell with the highest value of C32 among
  - o those cells that have the highest PRIORITY_CLASS that fulfill the criterion C31 ≥ 0, or
  - o all cells, if no cells fulfill the criterion C31 ≥ 0.
- Downlink signaling failure.

4.6.4 Network-Controlled Cell Change

The decisions about the reselection are made by the network when the mode is NC2 and the MS is in GMM Ready state. The decisions are made on the basis of the measurement reports that are sent by the MS to the network. The RXLEV and the RXQUAL values sent by the MS determine whether the cell reselection should be performed or not.

In case the network decides that the reselection should be performed, the packet cell change order message is sent to the MS. The cell description is included in the message, which provides information about the BCCH frequency and the Base Station Identity Code (BSIC) that is applicable in the target cell. The MS selects the cell as per the included description.

4.6.5 Measurement Reports

The MS makes the measurements and then sends the measurement reports to the network depending on the measurement-reporting parameters broadcast by the network. In the NC1 and the NC2 mode of operation, the reports are sent to the network periodically. The reporting period is decided by

- **NC_REPORTING_PERIOD_I:** The reporting period used when the MS is in the PI mode.
- **NC_REPORTING_PERIOD_T:** The reporting period used when the MS is in the PT mode.

The main information that is sent in the measurement reports are

- The received signal strength, RXLEV of the serving cell.
- The received signal strength, RXLEV of the neighbor cell.
- The average interference level, γ_{ch} of the serving cell during the idle periods.

4.7 Uplink TBF Establishment

When the MS or the network has any data to be transferred, a TBF is initiated. A data transfer in the uplink direction constitutes an uplink TBF, while the data transfer in the downlink direction leads to the establishment of a downlink TBF. This section deals with the procedures for the establishment of the uplink TBF, while the next section discusses the establishment of the downlink TBF. The signaling messages for the TBF establishment are exchanged on the control channels. The packet control channels are used for the signaling if these are defined in the serving cell. In case the packet control channels are absent, the GSM control channels are used. The procedures on both these channels are discussed in the subsequent sections.

As stated, when the MS has data to transfer to the network, an uplink TBF is established. In order to establish the TBF, some control information has to be exchanged between the MS and the network. Note that to transfer this signaling information also, an uplink TBF is needed. Hence, for any data, whether it is user data or signaling messages, an uplink TBF is required. The difference lies in the channels being used for the transmission and the RLC blocks that are used for the transfer of the information. For the control information, the RLC control blocks are used, whereas for the user data, the RLC data blocks are used.

In a nutshell, the establishment of an uplink TBF involves an access request from the MS to the network for the radio resources and the allocation from the network that is used by the MS for the data transfer. The network may use any allocation scheme for the TBF being initiated. These allocation schemes are discussed in the next section. The TBF establishment procedures are discussed in the subsequent sections.

4.7.1 Allocation Schemes for Uplink TBF

On receiving the access request for radio resources from the MS, the network makes an allocation according to the availability of resources. Different allocation schemes may be used. These are as follows:

- Dynamic allocation
- Extended dynamic allocation
- Single block allocation

If the "dynamic allocation" scheme is used, the network allocates the PDCH to the MS and also assigns an USF that is associated with the PDCH. On receiving the allocation, the MS starts monitoring the downlink blocks on the allocated PDCH and decodes the USF value being transmitted in these blocks in the downlink. Whenever the network wants to allocate an uplink block on a PDCH for a TBF, it includes the USF for the TBF on the preceding downlink block. Thus, USF-based allocation operates on different blocks of the same PDCH (B_n and B_{n+1} of TS_x) and not on same blocks of the different PDCH (i.e., not on B_n of TS_x and TS_{x+1}). If the decoded USF value is the same as the allocated USF value for the TBF, the MS transmits data on the next uplink radio block.

It is possible that the network allocates the same PDCH (timeslot) to different MS. In this case, different USF values are allocated to these MS. Consequently, different USF values are transmitted by the network in the downlink to allow the different MS to transfer data in the uplink direction. This mechanism allows the sharing of PDCH between the different MS. However, there is a fundamental limitation of this mechanism in the sense that for each uplink allocation, the MS must decode the downlink blocks to know the USF sent. As there are limits on the number of uplink and downlink slots that an MS can read, some TBF allocations do not comply with the MS multislot classes (i.e., the number of uplink/downlink slots an MS has to operate is higher than its capability to do so). While the USF granularity (which allows MS to monitor one USF and send four uplink blocks instead of one) alleviates the problem to an extent, it does not completely solve the problem.

The "extended dynamic allocation" of resources extends the dynamic allocation to allow higher uplink throughput and solves the problem mentioned earlier. In this case also, the MS may be allocated single or more PDCH(s) depending on the multislot class of the MS. For each of these PDCH(s), a USF value is also assigned to the MS. The MS starts monitoring the lowest allocated PDCH for the assigned USF, and then the next lowered numbered PDCH, and so on. When the MS decodes its assigned USF on a particular PDCH, it transmits the one or four data blocks (depending on the USF granularity) on the uplink radio block of this PDCH and all the allocated higher numbered PDCH(s). This scheme essentially does away with the need to read all USF since one USF allocation implicitly implies allocation for all higher numbered blocks.

The "single block allocation" scheme allocates a single uplink radio block to the MS. The MS may request this allocation type for measurement reporting. Alternatively, the network uses this allocation in order to trigger the two-phase access TBF establishment procedure. These procedures are discussed in detail in the subsequent sections.

4.7.2 Channel Access on RACH/PRACH

The channel access is performed using two modes, "One-phase Access" and "Two-phase Access". These modes essentially control the handshakes required to allocate resources to the MS (first one uses one handshake with two message exchange, while the latter has two handshakes with four message exchanges). While MS may request for a particular access mode, the network decides what mode is actually used. The one-phase access is called so because in this mode, the network allocates uplink resources to the MS as soon as it receives the access burst. In two-phase access mode, the first phase provides the MS with a single uplink block allocation. This block allocation is then used in the second phase to acquire uplink resources for actual data transfer. This technique allows the MS to specify more clearly its requirements and capabilities, which in turn allows the network to allocate resources according to resource availability and loading. The access mode and its relation with channel request are further discussed next.

The TBF establishment is triggered by sending channel request on RACH (if PBCCH is not configured) or sending packet channel request on PRACH (if PBCCH is configured in the network). The packet channel request on PRACH has two formats:

- **8-Bit packet channel request.** In this case, the MS sends the cause of the TBF establishment, the multislot class (optional), and some random bits. This is the only channel request format that is available for the circuit-switched procedures and can also be used in packet-switched procedures.

- **11-Bit packet channel request.** In this case, the MS sends the cause of the TBF establishment, the multislot class, the radio priority, and some random bits. This format is introduced specifically for the packet-switched procedures.

The packet channel request format to be used by the MS is decided by the network. The format to be used by the MS is indicated by the ACC_BURST_TYPE parameter broadcast in the PSI 1 message, where 0 indicates 8-bit format and 1 indicates 11-bit format. This parameter determines the format of the access burst not only on the random access channel but also on the uplink of the PTCCH channel (access burst messages are sent in this channel during the timing advance control procedure).

The message contents of the 8-bit and the 11-bit packet channel request message are discussed next. The contents are also listed in Tables 4.4 and 4.5, respectively.

- **Establishment cause.** This field indicates the reason for the establishment of the TBF. The cause values enable the network to allocate resources accordingly. The allowed establishment causes are explained next:
    - o **One-phase access (or one-phase access in RLC unacknowledged mode).** This indicates the intention of the MS to initiate a one-phase access. However, the network actually decides which mode must be used. In the unacknowledged RLC/MAC mode, the MS cannot propose a one-phase packet access (so the two-phase access must be used in unacknowledged mode or one-phase can be used in acknowledged mode). The RLC/MAC mode of operation is governed by the Quality of Service (QoS) profile of the MS, which is stored at subscription profile at the Home Location Register (HLR).
    - o **Two-phase access.** This indicates the intention of the MS to initiate a two-phase access. Ultimately, the network decides which mode is used. In the unacknowledged RLC/MAC mode, the two-phase access is used.

TABLE 4.4

Contents of the 8-Bit Packet Channel Request Message

Establishment Cause	*Prefix*	*Multislot Class*	*Priority*	*Random Bits*
One-phase access request	1	5 bits	N	2
Two-phase access request	01000	N	N	3
Page response	01001	N	N	3
Cell update	01010	N	N	3
MM procedure	01011	N	N	3
Single-block without TBF establishment	01100	N	N	3
One-phase access request in RLC unack mode	011010	N	N	2
Dedicated channel request	011011	N	N	2
Emergency call	011100	N	N	2

TABLE 4.5

Contents of the 11-Bit Packet Channel Request Message

Establishment Cause	*Prefix*	*Multislot Class*	*Priority*	*Random Bits*
One-phase access request	0	5 bits	2 bits	3
Two-phase access request	110000	N	2 bits	3
Page response	110001	N	N	5
Cell update	110010	N	N	5
MM procedure	110011	N	N	5
Single block without TBF establishment	110100	N	N	5
One-phase access request in RLC unack mode	110101	N	N	5
Dedicated channel request	110110	N	N	5
Emergency call	110111	N	N	5

- o **Page response.** This establishment cause is used by the MS in response to a paging. The TBF establishment for this procedure uses one-phase access mode.
- o **Cell update.** This establishment cause is used by the MS to trigger a cell update procedure. The TBF establishment for this procedure uses one-phase access mode.
- o **Mobility Management (MM) procedure.** This establishment cause is used for GPRS MM procedures. The TBF establishment for the MM procedures uses one-phase access mode.
- o **Short access.** As the name suggests, the short access is used when the MS wants to transfer an amount of data that is less than 8 RLC/MAC blocks.
- o **Single block without TBF establishment.** This mechanism is used to send a measurement report to the network. This procedure allows the MS to merely send a block without having to incur the overhead of establishing a TBF.
- o **Emergency call.** This establishment cause is used for emergency services.

It is important to note that the establishment cause listed here applies to 8-bit and 11-bit packet channel request on PRACH. The establishment causes of channel request are different and do not include all the causes listed here. The channel request on RACH for packet data includes only two causes, namely one-phase access and single-block access. The single-block allocation is then used for initiating two-phase access.

- **Priority.** This field indicates the radio priority of the TBF being established. This field is used in the 11-bit packet channel request only. The MS allocates higher priority to the TBF that is being established to transfer control information.
- **Multislot class.** This 5-bit field indicates the multislot class of the MS allowing the network to allocate resources to the MS according to its capabilities.

The contents of the channel request and the frame number on which the channel request is received are known as the packet request reference. This reference acts as the identity of the MS in the downlink message sent in response to the channel request.

4.7.3 Procedure on the PCCCH

When the PBCCH is defined in the serving cell, the MS uses the PCCCH for the TBF establishment. The TBF is initiated by a packet channel request message sent by the MS on the PRACH. The location of the PRACH in the 52 mulitframe is indicated to the MS by the network parameters that are broadcast in the PSI messages. The channel request contents were discussed in the previous section. This section continues the discussion further and elaborates upon the signaling exchanges for one-phase and two-phase access.

4.7.3.1 One-phase Access on PCCCH

Upon receiving the channel request from the MS, the network allocates the radio resources. The decision for the resource allocation is governed by the load supervision of the network. The other deciding factors are the multislot class and the radio priority of the MS. Accordingly, different allocation schemes may be used. The allotment is sent to the MS in the allocation message, the packet uplink assignment message on the Packet Access Grant Channel (PAGCH). The packet request reference is used to identify the MS in this message. As part of the assignment, the following information is provided to the MS:

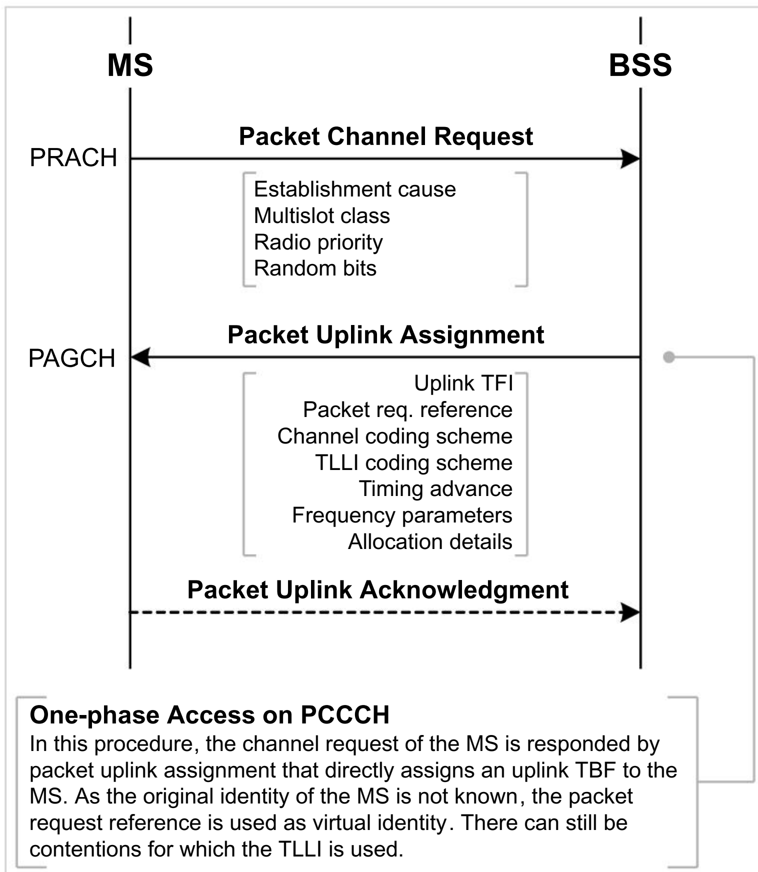
- **Uplink TFI.** This identifies the uplink TBF. Note that the TFI is independent of the number of PDCH allocated to the MS. Thus, this TFI is also called the global TFI.
- **Packet request reference.** This carries the contents of the packet channel request message along with the frame number on which the request was received. This parameter acts as a virtual identifier of the sender of the request.
- **Channel coding.** This refers to the coding scheme (i.e., CS 1, CS 2, CS 3, or CS 4) to be used for the data transfer.
- **TLLI block coding scheme.** This refers to the coding scheme to be used when the Temporary Logical Link Identifier (TLLI) is encoded in the data blocks. The coding scheme is either CS 1 or the same as that commanded for channel coding.
- **Timing Advance (TA) information.** This includes the TA value for the MS in case the initial TA is computed and is being sent back to the MS. If the GPRS continuous TA procedure has to be used, the MS is provided with the details of the PTCCH to be used (i.e., the Timing Advance Index (TAI) and the timeslot of the PTCCH on which the continuous TA procedure is executed).
- **Frequency parameters.** This includes the radio channel frequencies to be used by the MS for the TBF. In case frequency hopping is used, the details of the MA are also provided to the MS.
- **TBF allocation details.** This includes details of the timeslot(s) being allocated and the USF value(s) to be used in case of dynamic and extended dynamic allocation. In addition, the power control parameters to be applied and the TBF starting time are also provided to the MS.

The MS uses this uplink allocation for the data transfer. Optionally, the network can poll the MS requesting it to send the packet control Acknowledgment message. This polling procedure is desirable for the network to be sure that the MS indeed has got the uplink assignment message and will be using the allocated uplink resources. The polling is achieved by using the S/P bit and the Relative Reserved Block Period (RRBP) field in the downlink RLC/MAC control block.

The TBF establishment discussed above in which the MS sends only one access message is known as one-phase access TBF establishment. The message exchange for one-phase access on PCCCH is depicted in Fig. 4.8.

FIGURE 4.8

One-phase Access Procedure for Uplink TBF Establishment on PCCCH



4.7.3.2 Two-phase Access on PCCCH

The two-phase access (irrespective of whether requested by MS or not) is indicated by the network to the MS by using special allocation called the "Single Block Allocation" in the packet uplink assignment sent in response to the packet channel request. In this allocation, only a single block is allocated to the MS that allows it to send the information about its capabilities to the network. The packet uplink assignment for single block has fewer parameters as compared to the normal dynamic/extended allocation as the allocation is only for one block. Further, note that the TFI is not allocated to the MS till this time.

The MS responds with a packet resource request message on the allocated radio block. This message provides the following details to the network:

- **Access type.** This indicates the reason for requesting the access and can be two-phase access request, page response, cell update, or MM procedure.
- **TLLI.** The MS encodes its identity, a TLLI, in the packet resource request message. The TLLI is used for contention resolution purposes as discussed in subsequent sections.
- **Channel request description.** This describes some of the characteristics of the channel requested including the number of RLC octets to be sent, the priority, and the peak throughput.
- **Radio access capability.** Another information element that provides MS capability information to it is radio access capability. This information element includes parameters related to GPRS multislot class, support of extended dynamic allocation, supported bands, and radio frequency power capabilities among other things.

The network uses the packet resource request to allocate an uplink TBF as per the needs of the MS. The information provided by the MS are used in making an uplink allocation. The multislot information allows the network to make maximum utilization of the capabilities of the MS. Other parameters like throughput and priority are also used to provide QoS-based resource allocation. The TBF allocated is then communicated on the downlink PACCH carrying parameters that are similar to the uplink TBF establishment in one-phase access. However, some of the parameters like request reference and TLLI channel coding command are not relevant and not sent in the second phase. A key addition, however, is the TLLI, which is same as that sent by the MS in the packet resource request. The TLLI sent in downlink allows the MS to match with the TLLI sent in the uplink, thereby helping in breaking contention.

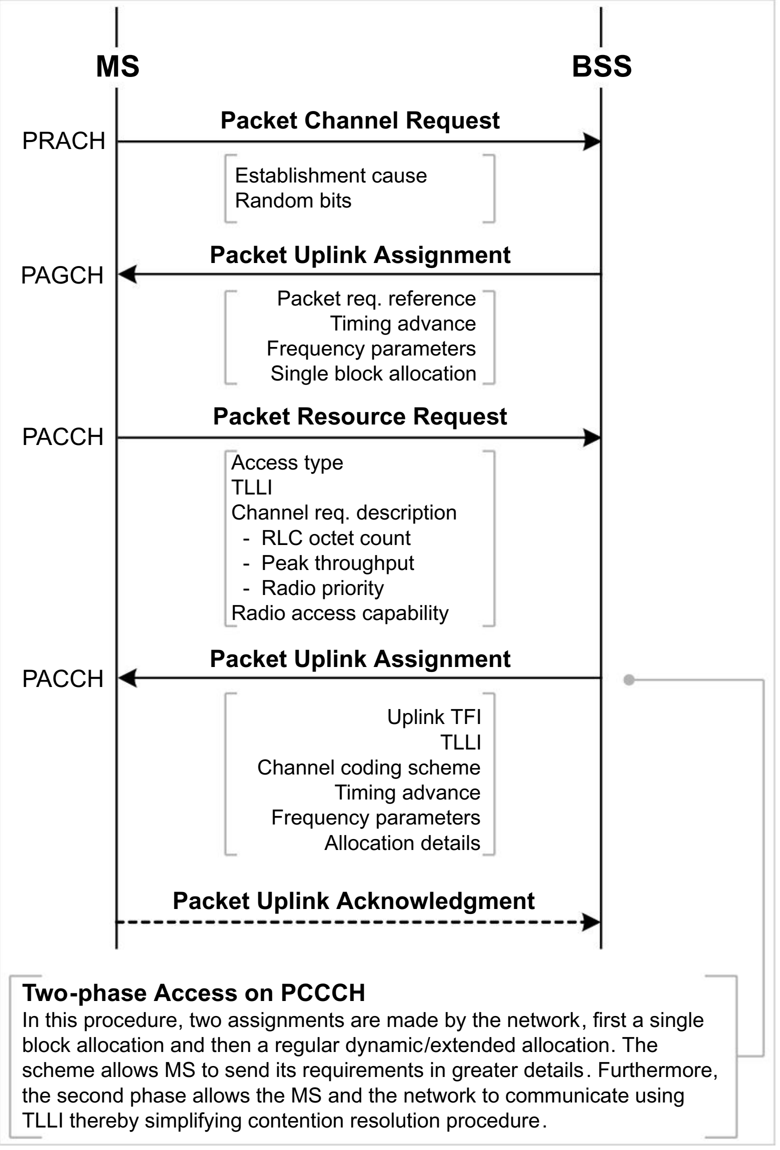
Like in one-phase access, the network can poll the MS by using the S/P bit and the RRBP field in the downlink RLC/MAC control block. This allows the network to receive the packet control Acknowledgment message, thereby being sure that the MS has indeed got the uplink assignment message and will be using the allocated uplink resources.

This type of access discussed earlier in which the MS needs to send two messages in order to establish the TBF is known as the “Two-phase TBF establishment”, as depicted in Fig. 4.9.

As discussed here, the two-phase TBF establishment is used by the MS when it is triggered by the network. In addition, it is mandatory for the MS to use this procedure to initiate a TBF in the RLC Unack mode. This is required in order to solve the contention resolution problem. This problem is discussed in more detail in Section 4.7.5. In this case, the establishment cause encoded by the MS in the access request would indicate a two-phase access request.

FIGURE 4.9

Two-phase Access Procedure for Uplink TBF Establishment



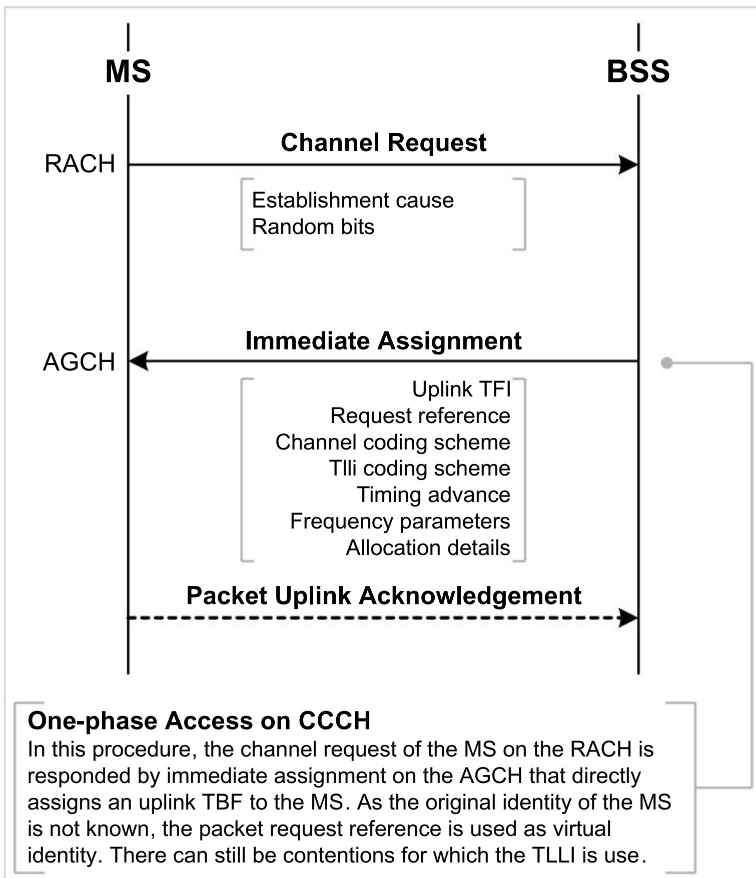
4.7.4 Procedure on the CCCH

In case the serving cell does not have any packet common control channel, the GSM CCCH is used for the TBF establishment procedure. Like PCCCH, the CCCH also offers two modes of access, namely one-phase and two-phase access.

In one-phase access, the MS sends the channel request message on the RACH to request for an uplink TBF establishment. The establishment cause in the message is sent as "packet access" indicating that resources are required for a packet data transfer. In response, the network allocates resources in the immediate assignment message sent on the AGCH. This message contains the uplink resource allocation for the MS. The parameters are similar to one carried in packet uplink assignment. Once the allocation has been received by the MS, the allocated PDCH are used for the data transfer. The network can optionally poll the MS to send an acknowledgment. The message transfer is as depicted in Fig. 4.10.

FIGURE 4.10

One-phase Access Procedure for Uplink TBF Establishment on CCCH



The two-phase access on CCCH is similar to the scenario described on PCCCH. The key difference between the two is that in CCCH, the first phase is on RACH/AGCH versus PRACH/PAGCH on PCCCH. Thereafter, in both the cases, PACCH is used to exchange packet resource request and packet uplink assignment message for second phase of message exchange.

4.7.5 Contention Resolution

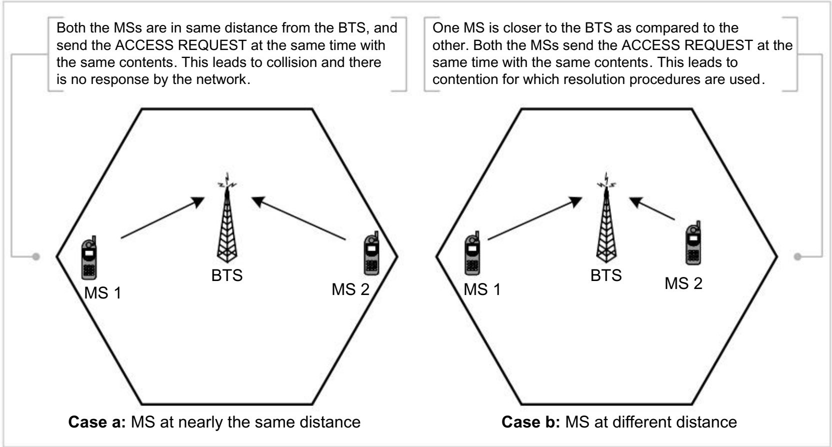
The channel request message that the MS sends for the establishment of the uplink TBF contains the cause value for the TBF, some random bits and optionally the multislot class of the MS, and the radio priority of the TBF being initiated depending on the message format being used. These parameters are not unique and can be same for multiple MS. Consider the case when multiple MSs belonging to the same multislot class initiate the TBF with the same establishment cause and the radio priority on the same occurrence of the RACH or the PRACH. In this case, the channel request from these MSs can be differentiated on the basis of the random bits that are encoded in the message. However, it might happen that the random bits generated by the MS are also the same. In this case the contents of the access message sent by the MS would be exactly the same, which would lead to a contention. This necessitates a contention resolution procedure.

Consider the scenario depicted in Fig. 4.11(a). If the mobiles are at nearly the same distance, the two request messages would be received by the Base Transceiver Station (BTS) at comparable signal level. A collision would, hence, be detected on the RACH or the PRACH and there would be no response to the channel request. The two MSs would then make another access attempt on another instance of the access channel. The probability of the same random bits being generated again is low and hence this is not a problem.

The problem arises when one of the MSs is closer to the BTS and the other is far away, as depicted in Fig. 4.11(b). In this case, the request message from the MS farther away would be received at a very low level as compared to the message from the MS closer to the BTS. The BTS would decode the request from the nearer MS and would not detect any collision. The network would respond to the access request by allocating resources in the packet uplink assignment message. Both the MSs would read this message in the downlink, identify their packet request reference, and assume that the message is intended for them. Both the MSs would then transmit the data blocks on the same PDCH. This situation leads to the concept of

FIGURE 4.11

Access Request from MS at Same or Different Distance from the BTS



contention resolution. First the one-phase access procedure and then the two-phase access procedure have been discussed here.

4.7.5.1 Contention Resolution During One-phase Access

After an MS has received the uplink allocation message, it starts transmitting the data blocks on the assigned PDCH. If the TBF establishment takes place using the one-phase access procedure, the MS encodes its TLLI in the data blocks. The TLLI uniquely identifies the MS in the routing area. When the MS sends the first RLC/MAC data block with the encoded TLLI, it starts a counter, N3104, and sets it to 1. With every block that is transmitted, N3104 is incremented. The maximum allowed value of this counter is calculated by the MS from the parameters broadcasted in PSI 1.

On receiving the first RLC/MAC data block with TLLI, the network sends a packet uplink ack/nack message as an acknowledgment of the data block. This message contains the TLLI value (specifically referred as contention resolution TLLI) received in the data block and the TFI value allocated to the TBF in the assignment message. The TLLI value is used to identify the MS and it enables contention resolution.

The contention resolution is successfully completed on the network side when it receives a data block with the encoded TLLI value. On the MS side, the contention resolution is completed on receiving the packet uplink ack/nack containing the same TLLI as sent in the uplink data blocks. After the contention resolution, the MS does not encode the TLLI value in the subsequent data blocks being transmitted. The procedure is depicted in Fig. 4.12.

However, if the MS receives a packet uplink ack/nack containing the TFI allocated to its uplink TBF but TLLI value that does not belong to this MS, the contention resolution fails. In this case, the MS stops transmitting on the TBF and re-initiates the packet access procedure. The contention resolution on the MS would also fail if the counter N3104 reaches its maximum value and the packet uplink ack/nack containing the TLLI of the MS is not received.

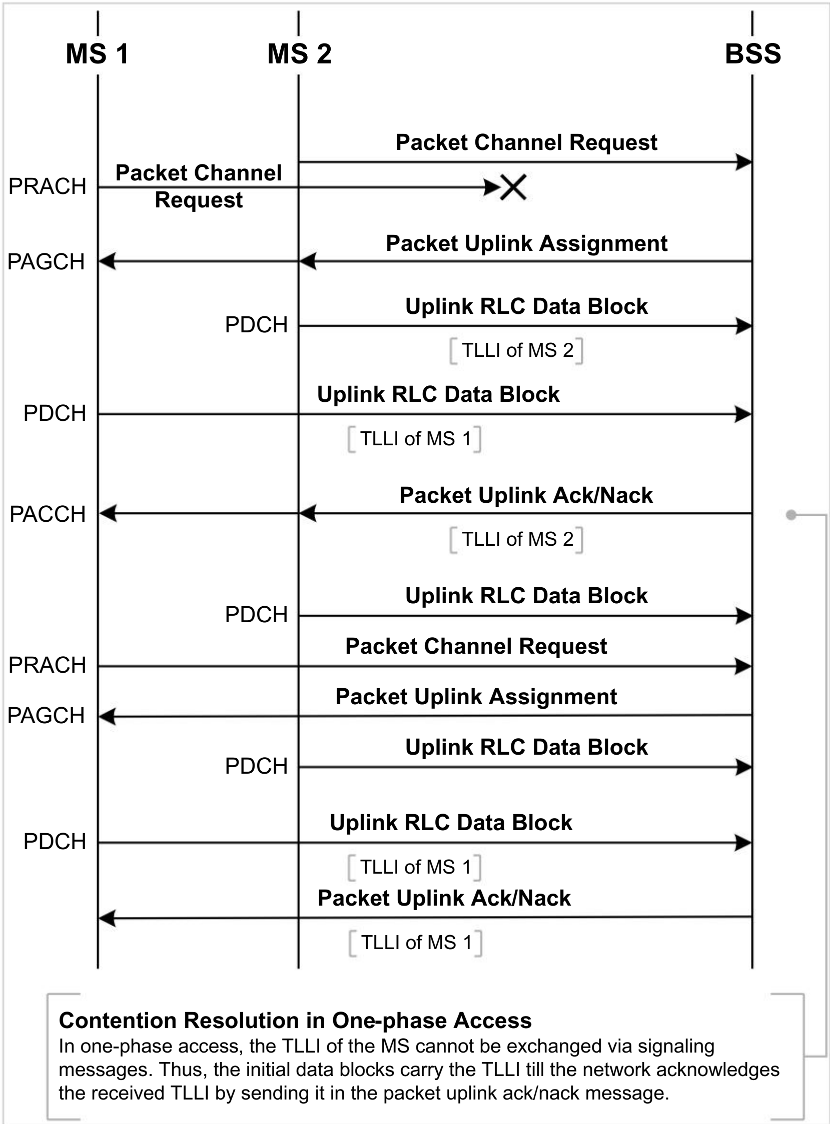
4.7.5.2 Contention Resolution During Two-phase Access

In case of two-phase access procedure, the contention resolution on the network side is completed before the data transfer starts. This is because the unique identification of the MS, the TLLI is exchanged between the MS and the network in the packet resource request message. On the MS side, the contention resolution is completed when the MS receives the packet uplink assignment sent in response to the packet resource request. These assignment messages contain the TLLI of the MS in order to identify it.

If the packet uplink assignment message sent by the network in response to the packet resource request contains a TLLI value that does not belong to the MS, the contention resolution procedure fails at the MS side and the MS initiates the data transfer by sending the access request again.

FIGURE 4.12

Contention Resolution in the One-phase Access Procedure



Consider a case when the TBF has to be initiated in the RLC unacknowledged mode. In this case if the TBF is established using the one-phase access procedure, the contention resolution would not be possible. This is because there would be no acknowledgment from the network. In order to allow the contention resolution in such TBF, it is mandatory to establish the TBF using two-phase access procedure.

4.8 Downlink TBF Establishment

When the upper layers on the network side have data to be transferred to an MS, these layers request for the establishment of a downlink TBF and specify certain parameters associated with the data transfer. These parameters include

- Radio priority
- QoS parameters, the QoS for the data transfer that includes the requested RLC mode
- DRX parameters
- Multislot class
- MS radio access capability

On receiving such a request from higher layers, the procedure for the downlink TBF establishment is initiated by the network. The MS for which the data is intended might be in PI state or in PT mode. If the MS is in the PI state, the network uses the common control channels to establish the downlink TBF. On the other hand, if the MS is in the PT mode, i.e., it is already involved in a data transfer, the network uses the associated control channel, PACCH, for initiating the downlink data transfer. The following sections discuss the procedures for the downlink TBF establishment. The procedures for the MS in PI mode are discussed first on the packet channels and then on the GSM channels. The procedures for the MS in the PT mode are explained in the subsequent sections of the chapter that discuss the concurrent TBF establishment.

4.8.1 Procedure on the PCCCH

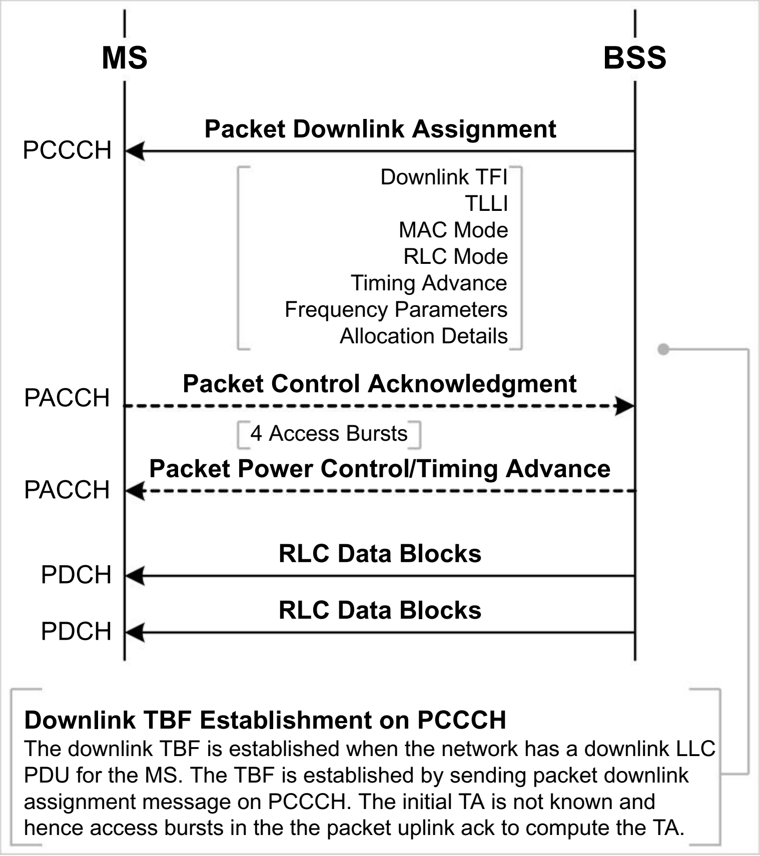
The downlink TBF establishment can be initiated on the packet channels as well as the GSM channels. If the packet channels are present, then these are used for the establishment procedure. Depending on the parameters provided by the upper network layers for the data transfer, the network assigns the radio resources to the MS. The resources allocated depend entirely on the network. This allocation information is provided to the MS in the packet downlink assignment message. This message is transmitted on the PCCCH block depending on the DRX mode of the MS. In case the MS is working in the DRX mode, the network calculates the paging group of the MS and the allocation message is sent on the PCCCH corresponding to the paging group of the MS. If, however, the MS is working in the non-DRX mode, there is no restriction on the downlink PCCCH on which the message can be sent. The message can be sent on any PCCCH to be used for paging. The signaling procedure is depicted in Fig. 4.13.

The downlink assignment message contains the information regarding the downlink TBF being established. If the TBF is being established for an MS in the PI mode, the TLLI is encoded in the assignment message and used for the identification of the MS. The information provided to the MS in the assignment message are:

- **Downlink TFI.** This identifies the downlink TBF.
- **TLLI.** This identifies the MS for which the downlink TBF is being established.

FIGURE 4.13

Downlink TBF Establishment Procedure on the PCCCH



- **MAC mode.** This indicates the MAC allocation mode (i.e., dynamic allocation or extended dynamic allocation).
- **RLC mode.** This indicates the RLC mode of operation, which can be acknowledged or unacknowledged mode.
- **TAI.** The initial TA value in downlink TBF establishment is not known as this is triggered by the network without receiving any access burst from the MS. Thus, if the GPRS continuous timing advance procedure is used, the MS is provided with the details of the PTCCH (i.e., the TAI and the timeslot of the PTCCH on which the continuous timing advance procedure is executed). The other alternative is to poll the MS by requesting it to send four access bursts in the uplink for the packet control acknowledgment. The network then responds with the packet power control/Timing Advance message that carries the TA. This allows the MS to get the first TA instead of waiting for up to 1.92 seconds for continuous TA update.

- **Frequency parameter.** This includes the radio channel frequencies to be used by the MS for the TBF. In case frequency hopping is used, the details of the MA are also provided to the MS.
- **allocation details.** This includes details of the timeslot(s) being allocated. In addition, the power control parameters to be applied and the TBF starting time are also provided to the MS.

After the downlink assignment, the initial TA can be computed using the packet uplink acknowledgment when used in the form of access burst. Thereafter, the MS waits till the TBF starting time, after which the MS monitors the assigned timeslots for the downlink TFI value. The data blocks containing the TFI value assigned to its downlink TBF constitute the downlink TBF of the MS. The MS acknowledges the downlink data blocks received using a packet downlink ack/nack message.

4.8.2 Procedure on the CCCH

When the packet control channels are not defined in the serving cell, the GSM common control channels are used for the downlink TBF establishment procedure. The CCCH is used for the signaling. The immediate assignment message is transmitted on the CCCH that is used for paging. Depending on the DRX mode of the MS, the CCCH is used. The assignment message provides information such as the timeslot allocated to the downlink TBF, the downlink TFI, and other parameters to the MS. The signaling procedure is as depicted in Fig. 4.14.

The packet downlink assignment message is carried in the immediate assignment message. This message provides all the allocation information to the MS. After the downlink assignment, the initial TA can be computed using the packet control acknowledgment when used in the form of access burst. Thereafter, the MS starts monitoring the assigned timeslots after the TBF starting time. The data blocks transmitted on the assigned timeslots contain the allocated downlink TFI value. This TFI value allows the MS to identify the blocks being transmitted for it.

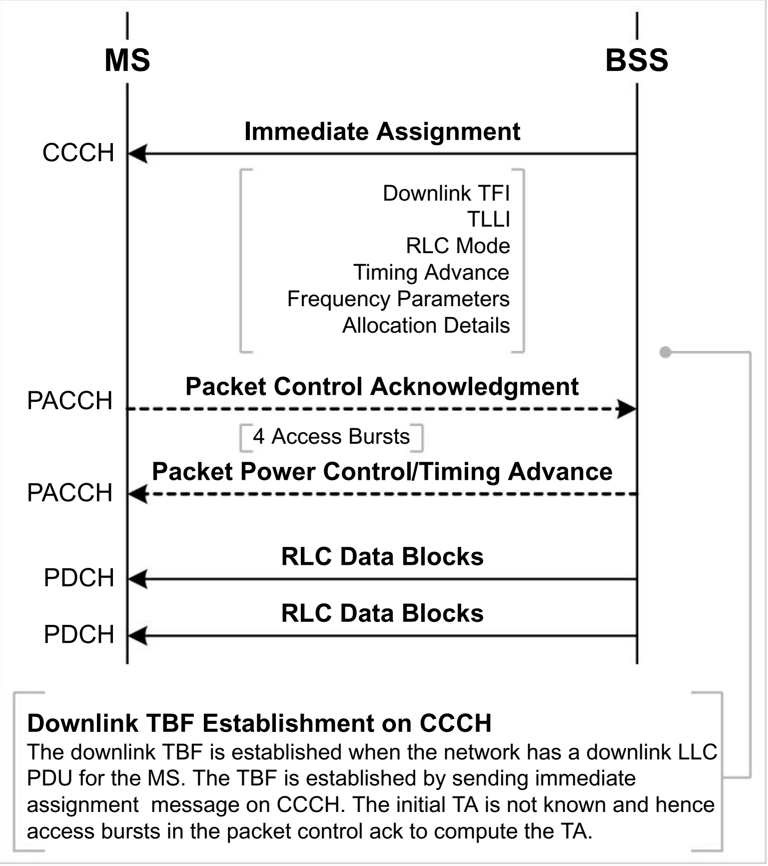
4.9 Concurrent TBF Establishment

It is possible for the MS to support more than one TBF at the same time in either direction. In other words, it is possible for an MS to be involved in more than one TBF in the same direction (uplink or downlink). Also, the MS may be involved in an uplink and a downlink TBF at the same time. Such TBFs are called concurrent TBFs.

It is possible for the MS to initiate an uplink TBF while it is involved in a downlink TBF. Similarly, a downlink TBF can be established for an MS that is involved in an uplink TBF. These procedures are discussed in the next sections.

Figure 4.14

Downlink TBF Establishment Procedure on the CCCH



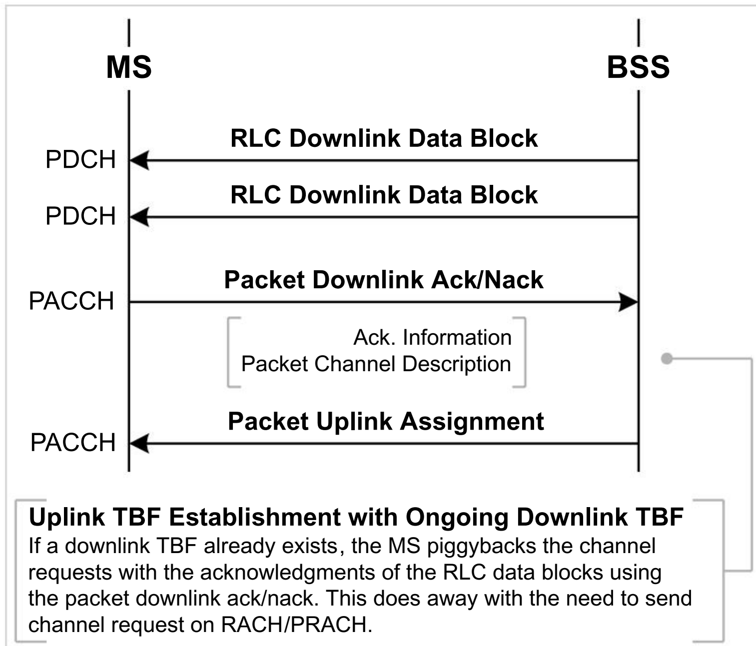
4.9.1 Uplink TBF Establishment for MS in Downlink TBF

An MS that is a part of a downlink TBF may have some data to transfer. In this case, the upper layers of the MS would request for the establishment of an uplink TBF. In the normal establishment procedure, the MS sends a channel request to the network on the access channel. However, for an MS that is a part of a downlink TBF, this is not required. In this case, the MS initiates the uplink TBF by sending a packet channel description in the packet downlink ack/nack message. On receiving such a request from the MS, the network assigns the radio resources to the MS on one or more PDCH in the packet uplink assignment or the packet timeslot reconfigure message that are transmitted on the PACCH. In these assignment messages, the MS is identified by the downlink TFI associated with the ongoing downlink TBF of the MS. It is to be noted here that in case the uplink TBF had been initiated by an MS in PI mode, the packet request reference and the TLLI (in case of two-phase access) would have been encoded in the packet uplink assignment messages.

The procedure for the establishment of the uplink TBF is depicted in Fig. 4.15. Once the MS receives the allocation for the uplink TBF, it keeps monitoring the allocated PDCH for the allocated USF values. Additionally, it also keeps receiving the data blocks that constitute the downlink TBF on the PDCH allocated for the downlink TBF. The concept should be clear with the following example.

Figure 4.15

Establishment of Uplink TBF for an MS Involved in Downlink TBF

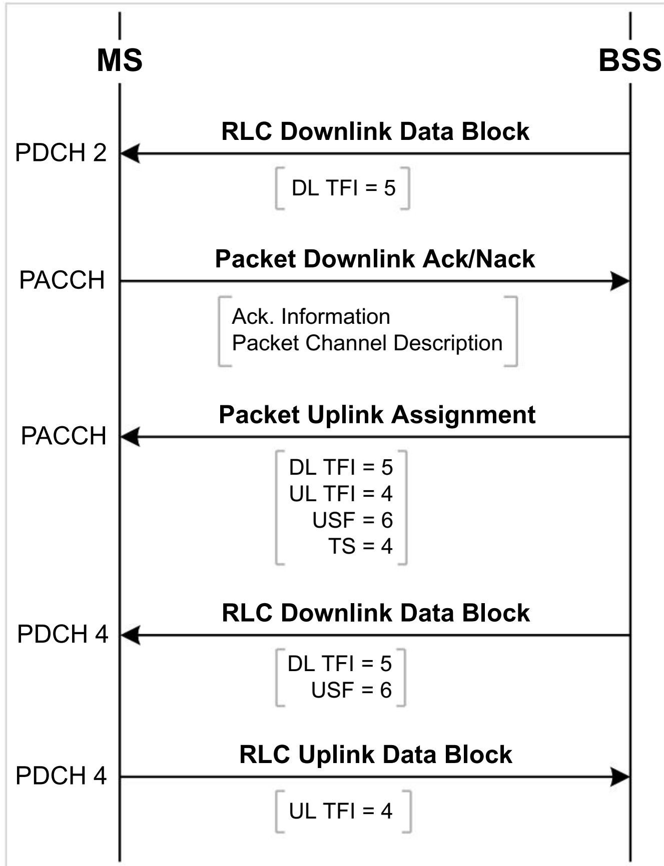


In the example presented in Fig. 4.16, the MS is a part of a downlink TBF identified by the downlink TFI value 5 and the allocated PDCH is TS 2. When the MS needs to initiate the uplink TBF, it sends the channel description in the packet downlink ack/nack message on the TS 2 in the uplink direction. In this case, the channel is acting as the PACCH. In response, the network also sends the packet uplink assignment, with the encoded TFI value as 5 on the TS 2. In this allocation for the uplink TBF, the MS is allocated to uplink TFI value 4, USF value 6, and TS 4. Note that the downlink RLC control block has a TFI and a direction bit. This indicates that this is for an MS with DL TFI of 5. The message in turn has an allocation that has TFI of 4.

Now in this case, the MS would keep monitoring the downlink blocks being sent on TS 2 for downlink TFI value 5 (as part of downlink TBF) and TS 4 for the USF value 6 (as part of uplink TBF). In case the MS finds the USF value 6 in any downlink block (data block, D1), it would transmit data on the next uplink data block with the TFI value encoded as 4. Please note that the downlink data block, D1, contains the downlink TFI value as 5, but this block

FIGURE 4.16

Example of Establishment of Uplink TBF with Ongoing Downlink TBF



does not belong to the downlink TBF of the MS being discussed. This is because D1 is being transmitted on the TS4, whereas the MS under discussion has been allocated TS 2 for the downlink TBF.

Note that in the RLC acknowledged mode, the data blocks transferred as part of a TBF are acknowledged by the receiving end. In the case of the uplink TBF, the network sends the receive status of the data blocks in the packet uplink ack/nack, whereas in the case of the downlink TBF, the MS sends the acknowledgment in the packet downlink ack/nack message. These messages provide information to the transmitter end whether a particular data block has been received successfully by the receiver or not. In case of unsuccessful reception of the data block, the transmitter retransmits the data block that has been nacked.

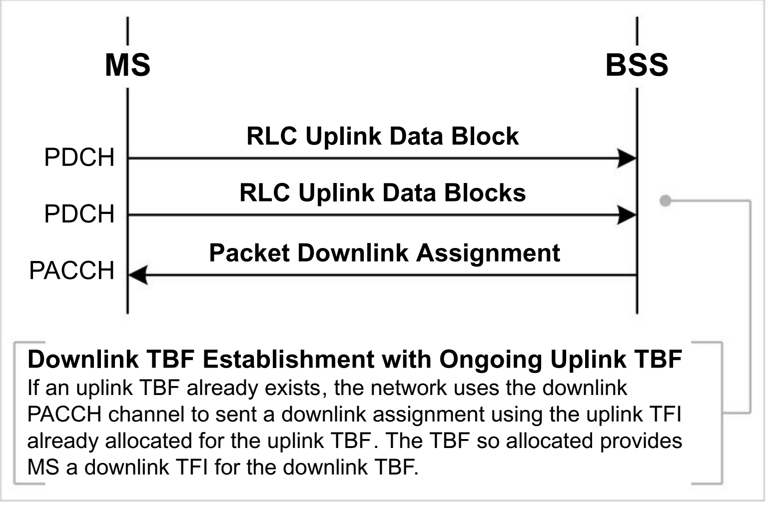
These messages consist of an acknowledgment bitmap that provides the information about the reception of a data block. Each bit provides information of a data block. A value "1" indicates an ack for the data block, whereas a value "0" indicates a nack for the same. The transmitter would resend all the data blocks that have the status as "0". The content of the acknowledgment message and the decoding is discussed in detail in Chapter 5 dealing with the RLC layer.

4.9.2 Downlink TBF Establishment for MS in Uplink TBF

It is possible for the network to initiate a downlink TBF for an MS that is involved in an uplink TBF. The network transmits the packet downlink assignment or the packet timeslot reconfigure message on the PACCH associated with the uplink TBF. The downlink TFI is assigned for the TBF being established. The procedure for the establishment of the concurrent TBF is as depicted in Fig. 4.17.

FIGURE 4.17

Establishment of Downlink TBF for an MS Involved in Uplink TBF



In the packet downlink assignment message, the MS is identified by the uplink TFI value of the ongoing uplink TBF. Once the downlink TBF starts, in addition to the data blocks belonging to the uplink TBF, the MS would also have to transmit control blocks (e.g., acknowledgment messages) belonging to the downlink TBF in the uplink direction. In this case, the control blocks are given preference over the data blocks.

4.10 Uplink TBF Release

Just as there are well-defined procedures for the establishment of a TBF, there are procedures for the release of a TBF also. The TBF release could be due to a variety of reasons. The most obvious reason would be the transfer of all data that had to be transferred as part of the TBF. The release could also be due to abnormal conditions. The aim of the present section is to discuss the TBF release procedures in the uplink while the next section describes the downlink TBF release procedure.

In the dynamic allocation of resources to an MS for the uplink TBF, the network indicates the use of the uplink resources to the MS using the USF value. However, in case of an open-ended TBF (i.e., where MS has not informed the number of RLC blocks to be sent), the network has no idea about when the entire data would be transferred and the TBF would terminate. In this situation, the network may allocate the resources to an MS that does not require the uplink radio resources and hence the resources would get wasted.

In order to avoid this waste of resources, the MS uses the countdown process to indicate the number of uplink RLC data blocks to be transmitted as part of the uplink TBF. During the end of the TBF, the MS starts transmitting a countdown value, CV value, in the uplink RLC data blocks. The CV value is calculated depending on the number of blocks to be transmitted, the number of PDCH allocated to the MS by the network, the coding scheme being used for the data transfer, and the block sequence numbers.

$$\text{Integer } x = \text{round}\left(\frac{\text{TBC} - \text{BSN}' - 1}{\text{NTS} \times \text{K}}\right)$$

$$\text{then, CV} = \begin{cases} x, & \text{if } x \le \text{BS\_CV\_MAX} \\ 15, & \text{otherwise} \end{cases}$$

where,

TBC, total number of RLC data blocks currently to be transmitted in the TBF.

BSN, absolute block sequence number of the RLC data block, with range 0 to (TBC – 1).

NTS, number of timeslots assigned to the uplink TBF, with range 1–8.

K, value is 1 for GPRS (its role is mainly for EDGE)

round(), the function rounds upward to the nearest integer.

BS_CV_MAX, this is a parameter broadcast in the PSI 1 message.

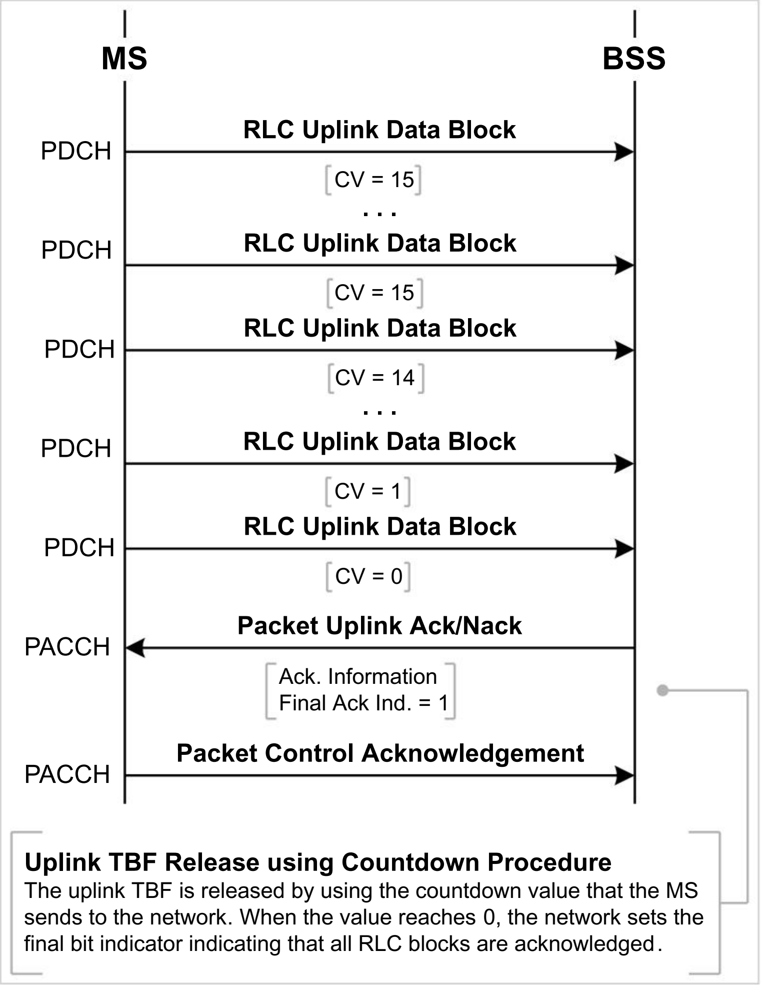
As depicted above, the value of the CV that is encoded in the uplink RLC data blocks is 15 if "x" is greater than the parameter BS_CV_MAX. However, as soon as the calculated value "x" is less or equal to the BS_CV_MAX, the countdown procedure starts and the encoded value of the CV is as calculated from the formula given here. The value BS_CV_MAX that controls the start of the countdown procedure is specified by the network in the PSI 1 message transmitted on the PBCCH.

The MS sends the last uplink RLC data block with the CV value encoded as 0, hence indicating to the network that the uplink TBF has ended. If the network has received all the data blocks successfully, it sends an acknowledgment for all the unacknowledged data blocks in the packet uplink ack/nack message with the final ack indicator set as 1. The network may also poll the MS to make sure that the MS has successfully received the packet uplink ack/nack message.

When the MS receives the packet uplink ack/nack with the poll bit set, it sends the packet control acknowledgment message and then releases the uplink TBF. The procedure for the uplink TBF release is depicted in Fig. 4.18.

Figure 4.18

Normal TBF Release After the Successful Completion of the TBF



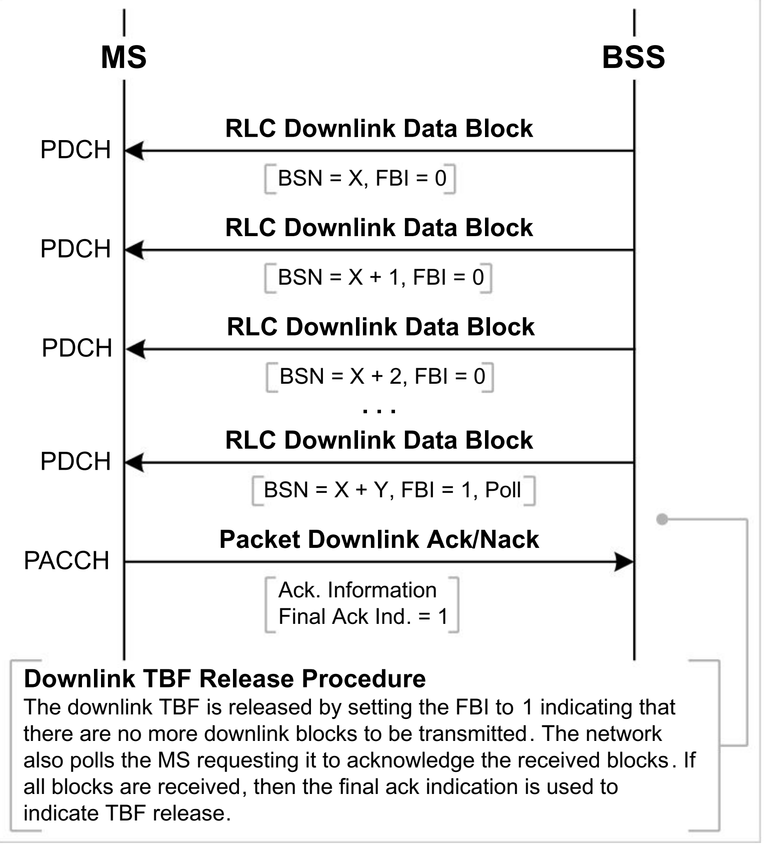
The network can initiate the uplink TBF release by sending the packet TBF release on the PACCH associated with the uplink TBF. If the cause for the release is indicated as a "normal release," the MS completes the countdown procedure and then releases the TBF. If, however, the cause is an "abnormal release," the MS aborts the uplink TBF and performs the abnormal release. The TBF is re-initiated by the MS using the PCCCH or the CCCH, as available in the serving cell.

4.11 Downlink TBF Release

In the downlink direction, the network does not need any information from the MS regarding the pending data blocks to be transmitted as part of the downlink TBF. Since all the required information is available with the network, the situation where downlink radio resources are unused never arises.

When the network transmits the last RLC data block of the downlink TBF, the Final Block Indicator (FBI) is set to 1 in the downlink RLC data block. This indicates to the MS that it is the last RLC block of the downlink TBF. The network sets the poll bit in the data block requesting the MS to send the acknowledgment of the downlink data blocks. In response, the MS sends packet downlink ack/nack message that acknowledges all the data blocks received. If all the data blocks have been decoded successfully, the MS sets the final ack indicator to the value 1. The downlink TBF is then released, as depicted in Fig. 4.19. The network can also initiate an immediate abnormal release of the downlink TBF by sending a packet TBF release on the PACCH. On receiving this message, the MS stops monitoring the assigned downlink PDCH and the TBF is released.

FIGURE 4.19 Normal Downlink TBF Release when the TBF has been Successfully Completed



Conclusion

This chapter provides an overview of the MAC layer and the functionalities provided by this layer. The MAC layer acquires the radio resources for the MS which uses it for the uplink and the downlink data transfers.

The MAC layer allows for the sharing of the same resources between the multiple MSs. The concepts of USF and the TFI were introduced that allow for the sharing of resources in the uplink and the downlink directions, respectively. In addition, it also performs the functions related to the cell reselection in order to ascertain that the MS is always served by the best possible cell.

The main function is to establish the TBF in the uplink or the downlink directions. It is possible to establish the TBF using the GSM channels in case the packet channels are not defined on the serving cell. The different procedures for these are described in detail.

The procedures for the release of the TBF are also described.

Further Reading

The main reference for this chapter is the specification [3GPP TS 44.060] that describes the functioning of the MAC and the RLC layers. The procedures for the establishment of the uplink and the downlink TBF and all the other functionalities provided by the RLC/MAC are described in this specification.

[3GPP TS 45.008] is another important specification that deals with radio subsystem link control. It covers a host of important topics like power control techniques applicable for MS and BTS and measurements in idle and active mode among other things.

[3GPP TS 45.010] deals with synchronization aspects in GSM and covers topics like timing advance.

[3GPP TS 43.064] is another important specification that provides an overview of the GPRS air interface and the RLC/MAC procedure.

Chapter 5

Radio Link Control Layer

5.1 Introduction

The Radio Link Control (RLC) and Medium Access Control (MAC) layer work very closely for the radio resource management in GPRS. While the MAC layer acquires the radio resources for the Mobile Station (MS), it is the responsibility of the RLC to utilize these resources for data transfer. The RLC layer predominantly acts as an interface between the higher layers and the MAC layer to exchange higher layer Protocol Data Units (PDUs) over the air interface. RLC is responsible for the segmentation of the upper layer PDU (i.e., Logical Link Control (LLC) PDU) into RLC/MAC data blocks for transmission at the sending side. These data blocks are transmitted on the resources that are acquired by the MAC layer. At the receiver end, the RLC reassembles the received data blocks to form the higher layer PDUs. The layer also provides the Backward Error Correction (BEC), thus enabling the selective retransmission of the unacknowledged RLC blocks. This is possible because of the acknowledged mode of operation of the RLC layer. Another key function of RLC layer is link adaptation whereby a particular coding scheme is selected based on channel conditions.

The RLC transmits the control information as well as the data blocks for the Temporary Block Flow (TBF). These blocks are encoded using the coding schemes defined in GPRS. While all the coding schemes can be used for the data transfer, only CS-1 coding is used for the control messages.

5.2 Modes of Operation

The RLC data blocks are transmitted in sequence and contain a Block Sequence Number (BSN). This number allows the reassembly of the blocks and selective retransmission of the blocks as well. These blocks to be retransmitted are determined by the acknowledgments sent by the receiver. This behavior is controlled by the "mode of operation" of the RLC layer. The following modes of operation are possible for RLC layer:

- **Acknowledged mode**. In this mode of operation, the receiver acknowledges the data blocks received. On receiving the acknowledgment message, the transmitting end retransmits the unacknowledged data blocks, if any.
- **Unacknowledged mode**. In this mode of operation, the acknowledgments are received from the receiving end but these are ignored. The data blocks are not retransmitted.

In the unacknowledged mode, the data blocks are sent to the upper layers in the same order as they are received. On the other hand, in case of acknowledged mode of operation, the data blocks are first collected and then transferred to the upper layers. The upper layers receive the data blocks in the same order as they were originally transferred.

The actual mode of operation used at RLC layer may differ from one data transfer to another and is determined at the time of TBF establishment. In the uplink direction, the mobile sets the RLC_MODE for the data transfer being initiated. In case of data transfers established using one-phase access procedures, the RLC mode is the acknowledged mode, by default. The MS can set the mode in the Packet Resource Request in case of the two-phase access procedure. If concurrent TBFs are being established, the mode is set in the Packet Downlink Ack/Nack message. In the downlink direction, the network sets the RLC mode for the TBF in the Packet Downlink Assignment message. The RLC_MODE field has a value of 0 for acknowledged mode and 1 for unacknowledged mode.

5.3 RLC/MAC Block Structure

While the MAC layer acquires the radio resources, the RLC layer provides the function of transmitting and receiving the data blocks constituting the TBF on these radio resources. The most important function is the segmentation and re-assembly of the data blocks. During segmentation, information is introduced in the transmitted block that allows the re-assembly at the receiving end. Also, every RLC block has a MAC header that provides information about the resource allocation, scheduling of the message, and so on. These are discussed in more detail in the present section. This section also discusses the structure of the data block and the control blocks transmitted as part of a TBF. The structure for the uplink and the downlink blocks is different and this has been discussed.

5.3.1 Control Block Structure

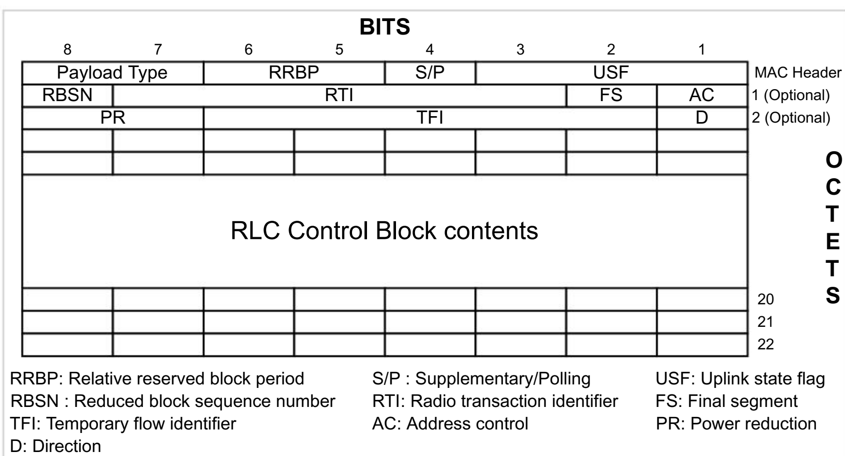
The signaling messages are carried in the control blocks. This includes messages for the establishment of the TBF as well as messages to acknowledge the data transfer. The coding scheme CS1 is the only coding scheme allowed for the control channels, and hence the uplink and the downlink control blocks can carry a maximum of 22 octets of control information. The basic structure of control block comprises the MAC header (1 octet) and RLC/MAC control information (22 octets). The exact structures of the control blocks in the uplink and the downlink directions different and explained next.

Figure 5.1 depicts the block structure for the downlink control block. As depicted above, the octets available to carry the control information may vary from 20 to 22 octets, depending on whether the optional elements are encoded or not. The information that is encoded is explained next.

- **Payload type:** This field indicates the type of payload being carried in the block. The payload may be a data block or a control block.
- **Relative Reserved Block Period (RRBP):** The network uses this field to schedule the transmission of a packet control ack message or Packet Associate Control Channel (PACCH) block from the mobile. This parameter indicates the position of the uplink radio block on which the mobile should transmit the message. The uplink radio block is relative to the frame number on which the downlink message with the parameter is received by the mobile. In other words, if a downlink message is received by the mobile at, say, frame number FN, the value of the parameter RRBP determines the radio block relative to FN, on which the uplink message should be transmitted. The 2-bit field indicates different delay depending upon exact values. The minimum is "00" that indicates 13 frame delay, while the maximum is "11" that indicates 26 frame delay.
- **Supplementary/Polling bit (S/P):** This single-bit parameter indicates whether the RRBP is valid or not. A value of 1 indicates it to be valid, whereas value of 0 indicates its invalidity.
- **Uplink State Flag (USF):** This field is used in the dynamic allocation of resources to indicate the user of the next radio block on the same timeslot in the uplink direction.
- **Reduced Block Sequence Number (RBSN):** This bit carries the sequence number of the downlink RLC/MAC control block. The RBSN bit is encoded as a binary number with the value of 0 or 1.

Figure 5.1

RLC/MAC Downlink Control Block Structure



- **Radio Transaction Identifier (RTI):** A control message may be segmented into a number of RLC/MAC control blocks and then transmitted on the radio interface. The RTI identifies the group of blocks that are carrying the different segments of the control message. The RTI field is 5 bits in length with the range of 0–31. Note that segmentation and re-assembly is not possible in uplink, and therefore this bit is not relevant here.
- **Final Segment (FS):** This single-bit field is used to indicate the block carrying the FS of the control message. Value of 0 indicates that current block does not contain the FS of an RLC/MAC control message, while the value of 1 indicates that the block is carrying the FS.
- **Address Control (AC):** As depicted in Fig. 5.1, octet 2 carrying the Tempory Flow Identifier (TFI) and the D bits is optional. This field indicates whether the octet is present in the block or not.
- **Power Reduction (PR):** This field indicates the power level reduction of the RLC block.
- **TFI:** The TFI identifies a TBF to which the control message belongs.
- **Direction (D):** The same TFI values could be used in the uplink and the downlink directions on the same timeslot at the same time. This is because the PACCH are bidirectional in nature. Thus, the direction bit indicates whether the TFI identifies an uplink TBF (value 0) or a downlink TBF (value 1).

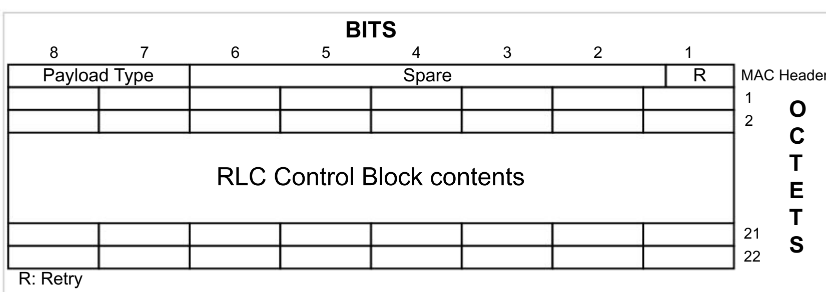
Figure 5.2 depicts the block structure in which the control information is carried in the uplink direction. The MAC header of one octet carries the following information:

- **Payload type.** This field indicates the type of payload being carried in the block. The payload may be a data block or a control block.
- **Retry (R) bit.** This indicates whether the mobile has sent an access request once or more in the recent access. This access request may be a channel request or a packet channel request message.

If one analyses this structure, it may be felt that it does not have any information, not even the basic information like TFI. Then the answer is to identify the source or the flow; the relevant information is included in the payload itself. For example, the TFI is a part of the packet downlink ack/nack message, and the packet control ack message carries the Temporary Logical Link Identify (TLLI) for MS identification (for those messages that use normal burst format). Few other messages like packet measurement report, packet enhanced

FIGURE 5.2

RLC/MAC Uplink Control Block Structure with MAC Header



measurement report, and packet uplink dummy control block carry the TLLI for MS identification. In summary, the uplink control format is simple and the relevant information is part of the payload.

Furthermore, if one compares the uplink and downlink frame formats, it is observed that the uplink has a very simple structure, while the downlink has much more elaborate structure. This difference arises from various factors including scheduling of uplink blocks and the associated control by the network (thereby requiring TFI, D, RRBP, and S/P fields), the segmentation and re-assembly in downlink (requiring RBSN, RTI, and FS), and few other reasons like in-band signaling for power control.

5.3.2 Data Block Structure

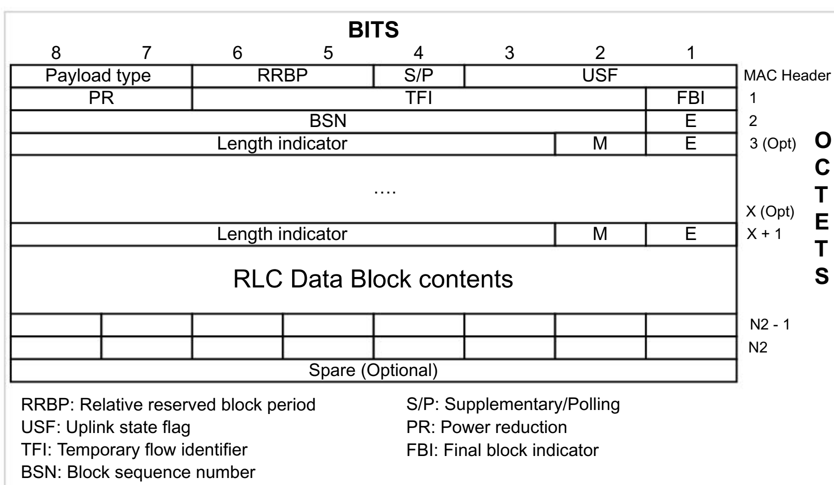
Any user data that has to be transferred as part of a TBF is transmitted on the radio interface as RLC/MAC data blocks. The coding scheme that is used for the data transfer in a TBF depends on the network allocation. The number of octets to be transferred in a data block hence varies from one TBF to another.

As part of the segmentation of the data, some parameters are introduced along with the data payload in order to assist in the re-assembly process at the receiving end. The data blocks may be transferred in the uplink direction or in the downlink direction. The basic RLC/MAC block has a MAC header and a RLC data block. The RLC data block in turn has three components; an RLC header, RLC data unit, and spare bits. There are different block structures in both uplink and downlink directions, as explained next.

Figure 5.3 depicts the downlink data block structure. Most of the parameters in this block are the same as the control block transmitted in the downlink direction (payload type, RRBP, S/P, USF, PR, and TFI). The parameters that are used exclusively in the downlink data block are discussed next.

Figure 5.3

RLC/MAC Downlink Data Block Structure



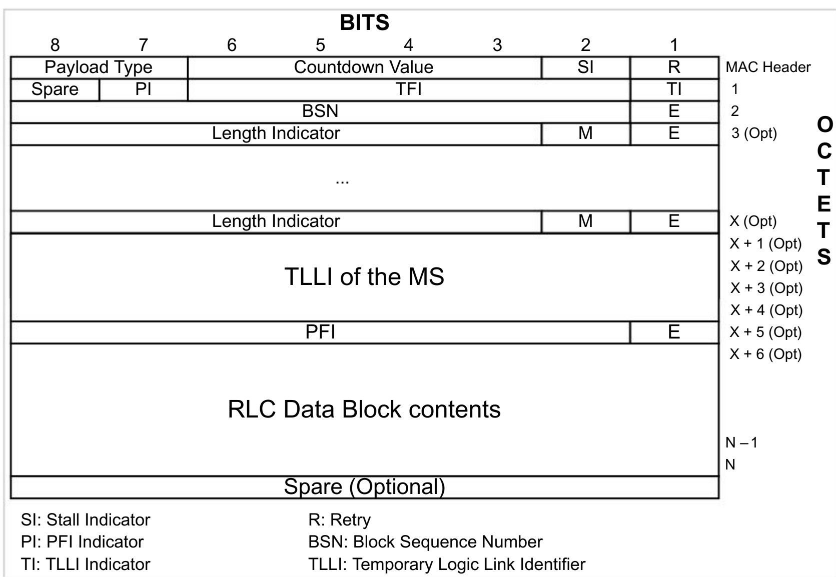
- **Final Block Indicator (FBI):** This bit with value "1" indicates that the block carrying this bit is the last data block to be transmitted as part of the downlink TBF.
- **BSN:** The data transferred as part of a downlink TBF is carried as blocks. Each of these blocks is numbered from 0 to 127 using 7 bits to allow the flow control and retransmission of lost blocks.
- **More (M) bit:** The M bit, along with the E bit and the Length Indicator (LI), is used to delimit LLC PDUs within a TBF. When the M bit is present (i.e., set to value of 1), it indicates whether another LLC PDU follows the current one within the RLC data block or not.
- **LI:** The LI is used to delimit upper layer PDUs within the RLC data block. The first LI indicates the number of bytes of the RLC data field belonging to the first upper layer PDU, the second indicates the number of bytes of the RLC data field belonging to the second upper layer PDU, and so on. The length field itself is optional and present only if the previous octet indicates that the header bytes are extended (i.e., E = 0).

Figure 5.4 depicts the uplink RLC/MAC data block structure. The important parameters for this block structure are discussed next.

- **Countdown Value (CV):** The MS uses this value to indicate that the ongoing uplink TBF is nearing completion. A value of "0" indicates the last block to be transferred as part of the uplink TBF. The CV field is 4 bits in length and is encoded as a binary number with the range of 0–15.
- **Stall indicator (SI):** This bit is used by the mobile to indicate to the network whether its RLC transmit window is stalled (value 1) or not (value 0).

FIGURE 5.4

RLC/MAC Uplink Data Block Structure



- **TLLI Indicator (TI):** This bit indicates whether the TLLI has been encoded in the transmitted RLC data block (value 1) or not (value 0).
- **TLLI:** The mobile encodes the TLLI value in the uplink data blocks for contention resolution. This is important for the one-phase access procedure.
- **Packet Flow Identifier (PFI) Indicator (PI):** As depicted in the Fig. 5.4, the octet that encodes the PFI (i.e., X + 5) is an optional octet. The PI bit indicates the presence (value 1) or absence (value 0) of this octet. If the TLLI is present in the data block, it is mandatory for the PFI to be present.
- **PFI:** This field contains an identity for Packet Flow Context (PFC).

If one compares the uplink and downlink frame formats, then certain fields are specific only to the downlink. This includes the final bit that indicates whether more blocks are available for transmission or not. Other downlink-specific fields include RRBP that carries scheduling information used in the uplink and the PR bit that carries power control related information. In the uplink, the CV field is similar to the final bit that allows network to know the number of blocks that remain to be sent in uplink. The TLLI is an other uplink-specific field that allows contention resolution when the ownership of the TBF has not been ascertained. Few other fields in uplink including PI and PFI are used for specific purposes.

5.4 RLC Operation during Control Message Transfer

The upper layer PDU that is transmitted using the services of the RLC layer may contain control information or application data. When control information has to be exchanged, the RLC control blocks are transmitted on the air interface. The flow of control information may be from the network to the MS or from the MS to the network. When the network sends any control information, the downlink RLC blocks are used. On the other hand, when the MS sends control information to the network, the uplink RLC blocks are used. It should be noted that only CS1 is allowed for any information to be transmitted on the control channels.

The RLC may segment the control information depending on the length of the control message. On the network side, if the control message does not fit into a single RLC block, then the information is split into multiple control blocks. When the last segment of the message is transmitted, the FS field of the block containing this segment is set. In case the FS of the message does not fit into an integral number of octets in the control block, filler octets are used for the remaining control block. This block is then transmitted on the air interface.

At the MS end, the RLC/MAC blocks are collected until all the blocks carrying the control message have been received. Once all the blocks have been received, the control message is obtained and necessary action is taken.

The MS does not segment any RLC/MAC message.

5.5 RLC Operation during Data Block Transfer

The main aim of the GPRS is to transfer data, either from the MS to the network or from the network to the MS. The RLC allows the transfer of the upper layer PDU in the form of the RLC data blocks. When the data are transferred from the MS, the uplink RLC data blocks are used, whereas in case the data are flowing from the network to the MS, the downlink RLC data blocks are used. The size of the data blocks is determined by the coding scheme to be used for the data transfer. The coding scheme for a TBF is determined at the time of the resource allocation in the assignment message. The RLC segments the upper layer PDU into RLC data blocks and transmits them on the air interface. Table 5.1 lists the possible coding schemes and the number of octets allowed. It also specifies the number of spare bits for each coding scheme.

Table 5.1

Coding Schemes and the Data Block Size

Coding Scheme	*Size of Data Block Without Spare Bits (in octets)*	*Spare Bits*	*Size of Data Block (in octets)*
CS1	22	0	22
CS2	32	7	32 7/8
CS3	38	3	38 3/8
CS4	52	7	52 7/8

The size of the upper layer PDU determines the number of data blocks required for the transfer. It may happen that the upper layer PDU to be transmitted is such that it does not fit into an integral number of RLC data blocks. In this case, the last data block would not be completely occupied and would have some empty octets. In order to utilize the resources more efficiently, the first data block of the next upper layer PDU to be transmitted is placed in the last data block. If however, the last empty block belongs to the last upper layer PDU, the rest of the block is filled with filler octets.

Each data block that constitutes a data flow has certain parameters that are encoded. One of the important parameter is the TFI, which identifies the TBF to which the data block belongs. Additionally, each block is allocated a sequence number, the BSN that is transmitted in the data block. These block numbers are allocated consecutively and they allow the receiving end to resegment the upper layer PDU. Also, the sequence numbers are used to acknowledge the blocks that have been correctly received at the receiving end. These concepts are explained in detail in the later sections of the chapter.

At the receiver end, the mode of RLC operation is important. As stated earlier, RLC can operate in two modes: the acknowledged mode and the unacknowledged mode. In the acknowledged mode, the data blocks are collected until all the data blocks belonging to the upper layer PDU have been received. These blocks are then reassembled into the upper layer

PDU and passed to the higher layers. If the RLC mode is acknowledged mode, the data blocks are delivered to the higher layers in the same order as these were transmitted originally. On the other hand, in the unacknowledged mode, the data blocks are delivered to the upper layers in the same order as they are received. In case some blocks are not received, fill bits "0" are filled in their place.

5.6 Peer-to-Peer RLC Operation

When a TBF is established between two RLC/MAC entities, one of them is the transmitter and the other is the receiver (note that TBFs in GPRS are unidirectional in nature). The RLC data transfer between these entities is governed by the rules of flow control. In other words, there is a limit to the number of blocks that the transmit side can send before an acknowledgment is required from the receiver. In case an acknowledgment is not received, then the unacknowledged blocks are retransmitted. The maximum number of blocks that can be transmitted without any acknowledgment is governed by the Window Size (WS) on the transmit side. This window is the "send window". The WS for GPRS is 64. Similarly, the receiver has a "receive window". Apart from the WS, each RLC/MAC data block has a sequence number, the BSN, which carries the sequence number of the packet. In addition to these parameters, some other parameters are maintained by both sides. The parameters that are important for the functioning of the RLC layer are listed in Table 5.2.

The transmit side maintains the details of the transmitted blocks and the acknowledged blocks. The parameters that are maintained are as follows:

- **Send State Variable, V(S).** This parameter indicates the sequence number of the next RLC data block to be transmitted. V(S) is set to the value of 0 at the beginning of each TBF in which the RLC endpoint is the transmitter. When the transmission of the blocks is in sequence, the BSN of a data block to be transmitted would be the same as the variable V(S).
- **Acknowledge State Variable, V(A).** This parameter indicates the oldest unacknowledged RLC data block.
- **Acknowledge State Array, V(B).** This array provides information about the acknowledgment status of WS number of data blocks relative to V(A) modulo Sequence Number Space (SNS), where WS is the window size. In other words, this array provides information whether the data blocks in the send window have been acknowledged or not. These values are updated from the acknowledgments received from the receiver entity in the Received Block Bitmap (RBB) of the packet ack/nack message.

The receiver end maintains the details of the received data blocks as well as the data blocks that have not been received. The parameters used at the receiver are as follows:

- **Receive state variable, V(R).** This variable indicates the sequence number that is one higher than the highest-numbered block that has been received. V(R) is set to the value "0" at the beginning of each TBF in which the RLC endpoint is the receiver.

TABLE 5.2

Parameters Important for the RLC Procedures

Category	*Parameter Name*	*Significance*	*Range*
General	BSN	Sequence number of the data block being transmitted as part of the TBF	0–127
	WS	Size of the transmit window or the receive window	64
	SNS	The number of distinct BSN possible	128
Transmitter end	Send state variable, V(S)	Sequence number of next RLC data block to be transmitted	0–(SNS–1)
	Acknowledge state variable, V(A)	Sequence number of oldest RLC data block that has not been acknowledged	0–(SNS–1)
	Acknowledge state array, V(B)	Array indicating the acknowledgment status of WS data blocks relative to V(A) modulo SNS	Pending_ Ack/Nacked
Receiver end	Receive state variable, V(R)	Sequence number one higher than highest RLC data block received	0–(SNS–1)
	Receive window state variable, V(Q)	Lowest sequence number that has not been received as yet	0–(SNS–1)
	Receive state array, V(N)	Array indicating the received status of WS data blocks relative to V(R) modulo SNS	Received/Not received

- **Receive window state variable, V(Q).** It indicates the lowest sequence block that has not been received as yet. Hence, this acts as the start of the receive window.
- **Receive state array V(N).** This array maintains the receive status of WS previous blocks, relative to the variable V(R) modulo SNS.

Consider a TBF transfer between two entities, A (transmit) and B (receive), under ideal conditions without any data loss as depicted in Fig. 5.5. The function of the variables V(A), V(S), V(R), and V(Q) is depicted in the figure. Initially, the send variable V(S) and the acknowledge variable V(A) at A are set to 0. Similarly, the receive variable V(R) and the receive window state variable, V(Q) at B are set to 0. When A transmits the first RLC data block (BSN = 0), then V(S) = 1 and V(A) = 0. This is because the next data block to be transmitted would have the BSN value as 1. Since the transmitted blocks have not been acknowledged, the V(A) remains 0. At the receiver end, the variables are updated such that V(R) = 1 and V (Q) = 1,

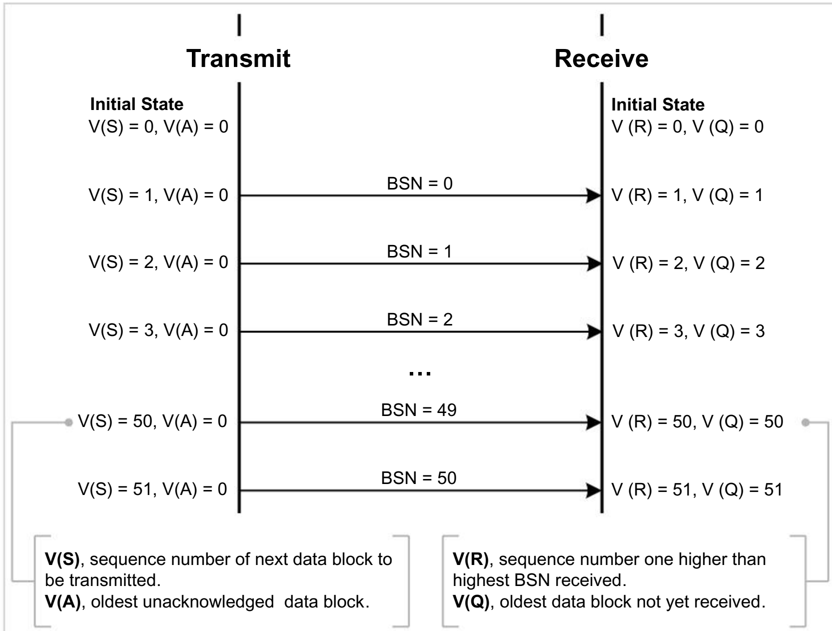


FIGURE 5.5

Function of Transmit and Receive RLC Parameters

indicating that the block with BSN = 0 is the highest block received as yet. Also, parameter V(Q) =1 indicates that BSN = 1 is the lowest block that has not been received also. As depicted in the figure, the parameter values at both transmit and receive ends get updated as the data blocks are transmitted.

In the scenario being considered, the data block starting from BSN = 0 to BSN = 50 have been transferred without receiving any acknowledgment from the receiver for any of the transferred blocks. The value of the parameter V(A) hence remains 0 at the transmit side.

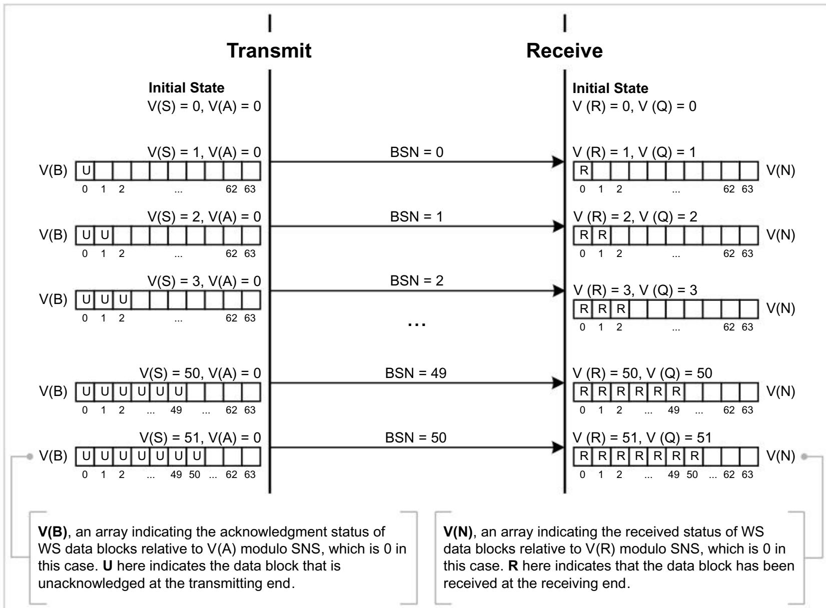
Another important aspect to consider is the acknowledge state array V(B) and the receive state array V(N). The values of these parameters for the data transfer discussed here are depicted in Fig. 5.6. Let us first consider the transmitter end. As the data blocks are transmitted, the transmitter keeps updating the acknowledgment status in the array V(B). In the present scenario, none of the blocks has been acknowledged and hence all the blocks (BSN = 0 to BSN = 50) are marked as “PENDING_ACK,” represented by “U.”

If the acknowledgments are not received for any of the data blocks lying in the window, the window gets stalled. In other words, the transmission of the higher sequence number data blocks would not happen and the transmitter would start retransmitting the unacknowledged blocks again. Once the acknowledgment is received for the blocks, the window gets unstalled and the transmission of the higher-numbered blocks starts.

In the present scenario, the transmit window is not stalled, i.e., [V(S) < V (A) + WS] modulo SNS, and hence the data block with BSN = V(S) (value 51) can be transmitted and the acknowledge state array V(B) is updated for the transmitted block as “PENDING_ACK.”

Figure 5.6

Function of V(B) and V(N) During Data Transfer



If no acknowledgments are received from the receive side, there may be situation when the sequence number of the block to be transmitted is such that V(S) = V(A) + WS. In this case, the transmit window is said to be stalled. In this case, the oldest unacknowledged data block is retransmitted followed by the next oldest block and so on. The retransmissions are repeated till the acknowledgment is received and the window is unstalled.

In the present case, the transmission of the blocks can continue till BSN = 63 without any acknowledgment from the receive side. At this stage, the sending side parameters are as follows: V(S) is 64, V(A) is still 0, and all 64 array indexes in V(B) starting from 0 to 63 are marked "U." The receiving side parameters are as follows: V(R) is 64, V(Q) is 64, and all 64 array indexes in V(N) starting from 0 to 63 are marked "R." After this, the window stalls and the transmit side starts retransmitting the blocks with BSN = 0, BSN = 1, and so on.

The acknowledgment for the data block is received in the packet ack/nack messages from the receiver. The uplink data blocks are acknowledged in the packet uplink ack/nack message from the network and the downlink data blocks are acknowledged by the mobile in the packet downlink ack/nack message. The acknowledgment messages are discussed in detail in the subsequent sections.

5.7 Acknowledgment and Polling

While the previous section focused on the sending side and how it impacts the parameters of the sender and receiver, this section talks about acknowledgments. Note that in the RLC-acknowledged mode of operation, the receiver provides information about the successful or unsuccessful reception of the data blocks. Based on these acknowledgments, the unacknowledged blocks are retransmitted by the transmitter. The acknowledgments are received in the ack/nack message. In particular, the MS sends the acknowledgments for the downlink blocks in the packet downlink ack/nack message, and the network acknowledges the uplink blocks in the packet uplink ack/nack message. In other words, the uplink message acknowledges downlink packets and downlink message acknowledges uplink packets.

The acknowledgments of the data blocks are set in the acknowledgment bitmap. The status of one block is represented by a bit. A bit "1" indicates successful reception of the data block (i.e., an "Ack"). On the other hand, a bit value "0" indicates an unsuccessful reception of the data block or a "Nack." The present section discusses the packet uplink ack/nack message in detail.

The message identifies the uplink TBF by the uplink TFI encoded in the message (Fig. 5.7). The channel coding indicator field indicates the channel coding scheme that the MS shall use when transmitting on the uplink. The Ack-Nack description contains the acknowledgment information about the data blocks constituting the TBF. The parameters sent as part of this information element are as follows:

- **Final ack indication.** This bit indicates whether the TBF has been received completely by the receiver or not. A value "1" indicates that all the blocks constituting the TBF have been received and hence the TBF is complete. On the other hand, a value "0" of this field indicates that some of the blocks have not been received successfully by the receiving side and should be retransmitted by the sending side. The information about these nacked blocks is contained in the RBB.
- **Starting Sequence Number (SSN).** This field contains the value of V(R) when this message was transmitted.
- **Received Block Bitmap (RBB).** This is the most important field in the message that provides the exact information about the data blocks that have been received successfully or unsuccessfully by the receiver. Every position in the bitmap represents a BSN, and the bit value indicates whether the block of that BSN was received

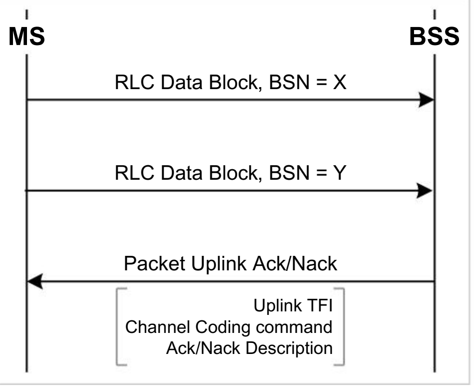


FIGURE 5.7

Acknowledgment of Uplink Data Blocks

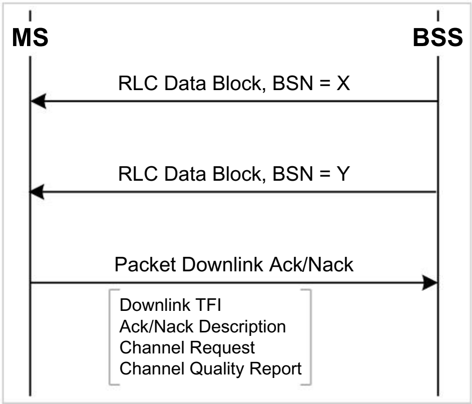
successfully (value 1) or not (value 0). The bitmap consists of 64 bits and hence can provide the acknowledgment status for 64 blocks. Note that the size of the bitmap is the same as the size of the GPRS window size, which is 64.

Apart from the fields discussed here, there are few other fields in the ack/nack message that can be used to update the timing advance and power control parameters.

Similarly, in the case of a downlink data transfer, the network polls the mobile to send the acknowledgment message for the downlink data blocks constituting the downlink TBF. The acknowledgment is done using the packet downlink ack/nack message. The network uses the supplementary/polling, S/P field, and the RRBP of the downlink data blocks to request for the ack/nack message. The signaling flow is as depicted in Fig. 5.8.

FIGURE 5.8

Acknowledgment of Downlink Data Blocks



In the downlink ack/nack message, the downlink TBF is identified by the downlink TFI. The packet downlink ack/nack message contains an ack/nack description that is similar to the packet uplink ack/nack message. Thus, the ack/nack description carries the final ack indication, SSN, and the RBB. Moreover, this message contains additional information from the MS that is used to report the channel quality in the downlink and to optionally initiate an uplink TBF. In particular, packet downlink ack/nack message carries the following important information elements (apart from the TFI and the ack/nack description):

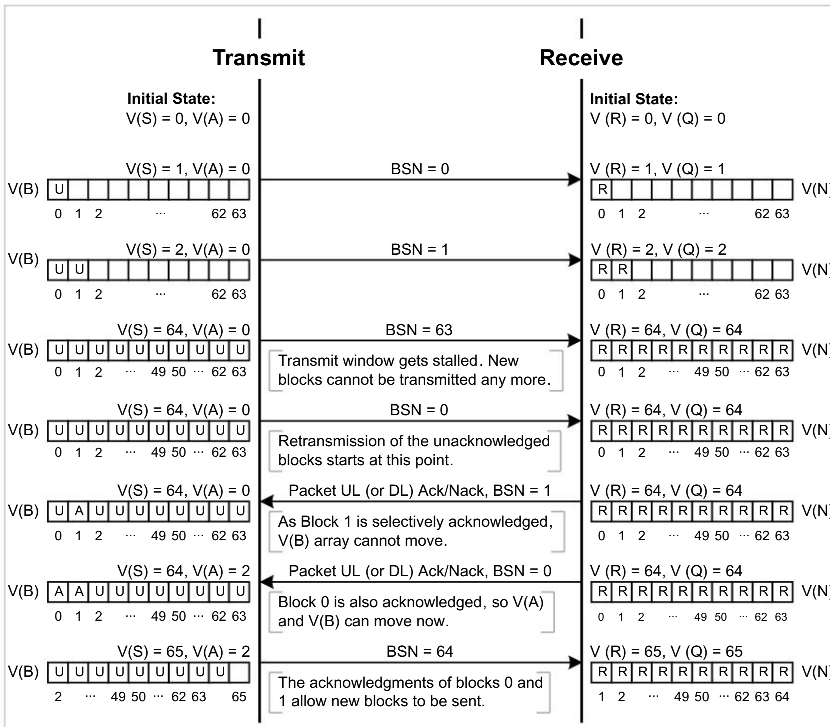
- **Channel request.** This field allows the MS to carry channel request for establishing an uplink TBF in the same message that carries acknowledgments for the downlink. This way, MS can avoid using the RACH/PRACH and use the PACCH instead, serving two purposes through one message.
- **Channel quality report.** This information element provides information about the measurements on the radio interface. The parameters include the C value (i.e., received signal level at the MS), signal variance, RxQual (i.e. received signal quality at the MS), and interference levels.

5.7.1 Acknowledgment of RLC Blocks

Consider the data transfer discussed in Section 5.6, where the scenario considered was the retransmission of the blocks from BSN = 0 when the send window gets stalled. Now in this section, consider that the receiver starts sending the acknowledgments for the data blocks in the packet ack/nack message (may be uplink or downlink ack/nack message depending on uplink or downlink TBF). The variables at both the transmitter and the receiver end get updated as the acknowledgments are received, as depicted in Fig. 5.9.

FIGURE 5.9

RLC Parameter Updation as the Data Transfer Progresses



As depicted in Fig. 5.9, after the window stalls at the transmit end, the retransmission of the unacknowledged blocks starts. Thereafter, in the example provided, the receiver sends the acknowledgment for the block with BSN = 1 in the acknowledgment message (i.e., packet uplink (or downlink) ack/nack). Note that it is not considered why the receiver selectively acknowledges BSN = 1 as the idea is just to highlight the use of different variables. It is possible that block BSN = 0 is lost and then retransmitted. Such exchanges are not part of this section and the reader is referred to next section for such exchange.

On receiving the ack, the transmit end updates the acknowledgment array and marks the BSN = 1 as acked, indicated by "A" in the V(B) array in the figure. The block, BSN = 0, however, is still unacknowledged and hence it is marked as "U," indicating that it is an unacknowledged block. The parameter V(A) indicates the oldest unacknowledged data block and hence it has the value 0. The array V(B) also does not progress due to this unacknowledged block with BSN = 0. In the next acknowledgment message, the BSN = 0 block is acknowledged. Once the block is marked as "A" in the array V(B), V(A) gets updated to the value 2, indicating that the oldest unacknowledged block has a value of 2. The window V(B) now progresses with respect to the parameter V(A) and moves to BSN value 65. Thereafter, as depicted in the figure, the block with BSN = 64 is transmitted and the values of the parameters V(S), V(R), and V(Q) are updated to 65. The array V(N) progresses to BSN values 1–64 as well.

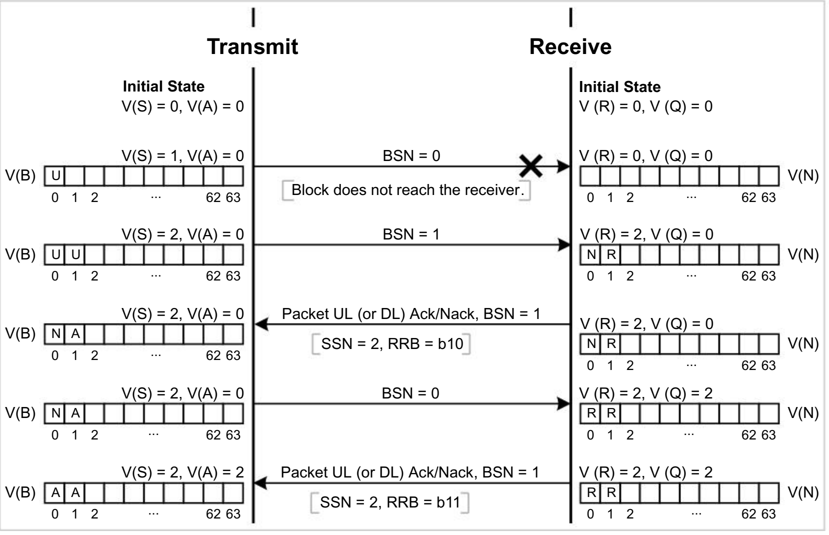
5.7.2 Negative Acknowledgment of RLC Blocks

As another example, consider a data transfer constituting two data blocks, BSN = 0 and BSN = 1. Fig. 5.10 depicts the transfer of these data blocks and the values of the RLC parameters at the transmitter and the receive ends. The first acknowledgment message from the receiver, as depicted in the figure, contains a bitmap with binary value b10. This indicates a positive acknowledgment for the BSN = 1 and a negative acknowledgment for BSN = 0. On receiving this acknowledgment message, the transmitter updates the acknowledgment array with the values N (indicating negative acknowledgment) and A (indicating positive acknowledgment) for the BSN values 0 and 1, respectively.

The negatively acknowledged block BSN = 0 is retransmitted by the transmitter. This time, the receiver sends a second acknowledgment message containing a bit 1 for the BSN = 0 indicating positive acknowledgment. The value V(A) gets updated on receiving the acknowledgment for the BSN = 0.

FIGURE 5.10

Retransmission of Negatively Acknowledged Data Blocks



Another thing to note in the figure is the value of V(Q). Since the block BSN = 0 is not received by the receiver end, the value of V(Q) is not updated even though the transmit end updates its values. On the successful reception of the BSN = 1, the parameters V(R) and V(Q) take values 2 and 0, respectively. The value V(R) = 2 indicates that the next higher value of BSN expected from the transmitter is 2. The value V(Q) = 0 indicates that the lowest value of the BSN not received is 0. On the retransmission of the BSN = 0, the value of V(Q) is also updated to 2.

5.7.3 Handling of Receive Block Bitmap

The packet downlink and the packet uplink ack/nack messages contain the ack/nack description field that provides information about the acknowledgment status of the RLC data blocks. Consider a data transfer that constitutes the transfer of BSN = 0 to BSN = 63. In response, the receiver sends an acknowledgment bitmap in the acknowledgment message that contains the **SSN as 64 and the RBB as 64-bit binary value of** 11111111110000001111111 1111111000000001111111111111111111111111101.

Table 5.3 Depicts the decoding of the acknowledgment status of the data blocks using the RBB and the SSN. The bitmap is indexed relative to the SSN as follows:

BSN = (SSN – bit_number) modulo 128, for bit_number = 1–64.

Table 5.3

Decoding of the Ack/Nack Description

Bit number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
BSN	63	62	61	60	59	58	57	56	55	54	53	52	51	50	49	48
Ack status	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
Bit number	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**
BSN	47	46	45	44	43	42	41	40	39	38	37	36	35	34	33	32
Ack status	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
Bit number	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**
BSN	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
Ack status	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
Bit number	**49**	**50**	**51**	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**	**61**	**62**	**63**	**64**
BSN	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Ack status	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1

The BSN values can range between the values (SSN – 1) mod 128 to (SNS – 64) mod 128. In this example, the BSN values 1, 26–33, and 48–53 have been negatively acknowledged by the receiver. All the other data blocks are positively acknowledged.

Conclusion

This chapter provides an overview of the RLC layer. The MAC layer functions very closely with the RLC layer. While the MAC layer acquires the radio resources, the RLC layer uses these resources and transfers the information on the radio interface. This information may be the control messages or it may be data.

If the control information has to be transferred, the RLC control blocks are used. On the other hand, in case of data transfer, the RLC data blocks are used. The structure of the control and the data blocks is different in the uplink and the downlink direction and is discussed in detail.

RLC can work in two modes of operation: acknowledged mode and the unacknowledged mode. In the acknowledged mode, the acknowledgments from the receiving entity allow selective retransmission of the data blocks that are negatively acknowledged by the receiving entity. The acknowledgment messages and the decoding are also discussed. The window concept of RLC and the parameters that are maintained at the transmitter and the receiver ends have been described as well.

Further Reading

The main specification referred in this book is the [3GPP TS 44.060]. This specification describes the RLC/MAC layer and the signaling procedures for the TBF establishment and the data transfer. It provides complete information about the structure of the data blocks and it describes the fields of the blocks. The specification also provides details of the flow control and acknowledgment between RLC peers.

Chapter 6

GPRS Mobility Management

6.1 Introduction

The very essence of a mobile network is the ability of a mobile subscriber to roam within and across multiple Public Land Mobile Networks (PLMNs), and still be reachable and able to access network services. To reach a mobile subscriber within a mobile network, it becomes essential for the network to maintain the current location of the subscriber. This maintaining of the location information forms the basis of the Mobility Management (MM) procedures. Sometimes, the term location management is also used in place of MM to refer to the same function. The two terms are also used interchangeably in this chapter to refer to the same concept.

This chapter first briefly describes the MM procedures used within a GSM network. This helps in easier understanding of the GPRS MM procedures (also called GMM procedures), which can be seen as an extension of the GSM MM procedures. To appreciate the GMM procedures, it is essential to understand the state machine used for GPRS MM. The GMM state machine is discussed, and this is then followed by a detailed discussion of the GMM procedures.

6.2 Mobility Management in GSM

In PSTN networks, the end terminals (telephones, etc.) are considered fixed entities and reachability to these terminals is through preconfigured fixed routes. However, this cannot be the case in a mobile network, where the terminals (mobile phones) are not stationary or fixed. To reach a mobile subscriber in a GSM network (e.g., in case of an incoming call for the subscriber), it becomes essential for the network to know the exact location of the mobile subscriber. However, tracking the location of each and every subscriber at a very fine geographical granularity is in itself a Herculean task. Considering the fact that a mobile subscriber is mostly idle and only occasionally active (i.e., participates in a call), a GSM MM state model is defined, which suggests a practical mechanism for tracking the subscriber location and movement.

6.2.1 GSM MM State Model

Figure 6.1 depicts a simplified three-state Finite State Machine (FSM), which defines when a GSM network tracks the mobile subscriber, and at what level of granularity. The following states are depicted in the state machine:

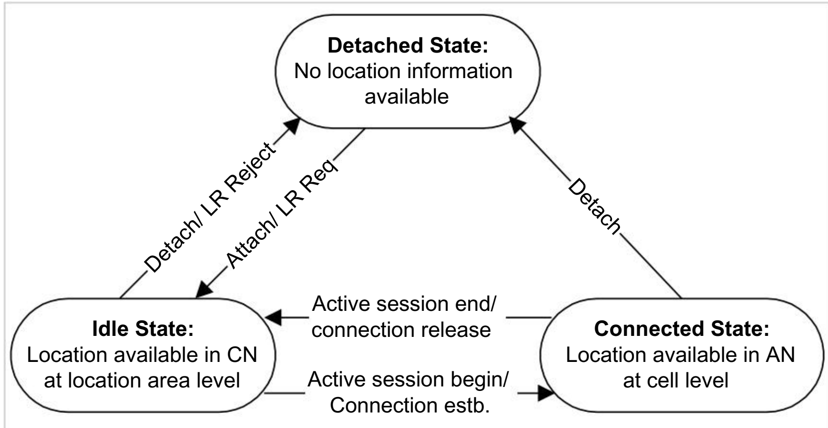
- **Detached state.** In this state, the network does not know the location of the subscriber. In other words, the subscriber is not reachable by the network. This may be the case when, for example, the subscriber's handset is in a "power-off" state, or when the SIM is removed. A transition from the "detached" state to one of the other two states happens when the subscriber registers his presence in the network (or attaches to the network, e.g., on power-on), via the Location Registration (LR) process.
- **Idle state.** In this state, the location of the subscriber is known to the network (the core network, to be precise, with which the subscriber is registered). However, there is no active session for the subscriber (i.e., no voice call, etc. in progress). Hence, the subscriber is assumed to be idle in this state. From the "Idle" state, a transition to the "connected" state happens when an active session is established for the subscriber. Participation in a session is generally characterized by the establishment of a connection, and hence the name "connected" state. From the "idle" state, a transition back to the "detached" state may happen, for example, due to powering-off of the handset, or removal of SIM from the equipment, or because of failure of the network to register the subscriber.
- **Connected state.** In this state, the location of the subscriber is known to the network (both the Access Network, and the Core Network). Further, there is an ongoing session active for the subscriber (e.g., a voice call). Thus, in this state, the subscriber is not considered to be idle, since he/she is involved in an ongoing session. From the "connected" state, a transition back to the "idle" state happens on completion of the ongoing session. A transition can also happen from the "connected state" to the "detached" state, for example, as a result of powering-off the handset, or removal of SIM from the equipment.

Besides the state transitions, Fig. 6.1 also depicts the granularity of the information available to the network as regards the location of the subscriber. In the "detached" state, no location information is available to the network, since the subscriber is not registered with the network. Location information is available to the network in both the "connected" and the "idle" states. However, the granularity of the location information available differs in both these states.

In the "idle" state, the location information of the subscriber is maintained by the core network (and not by the Access Network) only at the granularity of LA. However, in the

Figure 6.1

GSM Mobility Management State Model



"connected" state, the Access Network also knows the location of the subscriber, at a finer cell-level granularity. The logic behind this design is that the exact location of the subscriber is only required when there is an active session for it, and not otherwise. The information maintained in the different network elements is as depicted in Table 6.1.

Table 6.1

Maintenance of Location Information in GSM

Location Identifier	*Network Element*		
	BSC	*MSC/VLR*	*HLR*
Cell Identity	Yes*	No	No
Location Area Identity	No	Yes	No
MSC/VLR Number	No	N/A	Yes

*In "connected" state (refer to GSM MM state model).
N/A: not applicable.

6.2.2 GSM MM Procedures

GSM MM procedures can be considered to be a collection of two broad sets of procedures, namely

- Mobile Station (MS) Idle Mode Procedures, and
- MM procedures on the MS-Core Network interface

MS Idle Mode procedures in the MS include procedures for the MS to enter into the "idle" state from the "detached" state, as well as the procedures while within the "idle" state. When an MS is switched on, it tries to make contact with a PLMN. The MS selects the PLMN either automatically or manually as a result of user selection. The MS next looks for a suitable cell of the chosen PLMN and tunes to the control channel of that cell. It then registers its presence in the registration area of the cell by means of an LA or International Mobile Subscriber Identity (IMSI) Attach procedure, and moves out of the "detached" State.

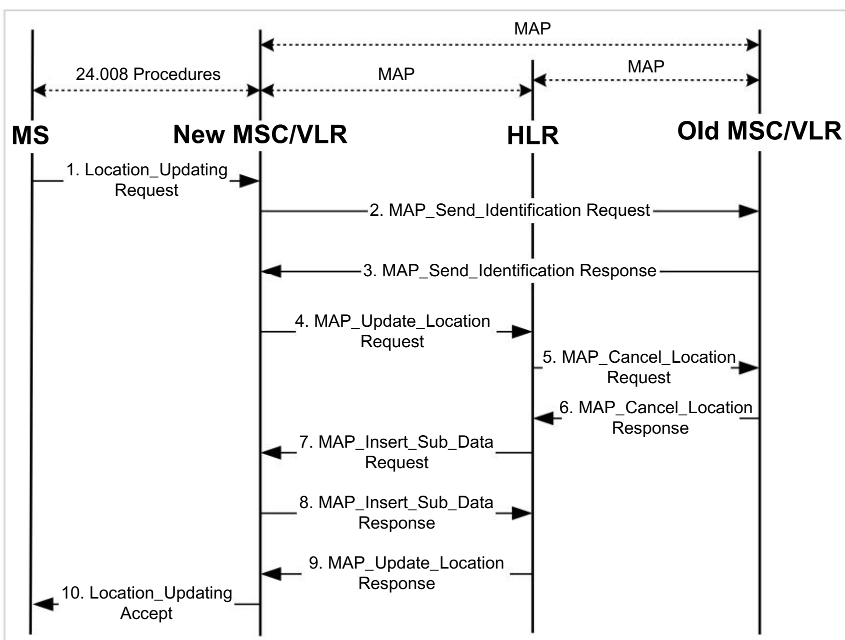
Once in the "idle" state, if the MS moves out of the coverage area of the selected cell, or if it finds a more suitable cell, it reselects the more suitable cell of the selected PLMN. If the new cell is in a different registration area, an LA request is made to register the new location information of the MS with the network. Similarly, if the MS loses coverage of the PLMN, it selects a new PLMN either automatically or as a result of manual user intervention, and repeats the idle mode procedures. MS idle mode procedures are discussed in more detail in [3GPP TS 43.022].

The MM procedures on the MS-Core Network interface (refer [3GPP TS 24.008]) are performed by the MM sublayer, which resides on top of the radio resource sublayer in the control plane (refer to GSM control plane architecture discussed in Chapter 1). MM procedures on the MS-Core Network interface are triggered as a result of the LR procedures, executed by the MS as part of its idle mode procedures. Figure 6.2 depicts the message flow diagram for the LA scenario (also called the normal location update procedure), which is used by the MS to update its LA information with the core network. The sequence of steps followed is

1. The MS initiates a Location_Updating Request message, which includes the Temporary Mobile Subscriber Identity (TMSI) to identify the subscriber to the MSC/VLR. In addition, the Location Area Information (LAI) of the previous LA in which the subscriber was registered is also included in this message.
2. The new VLR checks if the subscriber is known to it. The VLR may not have information about the subscriber if the subscriber has migrated to it for the first time. If this happens to be the case, the new VLR sends a MAP_Send_ Identification Request message to the

Figure 6.2

Normal Location Update Sequence Diagram



old VLR (which is determined from the LAI of the previous LA), requesting it to send the subscriber's IMSI. The TMSI is sent in this message to identify the subscriber to the old VLR.

3. The old VLR sends the IMSI to the new VLR as part of the MAP_Send_Identification Response message. Step 2 and 3 may however be omitted if the subscriber is known to the VLR, in which case the VLR already has the IMSI information of the subscriber in its own database.
4. At this point, the new VLR may, at its own discretion, initiate the authentication procedure to authenticate the mobile subscriber. The VLR next sends a MAP_Update_Location Request message to the Home Location Register (HLR) to update the location information of the subscriber in the HLR. The new MSC and VLR numbers are included in this request message. The subscriber is identified in the HLR by using the IMSI received in the message from the VLR.
5. On receipt of this message, the HLR sends a MAP_Cancel_Location Request message to the old VLR to indicate to it that the subscriber has moved out from its service area.
6. The old VLR, on receipt of this message, deletes the subscriber's information stored in its data store. It then acknowledges the message from HLR by sending it a MAP_Cancel_Location Response message.
7. Next, the HLR sends relevant subscriber data to the new VLR. This data is required by the latter to provide services to the subscriber. The subscriber data is sent as a part of the MAP_Insert_Subscriber_Data Request message to the new VLR.
8. The new VLR updates its datastore with subscriber data and sends a MAP_Insert_Subscriber_Data Response message to the HLR.
9. The HLR then acknowledges successful completion of the location update procedure, by sending a MAP_Update_Location Response message to the new VLR.
10. The new VLR in turn confirms the successful completion of the location update procedure by sending a Location_Updating Accept message to the MS, thus completing the location update procedure.

The same procedure of LR, as described earlier, is also repeated periodically, even if the MS has not moved or changed its LA. This is called the periodic LR procedure. It acts as a heartbeat mechanism between the network and the MS, whereby the MS reconfirms it presence to the network periodically. As expected, steps 2 and 3 of this procedure are not executed, since in case of periodic location updates, the MS is still within the same LA, and the VLR thus has the required information about the subscriber.

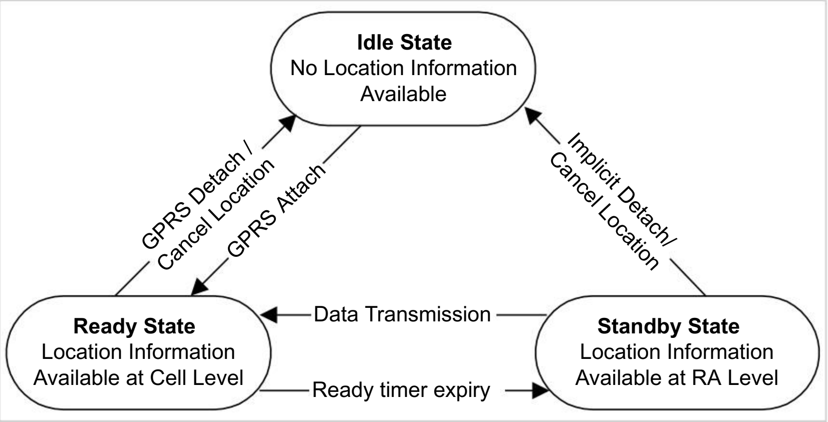
6.3 State Machine for GPRS Mobility Management

GMM state model involves a three-state FSM, which is similar to the GSM MM state model described in Section 6.2.1. However, the names of the GPRS states as well as certain transitions across states are slightly different from those in the GSM MM state model.

Figure 6.3 depicts the three-state GMM state model. The following states are depicted in the state machine:

FIGURE 6.3

GPRS Mobility Management State Model



- **Idle state.** In this state, the GPRS MS is not registered with the GPRS network. In other words, there is no GPRS mobility context established between the MS and the GPRS network (Serving GPRS Support Node (SGSN) to be precise). The GPRS network thus has no information of the subscriber and the subscriber is not reachable via the GPRS network. The GMM "idle" state is thus similar to the MM "detached" state. A transition from the "idle" state to the "ready" state happens when the subscriber registers his/her presence in the network, via the GPRS attach procedure.
- **Ready state.** In this state, the location of the subscriber is known to the network at the cell level granularity. Further, there is an ongoing session active for the subscriber (e.g., a data session like Internet access). The GMM "ready" state is thus similar to the MM "connected" state.

However, note that there is a subtle difference in the transitions followed in the MM and the GMM state models. In the GSM MM state model, a transition from the GSM "detached" state happens only to the GSM "idle" state, and never to the GSM "connected" state. However, in the GMM state model, a transition happens directly from the GPRS "idle" state to the GPRS "ready" state, when the MS attaches to the GPRS network. This difference in ideology stems from the fact that a GPRS subscriber is expected to attach to the GPRS network only when he/she has to participate in a GPRS session—in anticipation of which, the network moves the subscriber directly to the GPRS "ready" state. This is not the case with GSM, wherein a subscriber registers with the network only to be reachable, and may not necessarily do so when he/she has to initiate a voice call.

From the GPRS "ready" state, a mobile subscriber may move back to the GPRS "idle" state by explicitly detaching from the GPRS network. Similarly, from the network (or SGSN) perspective, a transition happens from the "ready" state to the "idle" state when either the subscriber detaches from the network or moves out of the SGSN coverage area (and into another SGSN area). In this latter case, the SGSN moves to the "idle" state upon receipt of a cancel location request from the HLR (see Sections 6.4.1 and 6.4.2).

- **Standby state.** In this state, the location of the subscriber is known to the network at the Routing Area (RA) level granularity. Further, there is no ongoing session active for the subscriber. The GMM "standby" state is thus similar to the MM "idle" state. A GPRS subscriber moves from the GPRS "standby" state to the GPRS "idle" state, as a result of an implicit detach (e.g., power-off of handset). Similarly, from the network (SGSN) perspective, a transition form the "standby" state to the "idle" state could happen either as a result of an implicit detach (when the subscriber is no longer reachable, e.g., due to power-off) or as a result of moving to another SGSN area. In the latter case, the SGSN moves to the "idle" state upon receipt of a cancel location request from the HLR (see Sections 6.4.1 and 6.4.2).

6.4 GPRS Mobility Management Procedures

Chapter 2 provided a brief overview of the GMM procedures. This section details the MM procedures for a GPRS subscriber, which differs slightly from those for a GSM subscriber. GMM procedures for the GPRS domain are discussed in [3GPP TS 23.060] specification.

For a GPRS subscriber, MM procedures involve interaction between the MS, SGSN, Gateway GPRS Support Node (GGSN), and the HLR. While the SGSN maintains the information of the current RA of the MS, the HLR maintains information of the current SGSN of the MS. Thus, the MS sends a location update (also called RA update, in case of GPRS network) to the SGSN on every RA change. The SGSN inturn updates the HLR in case of a change in SGSN area for the MS.

For interacting with entities in the external packet data networks, the GGSN acts as the gateway between the PLMN and the external networks. A Packet Data Protocol (PDP) context is maintained in the MS, the SGSN, and the GGSN. A PDP context contains information used for routing data packets between the MS and the GGSN, via the SGSN. The PDP context contains information about the PDP address assigned to the MS, the PDP type (IPv4, IPv6, X.25), Quality of Service (QoS), etc. (refer Chapter 7). On a change of SGSN area, the PDP context maintained in the GGSN also needs to be updated with the new serving SGSN information. This procedure is also a part of the GMM procedures.

An MS performs a GPRS attach before it can avail the GPRS network services. Moreover, when the MS no longer requires the services from the GPRS network, the MS may perform a GPRS-detach procedure. On receipt of the GPRS-attach message from the MS, the SGSN may initiate location updating toward the HLR if either the SGSN has changed since the last GPRS detach or if it is the first GPRS attach received from the MS. After an MS has successfully attached to the GPRS network, each update in its RA is conveyed to the SGSN through the "normal routing area update" procedure. The next few sections discuss the interaction between the MS, SGSN, GGSN, and the HLR for the different GMM procedures.

6.4.1 GPRS Attach Procedure

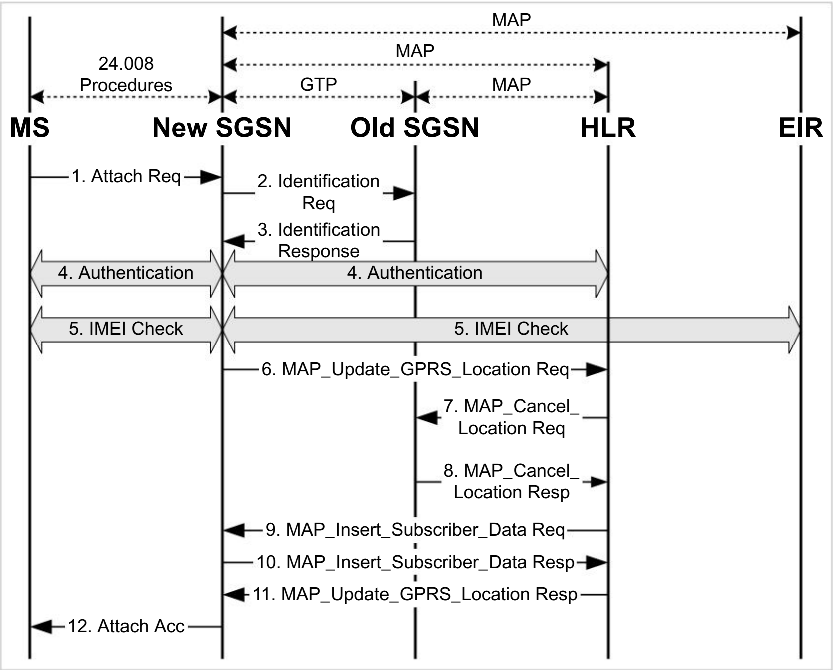
The MS uses the GPRS-attach procedure when it first attaches to the GPRS network. Using the GPRS-attach procedure, the MS provides its identity to the network. The identity provided to the network may be either the IMSI of the MS, or the P-TMSI (Packet-TMSI, similar to TMSI used in GSM networks). In case the MS provides the P-TMSI as the identity, it also includes the Routing Area Identity (RAI) associated with the P-TMSI. Since a P-TMSI is allocated to the MS by an SGSN, the RAI is the identity of the RA where the MS was registered when it was allocated this P-TMSI.

The following steps are followed for the GPRS-attach Procedure (see Fig. 6.4):

1. The MS sends the "attach request" message to the SGSN, which includes the identity of the MS.
2. In case the MS has identified itself by the P-TMSI, the SGSN first checks if the MS was previously registered with it or not. If not, the SGSN contacts the previous SGSN of the MS to request for the IMSI of the subscriber. The new SGSN in this case includes the P-TMSI of the subscriber in the "identification" request message sent to the old SGSN.
3. The old SGSN sends back the IMSI for the subscriber to the new SGSN using the "identification" response message. In case both the old and the new SGSN do not understand the P-TMSI, the MS is requested by the new SGSN to provide its identity, the IMSI. This is a part of the GPRS identification procedure, discussed separately in Section 6.4.6.
4. Authentication procedures are next followed to authenticate the MS. GPRS authentication procedures are separately discussed in Section 6.4.7.

Figure 6.4

Sequence Diagram for the "GPRS-Attach" Procedure



5. Optionally, the equipment of the MS (identified by the International Mobile Equipment Identify (IMEI)) is checked to verify that the equipment is not black-or gray-listed.
6. Subsequently, the SGSN checks if the MS had performed the GPRS detach in its region or not. In case the SGSN has changed since the MS performed the detach, or in case this is the first "GPRS attach" request from the MS, the SGSN informs the HLR about the location of the MS by sending a "MAP_Update_GPRS_Location" request message. The location update message to the HLR contains the SGSN number (an E.164 number), the SGSN address (IP address), and the IMSI of the MS.
7. The HLR initiates a "MAP_Cancel_Location" request message toward the previous SGSN, if any, where the MS was last registered. This is an indication to the previous SGSN that the MS has moved to a different SGSN, and the previous SGSN can purge/delete all data stored for the subscriber.
8. The old SGSN deletes subscriber records for the MS and acknowledges by sending a "MAP_Cancel_Location" response message to the HLR.
9. The HLR then sends the relevant subscriber data to the current SGSN through one or more "MAP_Insert_Subscriber_Data" request messages. The current scenario assumes that only one MAP message is required to transfer the complete subscriber data.
10. The new SGSN updates its datastore with the subscriber data and sends a "MAP_Insert_Subscriber_Data" response message to the HLR.
11. On successful completion of subscriber data transfer, the HLR sends a "MAP_Update_GPRS_Location" response to the new SGSN, indicating successful registration of the location information of the MS at the HLR.
12. The SGSN, in turn, indicates successful registration of location information to the MS by sending an "attach accept" message to the MS. This completes the GPRS attach Procedure.

6.4.2 GPRS Detach Procedure

Similar to the GPRS-attach procedure discussed in Section 6.4.1, the MS can perform the GPRS detach procedure when it no longer wishes to access the GPRS services. Once it is GPRS detached, the MS becomes inactive in the GPRS network. The GPRS detach procedure, when initiated by the MS, is also called the "explicit GPRS detach" procedure. This is so because in case of an explicit GPRS detach, signaling messages are exchanged between the MS and the network, just as in the case of a GPRS attach.

At times, the network may itself detach the MS from the GPRS network, without any notification to the MS. This is also called implicit GPRS detach (refer Section 6.4.9).

6.4.3 Normal Routing Area Update Procedure

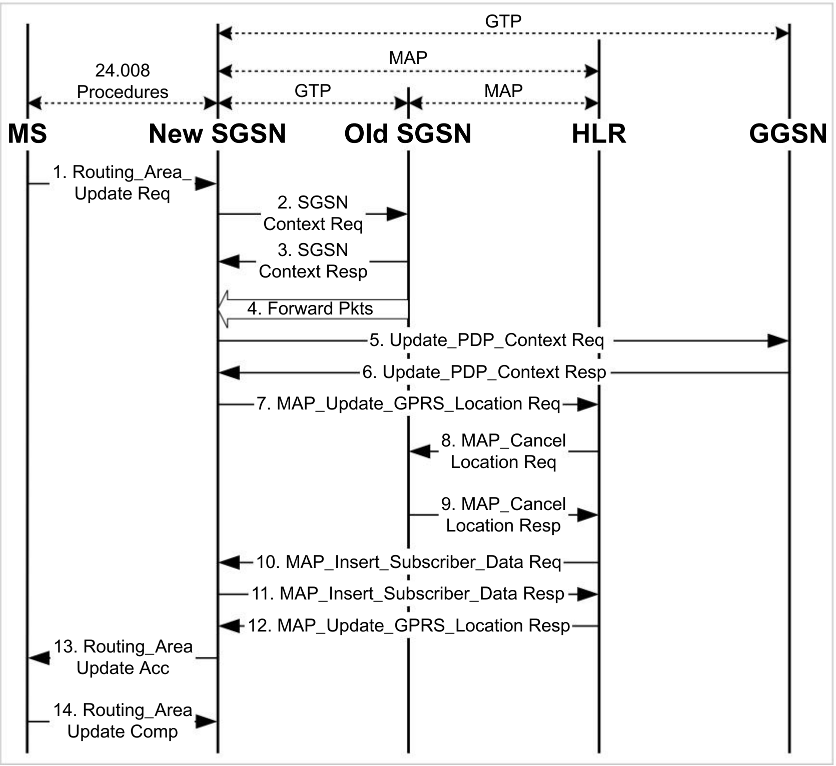
The normal RA update procedure is executed each time the MS moves into a new RA. On receipt of the RA update message, the SGSN checks if the RA or the SGSN of the MS has changed since the last update. In case the RA has changed, the SGSN modifies its information to reflect the current RA of the MS. In case the SGSN has changed since the last update, i.e., this SGSN has received an update from the MS for the first time, the SGSN also informs the HLR and the GGSN that the SGSN of the MS has changed.

In case the RA of the MS has changed but the SGSN area is still the same, the SGSN simply updates its current RA information stored for the MS. There is no need to inform either the GGSN or the HLR of the MS. In case of change in SGSN area, however, the GGSN and the HLR need to be notified. The GGSN routes incoming packets for the MS through the SGSN, and hence updation of the SGSN needs to be notified to the GGSN. The sequence of steps followed during normal RA update in case the SGSN has changed is as follows (see Fig. 6.5):

1. The MS sends the “Routing_Area_Update_Request” message to the SGSN with the old RAI and the old P-TMSI.
2. The SGSN checks if it has received a request from the MS for the first time, i.e., the SGSN has changed since the last update. Here, since this is the case, the SGSN requests the old SGSN of the MS for information about the MS, by sending a “SGSN_Context” request message. The information requested includes information about the PDP contexts of the MS.
3. The old SGSN responds by sending the PDP context information for the MS in the “SGSN_Context” response message.

FIGURE 6.5

Sequence Diagram for Inter-SGSN RA Update



4. Further, the old SGSN duplicates the buffered messages that it has received for the MS, and starts tunneling them to the new SGSN. This is done to ensure that the messages for the MS are not lost during the location update between the two SGSNs.
5. The new SGSN sends an "Update_PDP_Context" request to the GGSN to update its PDP context with the new SGSN address.
6. The GGSN acknowledges the change to the PDP context by sending an "Update_PDP_Context" response message to the new SGSN.
7. Next, the SGSN informs the HLR of the change in the location (SGSN) of the MS by sending a "MAP_Update_GPRS_Location" request message. The location update message to the HLR contains the new SGSN number (an E.164 number), the SGSN address (IP address), and the IMSI of the MS.
8. The HLR initiates a "MAP_Cancel_Location" request message toward the previous SGSN, if any, where the MS was last registered. This is an indication to the previous SGSN that the MS has moved to a different SGSN, and the previous SGSN can purge/delete all data stored for the subscriber.
9. The previous SGSN acknowledges deletion of the subscriber data by sending a "MAP_Cancel_Location" response message to the HLR.
10. The HLR then sends the relevant subscriber data to the current SGSN through one or more "MAP_Insert_Subscriber_Data" request messages. The current scenario assumes that only one MAP message is required to transfer the complete subscriber data.
11. The SGSN acknowledges receipt of the subscriber data by sending a "MAP_Insert_Subscriber_Data" response message to the HLR.
12. On successful completion of subscriber data transfer, the HLR sends a "MAP_Update_GPRS_Location" response message to the SGSN, indicating successful registration of the location information of the MS at the HLR.
13. The SGSN then sends the "Routing_Area_Update_Accept" message to the MS with the new P-TMSI.
14. The MS confirms the reallocation of the P-TMSI by sending a "Routing Area Update Complete" message back to the SGSN.

This completes the normal RA update procedure.

6.4.4 Periodic Routing Area Update Procedure

Periodic RA updates are sent by the MS to the SGSN to confirm its presence in the same SGSN area, even if there is no change in its RA information. This is similar to the periodic location updates sent by an MS to the MSC/VLR in a GSM network.

The periodic RA update procedure is similar to the normal RA update procedure discussed in Section 6.4.3. The only difference in case of the periodic RA update is that in all cases, the SGSN only updates its own information for the MS. There is no need to inform either the GGSN or the HLR of the MS update. This is similar to the case of normal RA update, where the SGSN area of the MS has not changed.

6.4.5 P-TMSI Reallocation Procedure

The P-TMSI reallocation procedure is used to ensure subscriber identity confidentiality. To provide the same, it is required that the IMSI is not made available or disclosed to unauthorized individuals and/or entities. This provides protection against tracing the location of a mobile subscriber by listening to the signaling exchanges on the radio interface for the IMSI. Thus, the temporary identity P-TMSI is used instead of the IMSI on the radio interface, as part of the signaling messages exchanged between the MS and the network.

However, it is still possible that some unauthorized individual or entity is able to derive the IMSI indirectly from listening to some other specific information on the radio interface. To preclude such a possibility, P-TMSI reallocation procedure is used. As a consequence of this procedure, the P-TMSI allocated to a mobile subscriber is frequently modified as a result of reallocation by the network. This helps to ensure that an unauthorized individual/entity is not able to corelate the messages transferred using different P-TMSI, as belonging to the same subscriber.

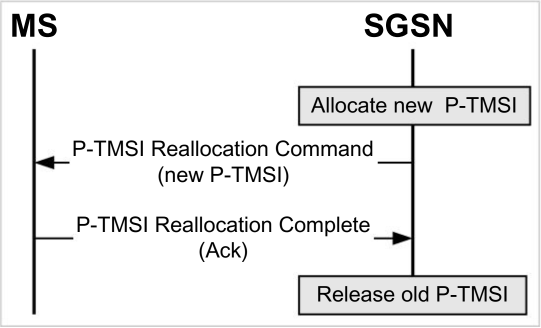
P-TMSI reallocation procedure can be executed either as a stand-alone procedure or as a part of other procedures. Generally, it is recommended that P-TMSI reallocation be done whenever there is a change in the RA of the MS. This results in the network sending the new P-TMSI to the MS as part of the RA update procedure. As a stand-alone procedure, P-TMSI reallocation is depicted in Fig. 6.6. In this procedure, the new P-TMSI allocated to the MS is sent as a part of a separate P-TMSI reallocation command message. Once the MS has acknowledged the new P-TMSI, the network may then delete the earlier allocated P-TMSI. More details on the P-TMSI reallocation procedure can be obtained from [3GPP TS 24.008] and [TS 23.060].

6.4.6 GPRS Identification Procedure

The GPRS identification procedure is used by the network to fetch the identity of the mobile subscriber. As mentioned in the discussion on P-TMSI reallocation, the mobile subscriber normally uses the P-TMSI to communicate with the network. However, under certain situations, the network may require that the mobile subscriber provides his/her actual

FIGURE 6.6

Stand-alone P-TMSI Reallocation Procedure

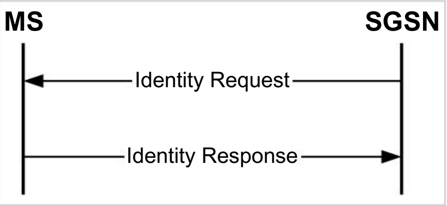


identity in the form of IMSI or IMEI. Such situations may occur in certain cases when the network does not know the P-TMSI. In such cases, the network may use the identification procedure to request the subscriber to provide its IMSI.

Figure 6.7 depicts the identification procedure used to fetch the actual identity of the MS. The network sends a request message to the MS for sending its real identity. This message includes information about the specific identity (IMSI or IMEI) that the network requires. Based upon network request, the MS indicates its IMSI or IMEI to the network in the response message. More details on the identification procedure are available in [3GPP TS 24.008] and [TS 23.060].

FIGURE 6.7

Identification Procedure



6.4.7 GPRS Authentication Procedure

The GPRS authentication procedure is similar to the authentication procedure followed in a GSM network, with the difference that in GPRS, this procedure is implemented in the SGSN entity instead of the MSC/VLR. The purpose of the authentication procedure is twofold. From the network operator perspective, the authentication procedure is used to protect the network against unauthorized use by corroborating that the subscriber is really the one as identified by the IMSI/P-TMSI provided by the subscriber. On the other hand, the authentication procedure also provides protection for the mobile subscribers by denying the possibility for intruders to impersonate authorized users.

Authentication procedure may be used at different stages of the subscriber's interaction with the network. Some of these include

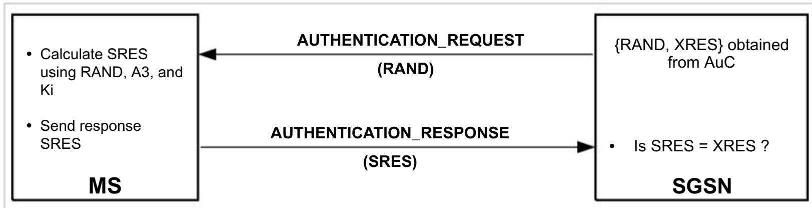
- Scenarios where subscriber-related information is intended to be changed in the HLR/VLR. This includes RA update scenario involving change of SGSN.
- Scenarios where an access is made to some network service.

Details on the authentication procedure implemented by the network and the MS are provided in [3GPP TS 43.020]. GPRS-specific details can also be obtained from [3GPP TS 23.060]. In a nutshell, the procedure involves the following sequence of steps for subscriber authentication (see Fig. 6.8):

- The network transmits a random nonpredictable number (called RAND) to the MS.
- The MS then computes the signature of RAND (called SRES) by using a publically known algorithm (called A3) and some secret information (the individual subscriber

FIGURE 6.8

Authentication Procedure



authentication key, called Ki), which is known only to the actual subscriber and the network.

- The MS then transmits the signature SRES to the network.
- The network then tests the SRES for its correctness.

Note that only the actual subscriber will be able to calculate the SRES correctly, since only the real subscriber will be aware of the secret subscriber-specific authentication key (Ki). Within the network, the function of computing the XRES using the RAND, Ki, and A3 is done at the AuC. The AuC then transmits the (RAND, XRES) pair to the SGSN, which uses this for subscriber authentication.

6.4.8 Combined RA/LA Update and Attach Procedure

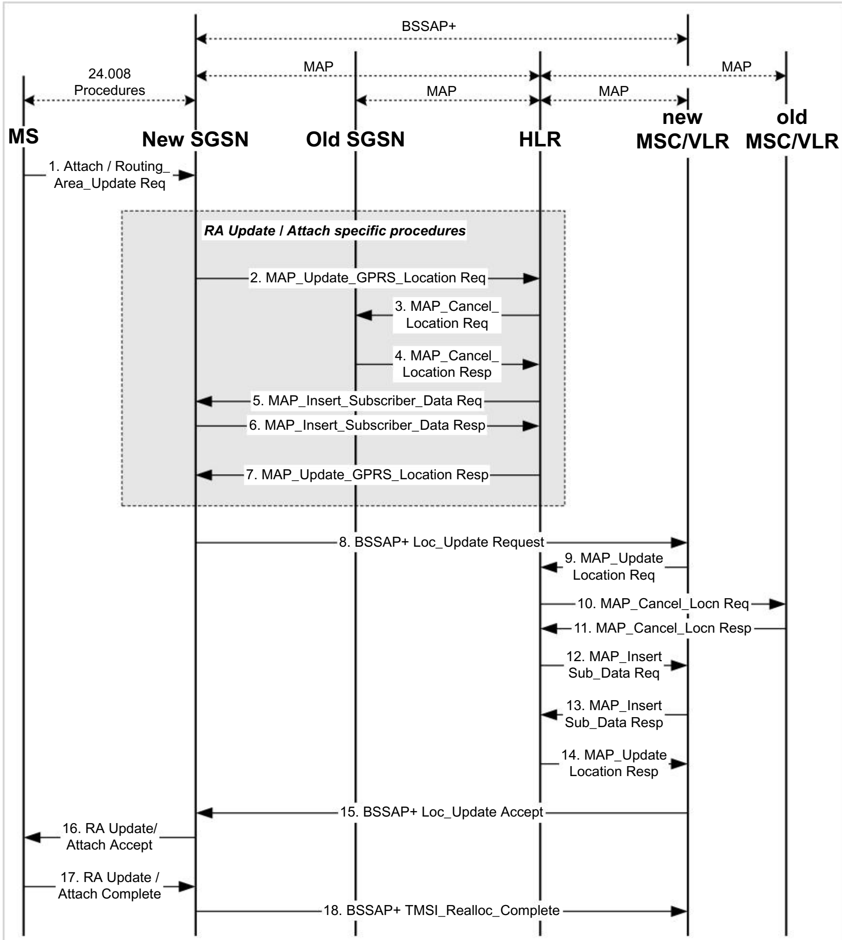
The GPRS network architecture defines an optional Gs interface between the VLR and the SGSN. The Gs interface was first discussed in Chapter 2. It is used to facilitate combined LA/RA update and combined IMSI-attach and GPRS attach by the MS. This helps in saving the radio resources when the MS is attached to both the GSM and the GPRS networks. Also, the presence of the Gs interface allows the paging of the MS for the GSM network via the SGSN. This section discusses the steps involved in a combined LA/RA update, or a combined IMSI/GPRS attach from the MS.

In case when the MS is attached to both the GSM and the GPRS networks, the LA and RA updating is done in a coordinated way to save radio resources. As discussed in Section 6.4.3, when an MS enters a new RA, it sends a "Routing_Area_Update" message to the SGSN. In case the LA of the MS has also changed (i.e., the new RA belongs to a different LA), the "Routing_Area_Update" message from the MS contains the LA update as well. The SGSN then forwards the LA update to the MSC/VLR. The MSC/VLR may return a new TMSI, which is sent to the MS by the SGSN. Similarly, the MS can perform a combined IMSI attach and a GPRS attach by sending a request to the SGSN only. The SGSN then forwards a location update request to the MSC/VLR indicating IMSI attach.

Figure 6.9 depicts the scenario of combined LA/RA update or combined "attach." Note that the messages exchanged between the SGSN and the VLR are the same, and the location update request from the SGSN to the MSC/VLR indicates whether it is a case of IMSI attach, or normal LA update. The BSSAP+ protocol messages are used between the MSC/VLR and the SGSN.

Figure 6.9

Sequence Diagram for Combined Attach/Update



6.4.9 Purge MS Procedure

Purge MS procedures in the GPRS network are defined to be similar as in the case of a GSM network. When an MS is deactivated (powered off), there are two mechanisms by which the network gets to know it. One method involves informing the network through the use of an explicit *GPRS-detach* procedure (Section 6.4.2). However, if the MS does not perform the GPRS-detach procedure when deactivated (e.g., in case when the battery of the MS was suddenly removed), an alternate mechanism is used to identify a case of potential MS deactivation, called the *Purge MS* procedure, as discussed in this section.

In case the MS is deactivated, it will no longer be able to initiate periodic RA update procedures. When the MS misses a preconfigured number of periodic RA updates, the network assumes that the MS may have been deactivated. This is also sometimes called

Implicit GPRS detach. In case the MS misses a preconfigured number of periodic RA updates, the network assumes that the MS may have been powered off or has moved out of its area. In this scenario, the SGSN sends a "MAP_Purge_MS" request toward the HLR, requesting the HLR permission to delete the subscriber data from the SGSN and to freeze the P-TMSI that was allocated to the subscriber. The sequence of steps followed at HLR on receipt of a "Purge MS" message is as follows:

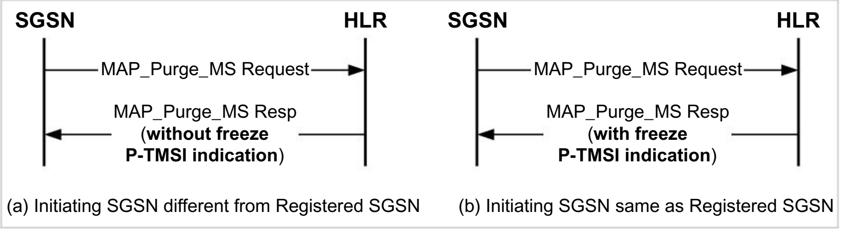
1. Check if the Purge MS message is received from the same SGSN as the one where the subscriber is currently registered. If not, then the HLR simply responds back to the SGSN with success, indicating that the SGSN may delete the subscriber data.

   This scenario is possible, for example, when the subscriber has moved into a new SGSN area, but the old SGSN could not process the "MAP_Cancel_Location" message sent by the HLR as part of the RA update scenario.
2. If the Purge MS message is received from the same SGSN as the one where the subscriber is currently registered, then the HLR marks a flag ("MS Purged for GPRS") for the subscriber in its database and sends back a positive response to the SGSN. This indicates to the SGSN that it may delete the subscriber data and freeze the P-TMSI allocated to the subscriber. The subscriber, from now on, is treated as an unreachable subscriber. The SGSN freezes the P-TMSI for some time, which is not allocated to any other subscriber immediately. This is done to safeguard against the original MS coming alive and using the P-TMSI again.

Once the "MS Purged for GPRS" flag is set for the subscriber, the flag is unset as a result of a new RA update message received from a SGSN for the subscriber. This means that the subscriber has become reachable again and the flag can be unset. Till the time the flag is unset, the subscriber is treated as being unreachable. The interaction between the SGSN and the HLR for Purge MS scenario is depicted in Fig. 6.10.

FIGURE 6.10

Purge MS Procedure



CONCLUSION

The main objective of the GMM procedures is to maintain and manage the location information for a mobile subscriber in a GPRS network. These procedures also facilitate subscribers to roam into a network operated by different service providers. The GMM procedures are best understood by means of a three-state GMM state model, which defines three states for an MS: idle, standby, and ready. The GMM state model has been discussed in detail in Section 6.3.

GMM procedures involve multiple different procedures, prime among which are the GPRS attach and the RA update procedures. The GPRS attach is performed by the MS when it first registers with the GPRS network. Once attached to the GPRS network, the MS informs of its RA change to the network using the RA update procedure. These, and some other supplementary GMM procedures, have been discussed in Section 6.4.

Further Reading

GMM procedures are an extension of the GSM MM procedures. While this chapter does provide some introductory information about GSM MM procedures, a better understanding of the same can help in better understanding of the GPRS procedures.

[GSM N. Narang] is a good introductory book for GSM. Chapter 7 of this book can be referred for more details on the GSM MM procedures.

For specific details on the GMM procedures, the following 3GPP specifications can also be referred.

[3GPP TS 24.008] is a consolidated specification that discusses all procedures on the MS-core network interface. The GMM procedures on the MS-SGSN interface can be referred to from this specification.

[3GPP TS 23.060] is the master specification for all GPRS services and procedures. GMM procedures are discussed in detail in a separate chapter of this specification.

[3GPP TS 29.002] is a consolidated specification that describes all the messages used for core network procedures. This specification can be referred to obtain more detailed information about messages used for GMM procedures within the CN.

[3GPP TS 43.020] discusses the procedures related to network security, and an annexure within this specification includes a description of the GPRS authentication procedures.

Chapter 7

Session Management Procedures

7.1 Introduction

In any voice-based network including GSM, the prevalent notion of communication is a “call,” which is essentially a communication channel to exchange voice traffic. However, as GPRS is designed for exchange of user data, the notion of calls does not apply in the GPRS network. Instead, the concept of “session” is used. A session is used to avail the services in the GPRS network. For each session, a Packet Data Protocol (PDP) context is created, which defines the characteristics of that session. The PDP context information is maintained at the Mobile Station (MS), Serving GPRS Support Node (SGSN), and Gateway GPRS Support Node (GGSN). The management of PDP context, which includes the procedures necessary to activate, modify, and deactivate a PDP context between the MS, SGSN, and GGSN, comes under the Session Management (SM) functionality.

To explain the SM procedures, the information in this chapter is organized as follows: first, some of the basic concepts are explained. These include addressing, routing, encapsulation, decapsulation, tunneling, and packet filtering. These concepts form the basis of all SM procedures. Thereafter, the PDP activation procedure is explained. The description includes the MS-initiated activation procedure, network-requested activation procedure, and the secondary PDP context activation procedure. After this, the PDP context modification procedure, which can be initiated by various entities including MS, SGSN, and GGSN, is discussed. Only the important procedures are explained. The PDP context deactivation procedure is discussed next. Like the PDP context modification, the deactivation procedure can also be initiated by different entities. Only the important scenarios are discussed here.

7.2 Session Management Concepts

Before delving into specifics of SM procedures, it is important to explain some of the elementary concepts that are central to understanding these procedures. These concepts are as follows:

- Addressing
- PDP context
- Packet routing
- Tunneling and encapsulation
- Packet filtering

Each of the above concepts is explained next.

7.2.1 Addressing

The GPRS network enables the MS to communicate with entities of a Packet Data Network (PDN). In order to engage in such a communication, the MS must have an address that is valid in the PDN. It may be noted that the PDN lies outside the Public Land Mobile Number (PLMN); this implies that the addresses in the PLMN (like International Mobile Subscriber Identify (IMSI) or Mobile Subscriber ISDN Number (MSISDN)) are alone not sufficient for communication with the PDN entities. Since the most common PDN is based on the Internet Protocol (IP), an MS must have an IP address for communicating with other entities in the IP network. The IP address may be an IPv4 or an IPv6 address. Whatever be its nature, the MS must have an address, called the Packet Data Protocol (PDP) address, to communicate with entities in a PDN.

The MS can obtain the PDP address in different ways. The simplest way is to assign a "static PDP address" to it, which is allocated by the network operator of the Home PLMN (HPLMN). Since the allocation is static, the address is permanent in nature.

However, network addresses are scarce resources. It does not make sense to allocate an address permanently, more so when a subscriber may not be using it all the time. Thus, the addresses are generally allocated "dynamically" so that a small set of these addresses can be shared between many subscribers. There are three different ways in which an MS obtains a dynamic PDP address:

- The Home PLMN (HPLMN) operator assigns PDP address to the MS
- The Visited PLMN (VPLMN) operator assigns PDP address to it
- The operator of the PDN assigns a dynamic PDP address to the MS

As part of the subscription information, the HPLMN operator defines how the PDP addresses are assigned. Further, when the PDP address is assigned by the HPLMN or the VPLMN, it is the responsibility of the GGSN to allocate and release this address.

7.2.2 PDP Context

PDP address leads to the notion of "PDP context" because dynamic PDP address is allocated by the GGSN during the activation of PDP context. A PDP context can be viewed as a set of information maintained by the MS, SGSN, and GGSN. It contains a PDP type (that identifies the type of PDN, for example, IPv4), the PDP address (say a dynamically allocated IPv4 address), Quality of Service (QoS) information, and other session information (see Table 7.1). Activating a PDP context refers to creating the PDP context at the MS, SGSN, and GGSN so that the MS can communicate with an entity in PDN using the PDP address maintained in the PDP context. When the communication is over, the PDP context is deactivated. During the deactivation of PDP context, the dynamic PDP address is released. In case the PDP address is assigned by an external PDN, there are again two ways to achieve it. First, the GGSN obtains an address from the external PDP using protocols like Dynamic Host Configuration Protocol (DHCP) and provides it to the MS during the activation of the PDP context. Alternatively, the MS may obtain the PDP address directly from the PDN after the PDP context is activated.

Table 7.1

Elements of PDP Context

PDP Type	The PDP type is identified by the organization that is responsible for PDP type (e.g., European Telecommunications Standered Institute (ETSI) or Internet Engineering Task Force (IETF) and a PDP type number Point-to-Point Protocol (PPP), IPv4, or IPv6).
PDP Address	This contains the PDP address whose format is governed by the PDP type. There is a length field that defines the PDP address length.
Network Service Access Point Identifier (NSAPI)	This field holds the index of the PDP context. In the MS, this field is used to identify a PDP Service Access Point (SAP). In the SGSN/GGSN, the NSAPI is used to identify the PDP context associated with a Mobility Management (MM) connection.
QoS	This defines the QoS for a PDP context. There are three types of QoS: • **Subscribed QoS.** This is the QoS profile maintained at the Home Location Register (HLR). A subscriber cannot request QoS greater than the value subscribed for. • **Requested QoS.** This is the QoS profile requested by the MS at the beginning of a session. • **Negotiated QoS.** While a user provides the requested QoS, the network negotiates each QoS attribute to a level that is in accordance with the available network resources.

Contd.

TABLE 7.1

Contd.

QoS Parameters	The important parameters included in the QoS profile are as follows: • **Traffic class.** This parameter defines the nature of traffic. Four traffic classes are defined, namely conversational class, streaming class, interactive class, and background class. The first two categories of traffic class are intended for real-time traffic flows. The last two categories of traffic class are intended for traditional Internet applications (like E-mail and WWW). • **Maximum and guaranteed bit rate.** These set of parameters define the maximum and guaranteed number of bits delivered by and to the access network within a period of time, divided by the duration of the period. A traffic is said to be conformant if it does not exceed the maximum and guaranteed bit rate. • **Delivery order.** This parameter defines whether the bearer shall provide in-sequence Service Data Unit (SDU) delivery or not. • **Maximum SDU size.** This parameter defines the maximum size of SDU that can be carried by the network. • **SDU error ratio.** It indicates the fraction of SDUs lost or detected as erroneous. SDU error ratio is defined only for conforming traffic. • **Residual bit error ratio.** This parameter defines the undetected bit error ratio in the delivered SDUs. • **Delivery of erroneous SDU.** This parameter defines whether erroneous SDUs are to be delivered to the application layer or not. • **Transfer delay.** It indicates maximum delay in milliseconds for 95% of delivered SDUs during the lifetime of a bearer service. • **Traffic handling priority.** This parameter specifies the relative importance for handling of all SDUs belonging to one PDP context compared to the SDUs of other PDP context.
Tunnel Endpoint Identifier (TEID)	This field is used between SGSN and GGSN to identify a tunnel endpoint in the receiving GPRS Tunneling Protocol-Control Plane (GTP-C) or GPRS Tunneling Protocol-User Plane (GTP-U) endpoint and to identify a PDP context.

Contd.

Table 7.1

Contd.

Access Point Name (APN)	An APN is used to access a service associated with a GGSN. The name is translated by SGSN using Domain Name System (DNS) to obtain the IP address of GGSN that can provide the requested service. An APN consists of an "APN network identifier" which is mandatory and an "APN operational identifier" which is optional. The APN network identifier is a Fully Qualified Domain Name (FQDN) (e.g., "service.company.com"). The APN network identifier is used to identity which external network the GGSN is connected to and optionally a requested service by the MS. The APN operational identifier is also an FQDN ending with ".gprs." The APN operator identifier is used to identify which PLMN GPRS backbone the GGSN is located.

Irrespective of whether the PDP address is dynamic or static, it is ready for use only after the PDP context is activated. During the activation of the PDP context, the GGSN creates a mapping between the IMSI and the PDP address. The GGSN also keeps other information (e.g., QoS, Tunnel Endpoint Identfier (TEID), etc.) in the PDP context. When the GGSN receives a packet from an external network, it uses the PDP context (that has the TEID-PDP address mapping) to route the packet to the MS. The SGSN also keeps the PDP context information to route packets from the Packet Control Unit (PCU) to GGSN and vice-versa.

The PDP context is always activated by the MS. Even if a PDP context does not exist and GGSN receives a packet with a static PDP address, it requests the MS to activate the PDP context. This is done through the "network-requested PDP context activation" procedure. Upon receiving the request from the network, the MS activates the PDP context.

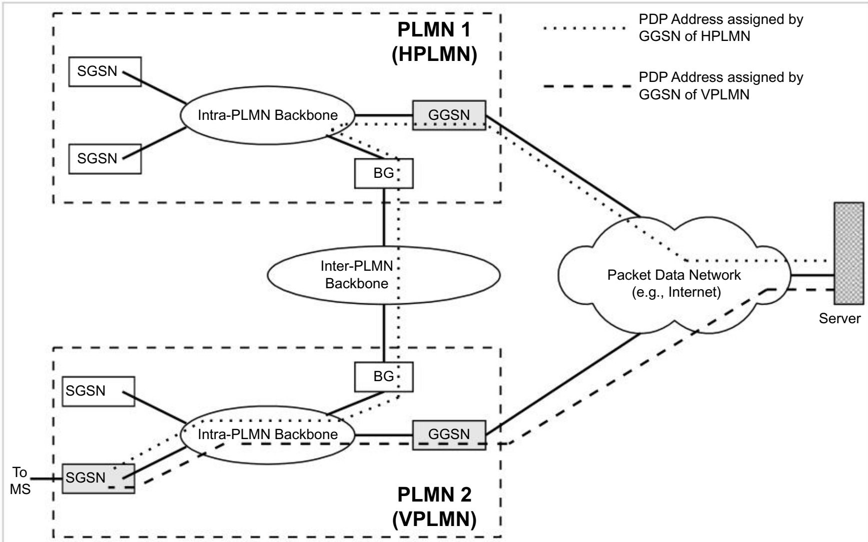
While the PDP context is always activated by the MS, it can be deactivated either by the MS or by the network (i.e., SGSN or GGSN). A PDP context is deactivated when data transfer is over and the PDP address is not required any more. Apart from activation and deactivation, there are also procedures to modify the attributes of the PDP context (e.g., QoS). The PDP context activation, modification, and deactivation procedures are explained in Sections 7.4, 7.5 and 7.6, respectively.

7.2.3 Packet Routing

After a PDP context is activated, the PDP address is available for data transfer. The PDP PDUs are routed (i.e., tunneled) between the SGSN and GGSN over the GTP-U/UDP/IP protocols. The exact path depends upon whether the MS is in HPLMN or VPLMN and whether the address is allocated by the GGSN of HPLMN or GGSN of VPLMN. Figure 7.1 depicts an example where the MS is roaming in a VPLMN. Two scenarios are shown in this figure. These are explained as follows.

FIGURE 7.1

Example of Packet Routing in GPRS Network



In case the GGSN of HPLMN assigns a PDP address to the MS, the packets are routed from the SGSN of the VPLMN to the GGSN of HPLMN. During the activation of the PDP context, the SGSN stores the address of the GGSN that provides connectivity to the external PDN. This address is used to tunnel the packets to the GGSN using the GTP-U. Since the exchange takes place between GSNs of two different PLMNs, an inter-PLMN backbone is used for communication. Further, the Border Gateway (BG) of the respective PLMN performs security functions and prevents unauthorized access. In reverse direction, when the GGSN of HPLMN receives a packet with a PDP address, it uses the PDP context to determine the SGSN to which this packet is to be forwarded. Note that as the PDP address is allocated by the GGSN of HPLMN, all packets destined for that PDP address are directed to that GGSN. The GGSN of HPLMN then takes further action.

It is also possible that the GGSN of VPLMN allocates a PDP address. Whether this is allowed or not is known by the SGSN when it queries the subscriber information from HLR. In case this is possible, the GGSN of VPLMN assigns a dynamic PDP address. All packets are then routed over the intra-PLMN backbone.

7.2.4 Encapsulation and Tunneling

The twin concepts of "encapsulation" and "tunneling" are central to the way data is exchanged in the GPRS network. These concepts are explained next:

Encapsulation is formally defined as "the addition of address and control information to a data unit for routing packets within and between the network(s)". In the context of SM, encapsulation enables packets of any PDP to be exchanged between the MS and the PDN. Encapsulation makes the PLMN transparent to the data protocols carried by it. This provides it the flexibility to transport packets of any protocol across the PLMN, without having to define protocol-specific mechanisms for such a transport. The encapsulation function is performed by the entity sending an encapsulated packet. The receiving entity decapsulates the packet (i.e., removes the address and control information) and obtains the original packet.

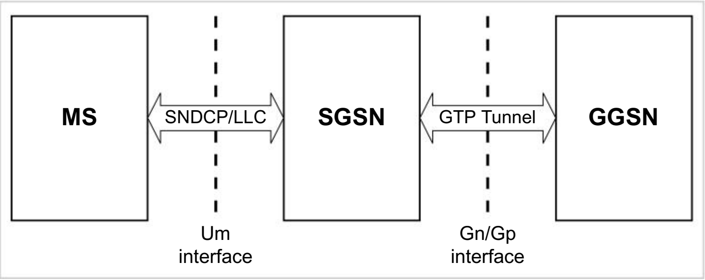
In the GPRS network, the PDP Protocol Data Unit (PDU) is encapsulated/decapsulated at the MS, SGSN, and GGSN. Between the MS and SGSN, the PDP packet is encapsulated using Subnetwork-Dependent Convergence Protocol (SNDCP). Between SGSN and GGSN, the PDP PDU is encapsulated using the GPRS Tunelling Protocol-User Plane (GTP-U).

Along with encapsulation, tunneling is used to exchange messages between the MS and GGSN. Formally, tunneling is defined as "the transfer of encapsulated data units within and between the network(s) from the point of encapsulation to the point of decapsulation." The tunnel is a two-way, point-to-point logical path used for information exchange.

In the GPRS network, the tunnels (referred to as GTP-U and GTP-C tunnels) are identified using Tunnel Endpoint Identifier (TEID). The GPRS Tunneling Protocol-Control Plane (GTP-C) is used for establishing GTP-U tunnels and exchanging TEID. The GTP-U tunnel exists between SGSN and GGSN. This is depicted in Fig. 7.2. Between SSGN and GGSN, the TEID is exchanged as part of PDP context activation procedure. The TEID is always assigned by the downlink entity. Thus, in the uplink direction (i.e., SGSN to GGSN), the TEID is allocated by GGSN. In the downlink direction (i.e., GGSN to SGSN), it is allocated by SGSN. The TEID is sent with every message as part of the GTP header. A PDP PDU is encapsulated along with a GTP header. This encapsulated PDU, called the GTP PDU, is carried as a User Datagram Protocol (UDP) PDU, which in turn is carried as an IP PDU. The TEID of the GTP header, along with the GSN address and port number, is used by the receiving entity to identify the GTP tunnel to which the PDU belongs. This information then gives the PDP context with which the PDP PDU is associated. The packet is then handled as per the parameters of the PDP context.

Figure 7.2

Packet Encapsulation in GPRS Network

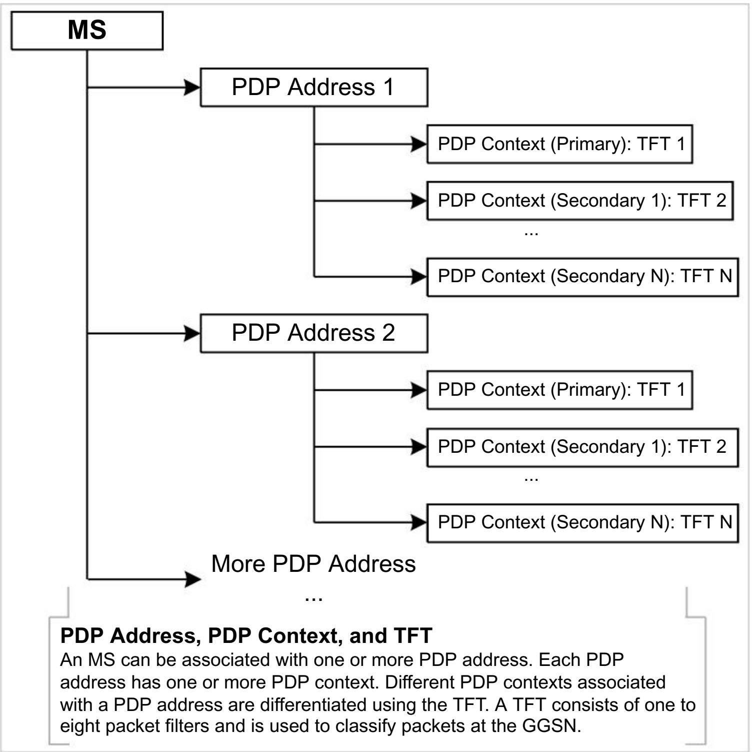


7.2.5 Packet Filtering

It is possible that a PDP address is associated with multiple PDP contexts (refer Fig. 7.3). For example, when the MS is participating in many application layer sessions, it will be getting

FIGURE 7.3

Relation Between PDP Address, PDP Context, and Traffic Flow Template



different flows for different applications, all with the same PDP address. Under such circumstances, the GGSN routes downlink PDUs based on the Traffic Flow Template (TFT) assigned to different PDP contexts. The TFT is used to classify packets at the GGSN and police incoming packets for downlink direction (i.e., GGSN to MS).

A TFT consists of one to eight packet filters, each uniquely identified by a packet filter identifier. A packet filter also has an evaluation precedence (similar to priority) that is unique within all TFTs associated with the PDP context that share the same PDP address. Besides the packet filter identifier and evaluation precedence, there is at least one attribute (used for filtering) from among the following:

- Source address and subnet mask
- Protocol number (IPv4)/next header (IPv6)
- Destination port range
- Source port range
- IPSec Security Parameter Index (SPI)
- Type of Service (TOS) for IPv4/traffic class and mask for IPv6
- Flow label for IPv6

Based on the settings of the packet filter attributes and the contents of the received PDP PDU, the GGSN does packet filtering and classification. Upon receiving a PDP PDU, the GGSN uses the packet filters with the smallest evaluation precedence to find a match. If no match is found, the packet filter with the next smallest evaluation precedence is inspected. When a match is found, the PDP PDU is associated with the corresponding PDP context. In case no match is found, the PDP context without any TFT is used to associate the PDP PDU. In case all PDP contexts have a TFT and no match is found, the PDP PDU is discarded.

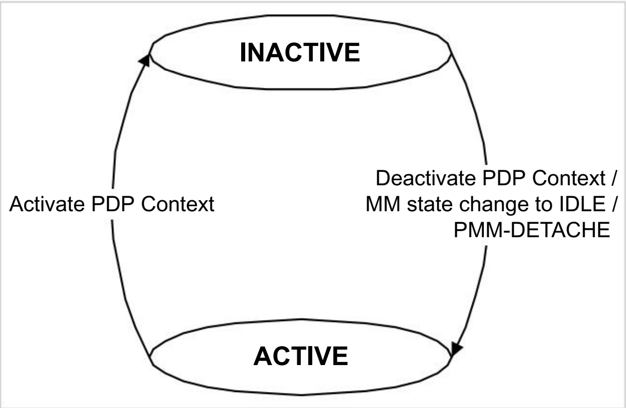
7.3 PDP Protocol States

To manage the PDP addresses, the MS, SGSN, and GGSN implement a PDP-state machine. Figure 7.4 depicts a simplified view of the PDP-state machine with two PDP protocol states (INACTIVE and ACTIVE) and the state transitions between these two states. Apart from these two states, there are intermediate states (like PDP-INACTIVE-PENDING and PDP-ACTIVE-PENDING). However, for the sake of simplicity, these intermediate states are not shown in the figure. The PDP states are explained as follows:

- **INACTIVE state.** This state indicates that the PDP address is currently not available for use (i.e., PDP context for the PDP address is not active). In this state, the PDP data cannot be exchanged by the MS. In case the GGSN receives a PDP PDU for a PDP address that is in an INACTIVE state, it may initiate a network-requested PDP context activation (see Section 7.4.3). This procedure applies to static PDP addresses. If GGSN is unable to serve the request (e.g., when it does not have information about the PDP address), it returns an error (e.g., Internet Control Message Protocol (ICMP) unreachable error). An MS makes transition to the ACTIVE state by activating the PDP context.
- **ACTIVE state.** In this state, the PDP address is available for data transfer. The GGSN contains a PDP context for the PDP address in ACTIVE state. The PDP context contains the necessary information to route a PDP PDU from and to the MS.

Figure 7.4

PDP Protocol States



7.4 PDP Context Activation Procedures

In order to communicate with PDNs, the MS activates a PDP context. The MS may initiate the PDP context activation procedure by itself or it may do so after receiving a request from the network. A PDP context applies to a specific QoS and NSAPI. A given PDP address can be associated with one or more PDP contexts. In case more than one PDP context exist for a given PDP address, the first PDP context is established using the "PDP context activation procedure" while the additional contexts are activated using the "secondary PDP context activation procedure." The secondary PDP context may be used to activate a PDP context while reusing the PDP address and other PDP context information from an already active PDP context, but with a different QoS profile. Both normal and secondary PDP context activation procedures are explained in this section. Further, if more than one PDP context exist for a PDP address, there is a TFT for the additional contexts. The TFT is sent transparently by the MS to the GGSN. As mentioned earlier, TFT is used for packet filtering and classification at GGSN.

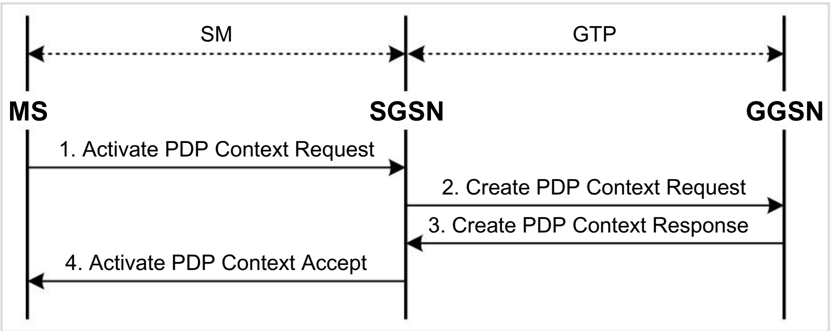
7.4.1 PDP Context Activation Procedure

The PDP context can be activated by the MS provided the MM connection exists. Figure 7.5 shows the messages exchanged during the PDP context activation procedure. The steps involved in this procedure are summarized as follows:

1. In order to request PDP context activation, the MS sends "activate PDP context request" message to the SGSN. The message contains the Network Service Access Point Identifier (NSAPI), PDP type, PDP address, Access Point Name (APN), and the requested QoS. The PDP address is sent only if the MS has a static address. In case the MS wants a dynamic address, the address field is left empty. The MS may also specify the APN to facilitate selection of a particular service or GGSN.
2. Upon receiving the "activate PDP context request," the SGSN validates the request. After this, the SGSN uses the APN to determine the GGSN that can handle the request. If a GGSN cannot be determined, an error response is sent to the MS. Else, the SGSN sends a "create PDP context request" to GGSN. In this message, the SGSN sends the

Figure 7.5

PDP Context Activation Procedure



information received from the MS along with a generated TEID that is to be used by the GGSN in the downlink direction. The SGSN, based on the availability of resources, may restrict the QoS. Thus, the QoS sent by SGSN to GGSN is called the "negotiated QoS" (and not the "requested QoS").

3. If the GGSN successfully processes the request, it creates an entry in the PDP context table. This entry allows the routing of packets between the SGSN and the external PDN. GGSN then creates the "create PDP context response" message and sends it to SGSN. The response message has a PDP address if it is dynamically allocated by the GGSN. In case the address is to be obtained from an external PDN, then a value of 0.0.0.0 is set to indicate that the MS shall get a PDP address after the completion of the PDP context activation procedure. The GGSN also sends a TEID to be used by SGSN in uplink direction.
4. The SGSN sends "activate PDP context accept" message with NSAPI, QoS negotiated, and other parameters to the MS. The procedure is complete when the MS receives this message. Depending upon the nature of the application, different QoS profiles can be requested by the MS. For example, for E-mail the QoS profile indicates delay-insensitive application. Similarly for real-time applications, the QoS profile indicates so. In case the network cannot provide the requested QoS profile, it can negotiate the QoS profile to a value close to the requested one. The MS has the option to accept the negotiated profile or deactivate the PDP context.

Once the PDP context activation procedure as described here is complete, the SGSN is ready to forward packets from MS to GGSN and vice-versa. Similarly, the GGSN is also ready to send and receive messages to/from MS.

7.4.2 Secondary PDP Context Activation Procedure

The secondary PDP context is activated when the MS wishes to reuse the PDP address and some other PDP context information from an active PDP context. The secondary PDP context must have a TFT. The TFT is used by the GGSN to differentiate between two PDUs, each with the same PDP address but associated with different PDP contexts. In case all previously activated PDP contexts already have a TFT associated with them, a secondary PDP context activation may choose not to send a TFT. Note that at most, only one PDP context associated with a PDP address may not have a TFT. All other PDP contexts must necessarily have a TFT.

The secondary PDP context activation procedure is similar to the steps depicted in Fig. 7.5. Some of the differences are as follows: first, the APN selection is not done. Note that since a GGSN is already selected, there is no need for the SGSN to select a GGSN that can provide a given service. Moreover, as the PDP address is already obtained, the dynamic PDP address allocation procedure by GGSN is not carried out.

7.4.3 Network-requested PDP Context Activation Procedure

As discussed in the previous subsections, the PDP context activation procedure is always initiated by the MS. However, it is possible that the GGSN may request the MS to initiate this procedure. This happens when a PDP PDU is received for a PDP address for which the GGSN of the HPLMN has some static PDP information (i.e., static PDP address). Note that even for network-requested PDP context activation, the MS actually initiates the PDP context activation procedure after receiving a request to this effect from the network.

Figure 7.6 shows the messages exchanged during the network-requested PDP context activation procedure. The steps involved in this procedure are summarized as follows:

1. The GGSN receives a PDP PDU for which there is no active PDP context. It then determines whether it has to initiate a network-requested PDP context activation procedure. This procedure is initiated when the received PDP address is a static address and the GGSN is the HPLMN of the MS under consideration. If the GGSN so decides, it sends the "send routing information for GPRS" message to HLR (using Mobile Application Port (MAP) protocol). Note that while the GGSN receives a PDP PDU containing a PDP address, it sends a request to HLR using the IMSI. For this, the GGSN maintains a mapping between the PDP address and the IMSI.
2. In case the HLR has the SGSN address for the IMSI under consideration, it sends the SGSN address in the response message to GGSN. Moreover, if the flag "MS Not Reachable for GPRS (MNRG)" is set in HLR, the HLR also sends "MS Not Reachable Reason (MNRR)" to GGSN. The MNRR indicates the reason why the MS is not reachable.
3. The GGSN takes further action based on information received from the HLR. In case the response contains the SGSN address and the MNRR is missing, the GGSN sends a "'PDU notification request" to the SGSN indicated by the HLR. The "PDU notification request" contains the IMSI, PDP type, PDP address, and APN. Even if the MNRR is present and indicates "no paging response," the GGSN still sends the "PDU notification request" to the SGSN, provided that the HLR has furnished the SGSN address. In all other cases (i.e., if the MNRR is different from "no paging response" or SGSN was not provided by HLR), the GGSN marks the MNRG flag for that MS.
4. The SGSN acknowledges the request by sending a response.

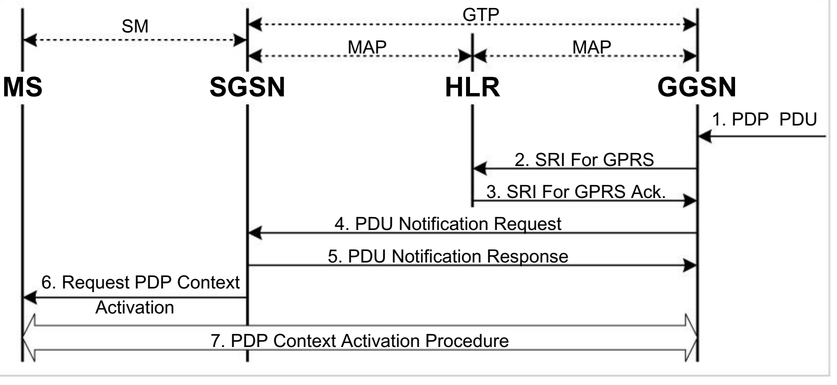


FIGURE 7.6 Network-requested PDP Context Activation Procedure

5. The SGSN also sends "request PDP context activation" message to MS with the information received from the GGSN. Before this message can be sent from the SGSN to the MS, the paging procedure is used to determine the latter's exact location.
6. If the MS accepts the invitation, it follows the PDP context activation procedure, mentioned in Section 7.4.1.

7.4.3.1 Unsuccessful Network-requested PDP Context Activation Procedure

It is possible that the network-requested PDP context activation procedure may not succeed. This may be because of an unreachable MS or due to other reason. In such a case, the specifications define different mechanisms that could ward off unnecessary enquiries to HLR. Among these, the mechanism using MNRG flag is most important and is explained as follows:

- When GGSN sends a "PDU notification request" message to SGSN and SGSN detects that the MS is GPRS-detached or if the IMSI is unknown, it send a "PDU notification response" message to the GGSN with the appropriate cause.
- The subsequent action at GGSN depends upon the cause received from SGSN. Two different course of actions are explained next:
    1. In case the cause indicated by SGSN is "MS is GPRS-detached," the GGSN sends a "failure report" message to the HLR. The HLR then sets the MNRG flag. HLR also stores the GGSN number and address. This is done to contact the GGSN in case some MS-related activity is later detected by the HLR. Now, when HLR receives another request from a different GGSN and it sees that the MNRG flag is set, the HLR sends a response containing the MNRR. The HLR also adds the number and address of this GGSN in a list of GGSNs that have to be contacted when some MS-related activity is detected. Upon receiving the response, the GGSN can avoid contacting the SGSN because the presence of MNRR indicates that the MS is not reachable for GPRS.
    2. In case the SGSN had reported "IMSI is not known," the GGSN may again query HLR requesting it to provide the SGSN address. If the SGSN address is same as obtained earlier, the GGSN sends "failure report" message to HLR. If the SGSN address is different, the GGSN again tries to perform network-requested PDP context activation procedure in the anticipation that the SGSN has changed since the previous attempt.
- Subsequently, when activity for MS is detected (e.g., when HLR receives a "ready for SM" or "update location" message), the HLR indicates this to all the GGSNs maintained for the given subscriber using "note MS present for GPRS" message. This allows the GGSN to know that the MS is reachable again and network-requested PDP context activation should be attempted. The procedure mentioned here has the advantage that a failure of network-requested PDP context activation procedure due to lack of activity is notified to the HLR. On subsequent queries, the HLR can then inform

that the MS is not reachable for GPRS. This prevents the GGSNs from making unnecessary requests to SGSN once they detect that MNRG is set. Since a GGSN may also maintain the MNRG flag, it can avoid attempting network-requested PDP context activation procedure for an MS that has the MNRG set.

7.5 PDP Context Modification Procedures

After a PDP context is activated, the SM procedures provide the means to modify the PDP context parameters. The modification procedure can be initiated by the MS, SGSN, and GGSN. The modification of subscriber data at HLR can also trigger the initiation of the PDP context modification procedure. Not all parameters of the PDP context are modifiable. Those that can be modified are

- QoS negotiated
- Radio priority
- Packet flow identifier
- PDP address (in case of GGSN-initiated modification procedure)
- TFT (in case of MS-initiated modification procedure)

The exact modification procedure followed depends upon the entity that initiates the request, and also on the nature of data modified. The following sections provide details of the PDP context modification procedure initiated by various entities.

7.5.1 MS-initiated PDP Context Modification Procedure

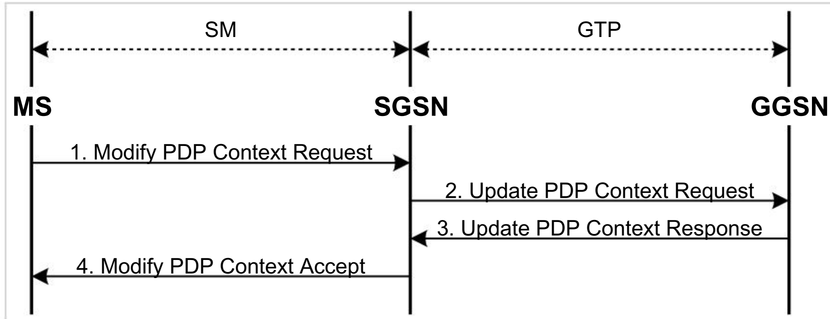
The MS can initiate the PDP context modification procedure to alter the negotiated QoS and/or TFT. The negotiated QoS indicates the desired QoS profile. The TFT can be added, modified, or deleted from a PDP context.

Figure 7.7 shows the messages exchanged during the MS-initiated PDP context modification procedure. The steps involved in this procedure are summarized as follows:

1. In order to modify the parameter(s) of a PDP context, the MS sends "modify PDP context request" message to the SGSN. The message contains the NSAPI, QoS requested, and/or TFT.
2. Upon receiving the "modify PDP context request," the SGSN validates the request and sends the "update PDP context request" to GGSN. In this message, the SGSN sends the information received from the MS.
3. The GGSN determines whether the QoS negotiated and/or TFT are compatible with the PDP context being modified. The GGSN may further restrict QoS negotiated as per its capabilities and the load of the system. The GGSN then stores the QoS negotiated. It also stores, modifies, or deletes the TFT of the PDP context as per the TFT in the request message. The GGSN then sends "update PDP context response" to SGSN.
4. The SGSN then sends "modify PDP context accept" with QoS negotiated, and other parameters to the MS.

Figure 7.7

MS-initiated PDP Context Modification Procedure



7.5.2 SGSN-initiated PDP Context Modification Procedure

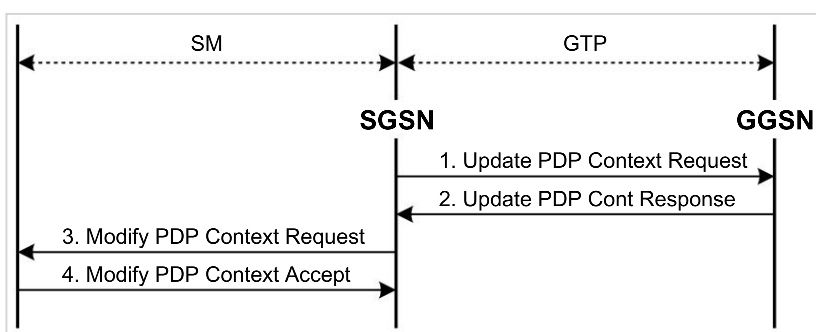
The SGSN can also initiate the PDP context modification procedure. This is done primarily to change the QoS of the PDP context.

Figure 7.8 shows the messages exchanged during the SGSN-initiated PDP context modification procedure. The steps involved in this procedure are summarized as follows:

1. In order to modify the parameter(s) of a PDP context, the SGSN sends "update PDP context request" message to the GGSN. The message contains the TEID, NSAPI, QoS negotiated, and other parameters.
2. The GGSN determines whether the QoS negotiated is compatible with the PDP context being modified. The GGSN may further restrict QoS negotiated as per its capabilities and the load of the system. It then stores the QoS negotiated and sends the "update PDP context response" to SGSN.
3. On receiving a success response, the SGSN selects the radio priority and identifies packet flow based on the negotiated QoS. The SGSN then sends the "modify PDP context request" to the MS.
4. The MS acknowledges this by sending an accept message.

Figure 7.8

SGSN-initiated PDP Context Modification Procedure



7.5.3 Other Context Modification Procedure

The GGSN can also initiate the PDP context modification procedure. This is done primarily to change the QoS of the PDP context or to change the PDP address. The procedure for this is similar to the SGSN-initiated PDP context modification procedure described earlier. The only difference is that in this case, the GGSN sends the "update PDP context request" message to the SGSN, and SGSN responds with "update PDP context request" message after the modification procedure is complete.

7.6 PDP Context Deactivation Procedures

When a PDP context is not required any more (e.g., when data exchange is complete), it is deactivated. The deactivation procedure can be initiated by the MS or by the network (i.e., SGSN or GGSN). The following section describes the PDP context deactivation procedure in detail.

7.6.1 MS-initiated PDP Context Deactivation Procedure

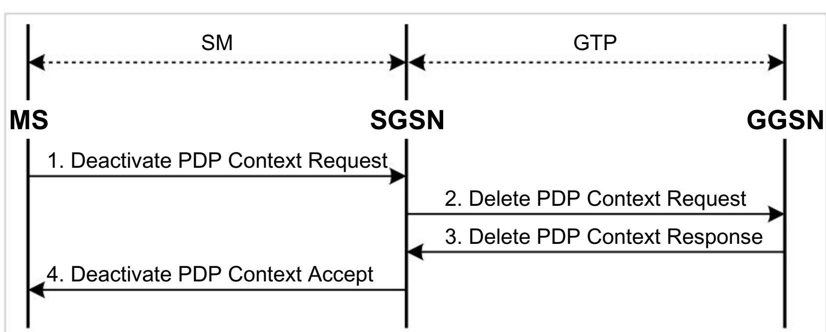
The MS carries out the MS-initiated PDP context deactivation procedure to tear down an existing PDP context when a PDP context is not required any more.

Figure 7.9 shows the messages exchanged during the MS-initiated PDP context deactivation procedure. The steps involved in this procedure are summarized as follows:

1. For PDP context deactivation, the MS sends a "deactivate PDP context request" message to the SGSN. The message may contain a tear-down indicator, which indicates whether only the PDP context under consideration is to be deactivated or all PDP context associated with the PDP address is to be deactivated.
2. Upon receiving "deactivate PDP context request," the SGSN validates the request. After this, the SGSN sends a "delete PDP context request" to GGSN. In this message, the SGSN sends the TEID, NSAPI, and the tear-down indicator.

Figure 7.9

MS-initiated PDP Context Deactivation Procedure



3. If the MS has included the tear-down indicator in the request message, the GGSN deletes all PDP context associated with the PDP address. If the GGSN has assigned a dynamic PDP address and the PDP context was the last context associated with the PDP address, then the GGSN releases this PDP address and makes it available. The GGSN then sends "delete PDP context response" to SGSN.
4. Thereafter, the SGSN sends "deactivate PDP context accept" message to the MS.

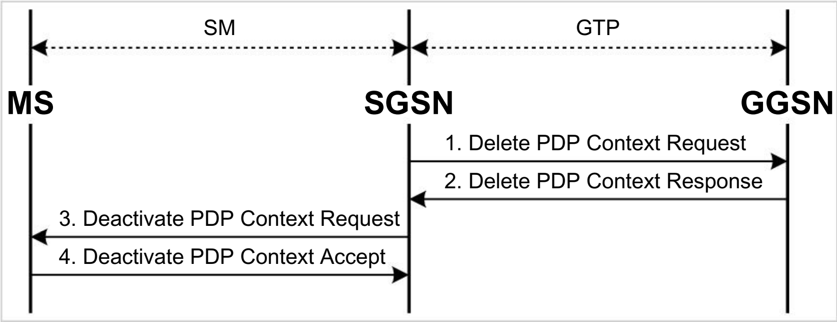
7.6.2 SGSN-initiated PDP Context Deactivation Procedure

The SGSN can also deactivate an active PDP context. Figure 7.10 shows the messages exchanged during the SGSN-initiated PDP context deactivation procedure. The steps involved in this procedure are summarized as follows:

1. The SGSN sends a "delete PDP context request" to GGSN. In this message, it sends the TEID, NSAPI, and optionally, the tear-down indicator.
2. If the SGSN has included the tear-down indicator in the request message, the GGSN deletes all PDP context associated with the PDP address. If the GGSN has assigned a dynamic PDP address and the PDP context was the last context associated with the PDP address, then the GGSN releases this PDP address and makes it available. The GGSN then sends "delete PDP context response" to SGSN.
3. Thereafter, the SGSN sends the "deactivate PDP context request" to the MS. In case the tear-down indicator is included in the request message, the MS deletes all PDP context associated with the PDP address.
4. The MS responds with "deactivate PDP context accept" message.

FIGURE 7.10

SGSN-initiated PDP Context Deactivation Procedure



7.6.3 Other Context Deactivation Procedure

Like the SGSN, the GGSN can also deactivate the PDP context. The procedure for this is quite similar to the SGSN-initiated PDP context deactivation procedure, described in Section 7.6.2.

Conclusion

This chapter explained the basics of SM, the heart of which is the concept of PDP context. A PDP context is a set of information maintained at MS, SGSN, and GGSN. The PDP context contains the PDP address, which is used to communicate with entities of a PDN. The PDP address is obtained during activation of the PDP context. The PDP context activation, modification, and deactivation procedures have been elaborated in this chapter. After PDP context is activated, the data packets are encapsulated using GTP-U and are tunneled between SGSN and GGSN. When data transfer is over, the PDP context is deactivated and the associated resources are freed.

Further Reading

The discussion in this chapter is primarily based on three 3GPP specifications. [3GPP TS 23.060] is the basic stage 2 specification for GPRS and describes the session management procedures and message flows in significant detail.

The session management part of [3GPP TS 24.008] provides details of the messages exchanged between the MS and the SGSN over the SM layer.

The [3GPP TS 29.060] provides details of the GPRS Tunelling Protocol (GTP) exchanges between SGSN and GGSN.

Chapter 8

GPRS User Plane

8.1 Introduction

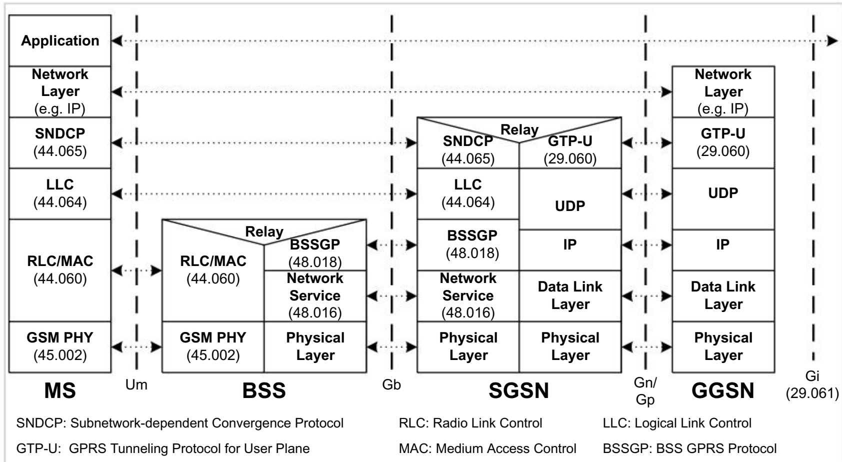
The GPRS network acts as an access network between the mobile subscriber and the external packet data network. The GPRS user plane provides a data transmission path to the external network according to the Quality of Service (QoS) parameters agreed during Packet Data Protocol (PDP) context establishment. This chapter discusses the role of GPRS in providing an end-to-end data transmission path between the mobile subscriber and the external packet data network. The complete end-to-end path can be seen as a combination of multiple smaller segments, each having its own associated user plane protocol stack (see Fig. 8.1). The first segment extends from the Mobile Station (MS) to the Serving GPRS Support Node (SGSN), and includes Logical Link Control (LLC) and Subnetwork-dependent Convergence Protocol (SNDCP) as the primary protocols. From the SGSN onward till the Gateway GPRS Support Node (GGSN) forms the next segment, where GPRS Tunneling Protocol for User Plane (GTP-U) is the prime user plane protocol. The final leg of the end-to-end path extends from the GGSN to the external packet data network. The focus of this chapter is to describe each segment of this end-to-end data transmission path, along with their associated key user plane protocols.

8.2 User Plane Between MS and SGSN

The user plane between the MS and the SGSN consists of two hops, one over the Um interface and the other over the Gb interface (see Fig. 8.2). While the air interface and Radio Link Control/Medium Access Control (RLC/MAC) protocols terminate within the Base Station Subsystem (BSS), BSS GPRS Protocol (BSSGP) and Frame Relay (FR) (used as Network Service

FIGURE 8.1

GPRS User Plane

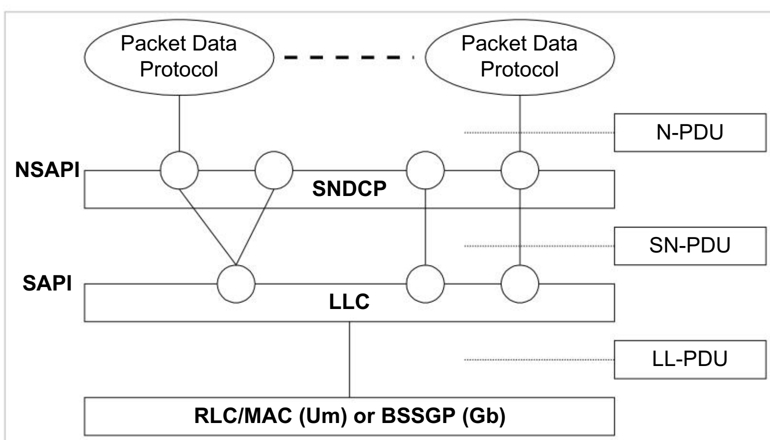


(NS) layer) carry RLC/MAC PDUs between the PCU (BSS) and the SGSN over the Gb interface. On the Gb interface, BSSGP resides over the NS layer. It uses the FR NS to provide virtual connections between the BSS and the SGSN. The BSSGP also conveys routing and QoS-related information between the BSS and the SGSN.

The intermediate hops (Um and Gb) between the MS and the SGSN are obscured by the LLC and SNDCP protocols (see Fig. 8.2), which function directly between the MS and the SGSN. The LLC provides a highly reliable link between the MS-SGSN, which is independent of the underlying radio interface. The ciphering procedure is applicable at the LLC layer. The LLC provides services to the SNDCP. It receives SNDCP PDUs (SN-PDU) over the SAPI interface, and after processing them converts them to LLC PDUs (LL-PDU), which are carried over RLC/MAC or BSSGP.

FIGURE 8.2

GPRS User Plane Between MS and SGSN



The SNDCP layer maps network-level characteristics onto the characteristics of the underlying network. This layer also performs functions like compression and decompression of user data, and segmentation and reassembly of network PDUs. It receives the network protocol PDUs (N-PDU) over the Network Service Access Point Identifier (NSAPI) interface, and after processing them, converts them to SN-PDUs, which are provided to the underlying LLC layer.

The key user plane protocols between the MS and the SGSN are discussed in more detail in the following sections.

8.2.1 SNDCP Layer

The purpose of the SNDCP layer is to map the upper layer packet data protocols (e.g., IP, PPP etc.) onto the characteristics of the underlying LLC layer. The following are some of the functions performed by the SNDCP layer:

- **Encapsulation/decapsulation of PDP N-PDUs.** The SNDCP layer provides a service interface, identified by NSAPI values, to the upper PDP layer. In one direction, the SNDCP layer receives N-PDUs from the upper PDP layer, and encapsulates it into an SN-PDU after filling the NSAPI value in the SN-PDU header. It then sends the PDU toward the LLC over the appropriate service interface Service Access Point Identifier (SAPI) provided by the LLC layer.

  In the reverse direction, the SNDCP layer receives an SN-PDU from the LLC layer, and after decapsulation of the SN-PDU, extracts and sends the N-PDU to the PDP layer on the appropriate service interface. The service interface used to send the N-PDU to the PDP layer is identified by the NSAPI value in the received SN-PDU header.

  The NSAPI and SAPI values to be used for a particular PDP are determined and established at the time of PDP context activation. These can be subsequently modified as a result of the PDP context modification procedure.

- **Segmentation and reassembly.** The SAPI service interface between the SNDCP and the LLC layer has a maximum upper limit defined for the length of the SN-PDU. This upper limit on the length of the SN-PDU is defined by a parameter "N201," which is negotiated at the time of LLC connection establishment, using LLC procedures (XID negotiation procedure). Since the N-PDU received by the SNDCP layer from the PDP layer may be of any length, one of the functions of the SNDCP layer is to segment the received N-PDUs into multiple SN-PDUs before transmission. In the reverse direction, the SNDCP layer reassembles the received SN-PDUs from the LLC layer into an N-PDU to be sent to the PDP layer.

  Two flags, "F-bit" and "M-bit" are defined in the SN-PDU header, which are used to determine the start and end of the multiple fragments of an N-PDU. The "F-bit" is used in the first segment of an N-PDU to indicate that it is the first segment. The "M-bit" is used in the last segment of an N-PDU to indicate that it is the last segment of the N-PDU. These flags are used in the receiving direction by the SNDCP layer to perform the reassembly of the SN-PDUs into a single N-PDU.

- **Compression of protocol control information.** The SNDCP layer provides an optional function to perform compression of the protocol control information carried as part of the N-PDU header. Since the bandwidth of the air interface is limited, it is beneficial to send compressed and meaningful information over the air interface. The SNDCP layer supports two mechanisms for protocol control information compression, as defined by IETF RFC 1144 and RFC 2507. While the former RFC defines the mechanism for compression of the Transmission Control Protocol/Internet Protocol (TCP/IP) header, the latter RFC defines the technique of performing compression of both the User Datagram Protocol (UDP)/IP and the TCP/IP headers.
- **User data compression.** The SNDCP layer also provides an optional function for the compression of user data carried as part of the N-PDU. If used along with the function for compression of protocol control information, then this function is applied after the protocol control information compression is done. The algorithm used within the SNDCP layer for user data compression is the V.42 bis algorithm. While the details of the user data compression algorithm are beyond the scope of this chapter, broadly, the algorithm consists of two functions—an encoder function and a decoder function. Both the encoder and decoder functions use a dictionary of code words, and use a string matching procedure to identify a mapping between a code word (compressed information) and the actual string which is transmitted/received.

For more details on the SNDCP layer, the reader is referred to [3GPP TS 44.065].

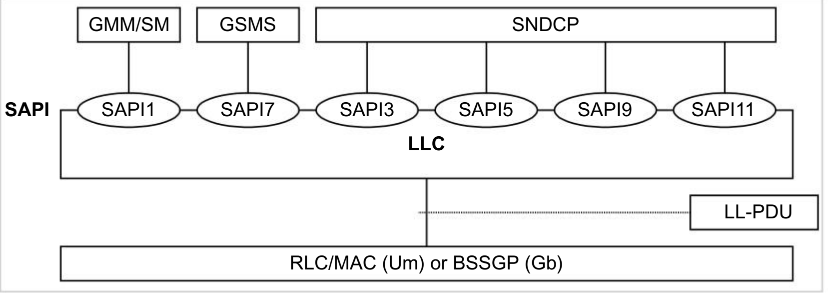
8.2.2 LLC Layer

The LLC layer is the logical link control protocol, which provides multiple logical links between the MS and the SGSN for operation of upper layer protocols. Multiple service identifiers are thus defined by the LLC layer for use by different upper layer protocols. In the GPRS control plane, LLC layer provides a service interface for control protocols GPRS Mobility Management (GMM) and Session Management (SM) (see Fig. 8.3). In the user plane, it provides a service interface to the user plane protocols GSMS and SNDCP.

Each logical link at the LLC layer is identified as a combination of two parameters, the SAPI, and the subscriber's Temporary Logical Link Identifier (TLLI). SAPI values are fixed and defined for upper layer protocols. The GMM and SM protocols both use the SAPI1. A

FIGURE 8.3

LLC Service Interfaces



protocol discriminator field present in the GMM/SM messages is then used to segregate between the messages of the two protocols, transmitted over the same SAPI. The GSMS protocol uses the service interface SAPI7. Four different service interfaces (SAPI3, SAPI5, SAPI9, SAPI11) are defined for use of the SNDCP. Each of these four service interfaces provides a specific nature of service depending upon the QoS parameters negotiated during the PDP context establishment.

While the SAPI values are fixed depending upon the upper layer protocol, the TLLI is derived from the subscriber's P-TMSI. The GMM layer controls the assignment of the TLLI value. This TLLI value is used for all communications between the LLC layer and the RLC and BSSGP layers.

The LLC layer uses similar principles as followed in the Link Access Protocol on D channel (LAPD) protocol. It provides both acknowledged and unacknowledged transmission modes. While the GMM/SM and GSMS protocols use the unacknowledged transmission mode of the LLC layer, the SNDCP layer uses both acknowledged and unacknowledged transmission modes, depending upon the QoS parameters negotiated during the PDP context establishment.

The following are the main functions of the LLC layer:

- **Data multiplexing.** The LLC layer can support multiple logical link connections for different upper layer protocols over different service interfaces. As discussed earlier, each logical link is uniquely identified using the SAPI and TLLI values. In the transmit direction, the LLC layer receives the upper layer frame and converts it to LLC frame format before transmission to the RLC or BSSGP layer. In the reverse direction, it receives a frame from the RLC or BSSGP layers and uses the SAPI value to route the frame to the correct upper layer entity.
- **Ciphering.** The LLC layer includes an optional ciphering function, which is used to cipher the information carried within the LLC frame. The ciphering algorithm and the ciphering key (Kc) are provided to the LLC layer by the GMM entity. This provides data confidentiality during transmission between the MS and the SGSN.
- **Error detection and recovery.** The LLC frame format contains a Frame Check Sequence (FCS), which is a three-octet information used for error detection. The FCS is used at the LLC layer for detecting any bit errors that may have happened in the LLC frame during transmission. In the acknowledged transmission mode, the LLC layer also provides error recovery by performing retransmission of frames that are not received (and hence, not acknowledged) by the receiver.
- **Flow Control.** The LLC layer also includes a flow control mechanism to control the flow of information between the sender and the receiver. This is particularly useful in case of a slow receiver so that the sender does not bombard the receiver with information. Explicitly defined supervisory frames (Receiver Ready (RR) and Receiver Not Ready (RNR)) are used by the receiver to indicate the sender whether it is ready to receive more information or not.

- **Window-based acknowledgment mechanism.** For acknowledged transmission mode, the LLC layer defines a window-based acknowledgment mechanism. Using this technique, a sender is able to send as many successive frames to the receiver as is the defined size of the window, without having to wait for an acknowledgment. Till an acknowledgment is received from the receiver for a transmitted frame, the sender buffers this frame in its transmission window. If the receiver does not acknowledge a frame, the sender uses this buffered frame for retransmission to the receiver.

  Clearly, the size of the transmission window at the sender is a matter of configuration. A smaller window size means that the sender has to wait often for the receiver to acknowledge the last few transmitted frames, before more frames can be transmitted. On the other hand, a larger window size requires a bigger memory at the sender side to buffer unacknowledged frames. The size of the transmission window is thus negotiated during LLC connection establishment, as part of the XID negotiation procedure. The size of the transmission window depends upon the memory available at the sending LLC entity.
- **XID negotiation.** At the end of a PDP context activation procedure, or in the scenario of an SGSN change, the LLC layer executes an XID negotiation procedure. This procedure is used to negotiate multiple SNDCP and LLC layer parameters between the MS and the SGSN side. The SNDCP layer parameters include parameters such as those used for user data compression and protocol control information compression. LLC layer parameters include parameters for defining LLC frame length, number of retries for unacknowledged frames, and the time between each retry.

For more details on the LLC layer, the reader is referred to [3GPP TS 44.064].

8.2.3 User Plane over Air Interface

While the LLC and the SNDCP operate directly between the MS and the SGSN, the SNDCP/LLC PDUs are carried over the GPRS air interface and Gb interface using other user plane protocols. Over the GPRS air interface, the RLC/MAC protocols form the key user plane protocols. In general, the RLC/MAC layer exists in both the GPRS user plane as well as the GPRS control plane. In the user plane, the RLC/MAC layer is used to transport the SNDCP PDUs encapsulated within the LLC PDUs over the air interface.

The RLC and the MAC protocols have already been covered extensively as part of Chapters 4 and 5 of this book. For more details on the use of RLC/MAC in the GPRS user plane over air interface, the reader is referred to Chapters 4 and 5 of this book.

8.2.4 User Plane over Gb Interface

While the RLC/MAC layer carries the SNCDP/LLC PDUs over the air interface, the BSSGP and the NS are used to relay the SNDCP/LLC PDUs over the Gb interface. This section describes the functions of the BSSGP and the NS in the GPRS user plane. However, before a description of the BSSGP and the NS layers, it is important to understand the addressing mechanism over the Gb interface, discussed in the next section.

8.2.4.1 Gb Interface Addressing Scheme

The Gb interface is based on the FR network, which forms the NS layer between the BSS and the SGSN. FR is a virtual circuit based networking technology and has its own addressing mechanism to identify virtual circuits. Like any other virtual circuit based networking protocol, FR virtual circuit identifiers only have local significance between two-FR nodes. End-to-end protocols operate over the NS layer, between the BSS and the SGSN, and thus define their own addressing scheme, which is used to identify communication channels end-to-end. A summary of the different identifiers used over the Gb interface is provided in Table 8.1. Figure 8.4 then depicts an example mapping between the various identifiers used over the Gb interface.

At the lowest layer of the Gb interface are the bearer channels, which are the physical channels carrying upper layer information. The bearer channels could be, for example, a defined set of $n \times 64$ kbps PCM channels on an E1/T1 physical link. FR virtual circuits are established on these bearer channels, and they are themselves identified using DLCI values.

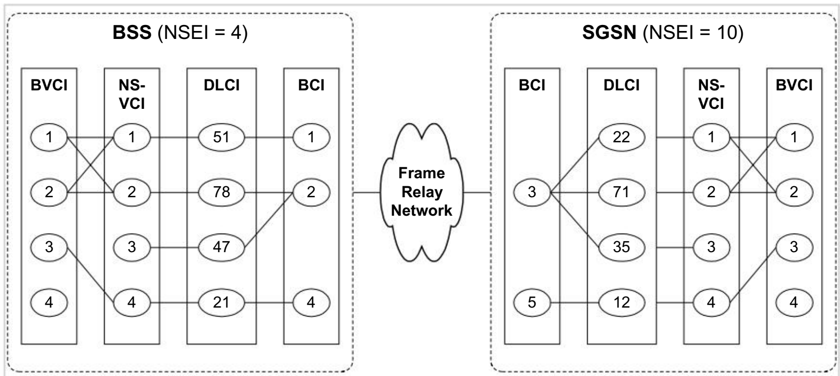
Table 8.1

Gb Interface Addresses/Identifiers

Identity	*Description*	*Composition*
BCI	Bearer Channel Identifier (BCI); identifies a BC, which is a physical channel carrying frame relay (NS) information	BC can be an $n \times$ 64 kbps channel on E1/T1 link
DLCI	Data Link Connection Identifier (DLCI); identifies an FR (NS) virtual circuit toward a particular destination; has local significance between FR endpoint and FR switch	10 bits (chosen by operator, or run-time assigned via signaling)
NS-VCI	Network Service Virtual Connection Identifier (NS-VCI); identifies an NS-VC, which is an end-to-end logical connection between BSS and SGSN; unlike DLCI, NS-VCI has end-to-end significance between BSS and SGSN	2-octet value; maps to a DLCI and BC identifier in the lower layer
NSEI	Network Service Entity Identifier (NSEI); identifies an NSE (BSS or SGSN), and its associated group of NS-VCs	2-octet value; unambiguously identifies one NSE
BVCI	BSSGP Virtual Connection Identifier (BVCI); identifies an end-to-end virtual connection between BSS and SGSN at the BSSGP layer	2-octet value; 0×0000 reserved for signaling functional entity

FIGURE 8.4

Gb Interface Addressing Scheme



As stated earlier, DLCI uniquely identifies a FR virtual circuit between two FR nodes (end nodes, or FR switches).

The GPRS specifications build their own addressing scheme on top of the FR addressing mechanism. First and foremost, it is important to identify the two GPRS network nodes connected via the Gb interface. This is done using an NSEI, which identifies a BSS or SGSN on the Gb interface. The NSEI then can be used to identify multiple BSS connected to the same SGSN. Each NSE uses a collection of end-to-end communication channels for information exchange over the Gb interface. Since the DLCI is only a local identifier, a new identifier NS-VCI identifies the channels end-to-end over the Gb interface.

The upper-most identifier over the Gb interface is the BVCI. Since the BSS services a collection of cells within the GPRS network, a separate logical channel is used over the Gb interface to carry data that belong to each cell. In other words, a BVCI identifies a channel carrying data specific to a particular cell serviced by a particular BSS. Two broad types of BSSGP Virtual Connections (BVC) are used, namely a signaling BVC carrying all signaling information between the two NSEs, and a set of traffic BVCs, carrying traffic for a particular cell of a BSS.

As depicted in Fig. 8.4, BVCs are mapped onto multiple NS-VCs, so as to provide redundancy of links, and also for load sharing perspective. In case one of the underlying links servicing an NS-VC becomes faulty, all traffic for the BVC is routed over the other functional NS-VCs, thus offering redundancy against link faults.

8.2.4.2 Network Service Layer

The NS layer on the Gb interface is based on the FR networking technology. The NS layer is divided into two sublayers, namely the Subnetwork Service (SNS) sublayer and the Network Service Control (NSC) sublayer (see Fig. 8.5). The SNS sublayer is the lower sublayer, and its functions are specific to the FR protocol. The SNS sublayer manages the FR protocol and its associated virtual circuits. The NSC forms the upper sublayer of the NS layer and is independent of the underlying FR protocol. The NSC sublayer directly interacts with the BSSGP layer sitting above the NS layer, and provides end-to-end control procedures between

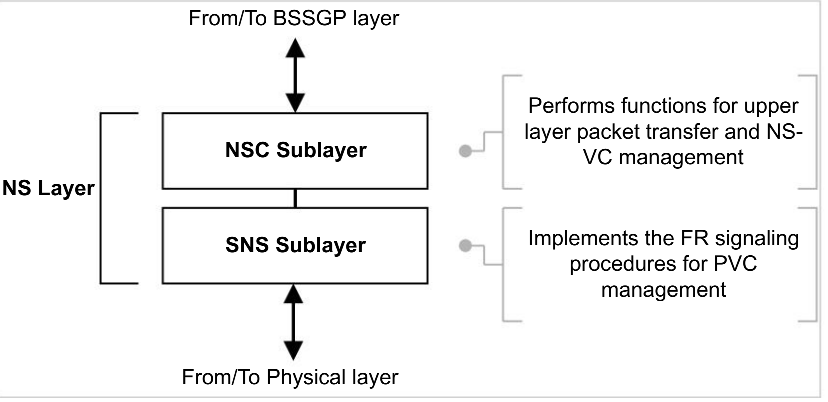


FIGURE 8.5 Sublayers of NS Layer

the BSS and the SGSN. The division of the NS layer into two sublayers allows the underlying networking protocol (currently FR) to be changed without impacting the other parts of the Gb interface.

The SNS sublayer of the NS layer is completely based on the FR protocol. The Gb interface is defined as a FR User to Network Interface (UNI) and typically uses FR Permanent Virtual Circuits (PVC). The functions of the SNS sublayer are thus related to the management of the PVCs between the BSS and the SGSN. This includes procedures for handling new PVC addition, detection of existing PVC deletion, and verification of PVC availability and link integrity. The procedures at the SNS sublayer are therefore the FR signaling procedures for PVC management.

The NSC sublayer functions above the underlying FR protocol. Its functions broadly include the following:

- **Transfer of upper layer packets.** The NSC sublayer interacts with the upper BSSGP layer and performs functions for the transfer of upper layer packets over the Gb interface. It uses the NS-UNIT DATA message for this purpose and encapsulates the upper layer Service Data Unit (SDU) into this message, before transmission over the Gb interface. The BVCI is also included as part of this message to identify the intended recipient at the BSSGP layer.
- **Management of NS-VC between BSS and SGSN.** The NSC sublayer is responsible for the management of the NS-VCs between the BSS and the SGSN. The NS-VCs at the NSC sublayer are statically configured via O&M mechanisms, and they map directly to underlying PVCs at the SNS sublayer. The NS-VC management procedures include functions for changing the administrative states of the NS-VCs (DEAD, BLOCKED & ALIVE, and UNBLOCKED). It also includes procedures for performing the reset of an NS-VC and for performing the test procedures on the NS-VC to verify its integrity and the end-to-end communication path.

For more details on the NS layer and its functions, the reader is referred to [3GPP TS 48.016].

8.2.4.3 BSSGP Layer

The BSSGP can be considered as the most important protocol on the Gb interface. It is used in both the control plane procedures and the user plane procedures performed over the Gb interface. The BSSGP layer is also responsible for the management of the BVC, on which all data and signaling for a particular cell is transferred.

Broadly, the BSSGP layer performs the following functions:

- **Transfer of LLC PDUs.** The BSSGP layer is used for the transfer of LLC PDUs over the Gb interface. The LLC PDUs are transferred over the Gb interface in unacknowledged mode. In the downlink direction (SGSN to BSS), the DL-UNITDATA message is used for carrying an LLC PDU from the SGSN to the BSS. In the uplink direction, the UL-UNITDATA message is used for carrying the LLC PDU in the reverse direction, from the BSS to the SGSN. The DL-UNITDATA and the UL-UNITDATA messages are exchanged on the BVC corresponding to the cell from which the LLC PDU has originated, or to which it is destined. The TLLI included in the DL-UNITDATA and UL-UNITDATA messages identifies the mobile from which the LLC PDU has originated, or to which it is destined.
- **GMM signaling procedures.** The BSSGP layer provides functions for certain mobility management procedures between the BSS and the SGSN. The BSSGP layer defines two PDUs, the PAGING-PS and PAGING-CS, using which the SGSN can request paging of an MS within a BSS area. The PAGING-PS PDU is used for paging a mobile for GPRS services and initiated by the SGSN. The PAGING-CS PDU is used for paging a mobile for non-GPRS services. This is used when the MSC/VLR requests the SGSN for paging of the MS using the optional Gs interface.

  Besides the PDUs for paging procedures, the BSSGP also provides PDUs for fetching the Radio Access (RA) capability of an MS, and for indicating the radio status to the SGSN. The RA-CAPABILITY-UPDATE PDU is used by the BSS to request the SGSN to provide the RA capability of the MS and/or its IMSI. The SGSN provides this information to the BSS using the RA-CAPABILITY-UPDATE-ACK PDU. The RADIO-STATUS PDU of the BSSGP layer is used by the BSS to inform the SGSN about the radio status, and the reason why the downlink transfer of an LLC PDU could not happen. Using this procedure, the BSS informs the SGSN about the failure reason for downlink LLC PDU transfer, which could happen due to the MS being out of coverage area or due to poor link quality, or other reasons.
- **Management of Gb interface virtual connections.** The BSSGP layer provides procedures for the management of the Gb interface virtual connections, and for the associated PDUs flowing over these connections. The first important procedure used over the Gb interface is the BSSGP flush procedure. This procedure is used by the SGSN to indicate to the BSS that the downlink data packet toward an MS is to be deleted, or rerouted, due to the MS having performed a cell reselection. In case the new cell of the MS is within the same BSS, the SGSN informs the BSS of the old and the new BVCI using the FLUSH-LL PDU. Using this information, the BSS is able to reroute the

downlink packet toward the correct cell. In case, however, the new cell of the MS is in a different BSS, the SGSN simply indicates to the BSS to delete the downlink packet. A procedure similar to the flush procedure is the LLC discarded procedure. Using this procedure, the BSS sends an LLC-DISCARDED PDU to the SGSN to indicate that one or more LLC frames have been discarded for a BVCI. This may occur due to different reasons, for example, cell reselection, or PDU lifetime expiry.

As part of the management procedures, the BSSGP layer also implements a flow control mechanism between the BSS and the SGSN, for flow control of downlink UNITDATA PDUs being transferred from the SGSN to the BSS. Since the BSS has limited buffering capability, it is essential to control the flow of DL-UNITDATA PDUs from the SGSN to the BSS, so as to limit the loss of LLC PDUs due to buffer overflow. Using the flow control procedure, the BSS informs the SGSN certain flow control parameters, using which, the SGSN is able to control its downlink transmission toward the BSS.

Another key management procedure of the BSSGP layer is the management of the BVC. Each BVC used for a particular cell could be in one of the two states, namely BLOCKED and UNBLOCKED. No data transfer is possible for a cell if its associated BVC is BLOCKED. BSSGP layer defines three different PDUs for the control of the BVC states, namely BVC-BLOCK, BVC-UNBLOCK, and BVC-RESET. While the BVC-BLOCK and BVC-UNBLOCK PDUs are sent by the BSS to the SGSN to perform BVC blocking and unblocking, respectively, the BVC-RESET PDU may be sent by either the SGSN or the BSS, leading to a BVC reset. The BVC reset procedure is used to reinitialize and synchronize the BVC context between the BSS and the SGSN.

- **Management of packet flow contexts.** The BSSGP layer also defines procedures for the management of packet flow contexts at the BSS. Each packet flow context within the BSS is identified using a Packet Flow Identifier (PFI), which is allocated by the SGSN. A packet flow context contains the BSS QoS profile, which is identical for one or more activated PDP contexts. BSSGP defines three procedures for the management of packet flow contexts within the BSS. These include separate procedures for packet flow context creation, packet flow context modification, and packet flow context deletion.

For more details on the BSSGP layer and its functions, the reader is referred to [3GPP TS 48.018].

8.3 User Plane within GPRS Backbone

The previous section described the nuances of the GPRS user plane segment from the MS to the SGSN. The second segment of the end-to-end user plane (see Fig. 8.6) resides purely within the intra-PLMN GPRS backbone and extends from the SGSN to the GGSN (over Gn interface), from where onward the packet is transmitted toward the external networks (over Gi interface). Alternatively, the packet may also be routed between two different GPRS packet domains, over an inter-PLMN GPRS backbone (over Gp interface). The primary protocol used within the GPRS backbone (Gn or Gp) for the transfer of user plane packets is the GTP-U protocol. The following section describes the role of the GTP-U in further detail.

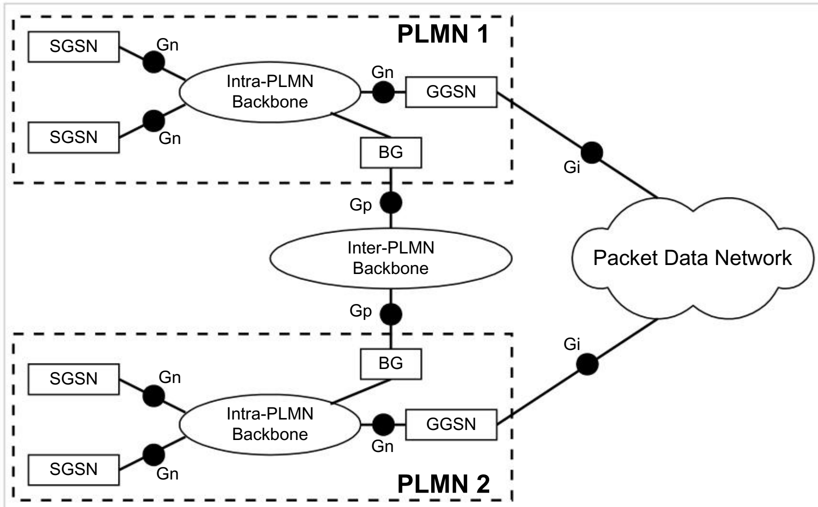


FIGURE 8.6

GPRS User Plane Within GPRS Backbone and Toward External PDN

8.3.1 GTP-U Layer

GTP-U is the main protocol used within the GPRS backbone for the transport of user plane packets between the SGSN and GGSN. GTP-U is a tunneling protocol, which resides above the UDP/IP layers (see Fig. 8.1). The user payload (typically IP packets) is encapsulated as GTP packets and exchanged using UDP/IP routing. Finally, at GGSN, the IP packets are de-encapsulated and exchanged as normal IP packets with the external PDN.

When the SGSN and the GGSN reside within the same GPRS packet domain, the GTP-U protocol is used as the user plane protocol over the Gn interface. GTP-U is also defined as the user plane protocol when interworking is required between two different GPRS packet domain networks. In other words, it is possible to use GTP-U as the user plane protocol in the inter-PLMN GPRS backbone to interwork between two separate GPRS packet domain networks. In this case, the GTP-U is used over the Gp interface.

A GTP-U tunnel is identified at either end (SGSN or GGSN) by using a Tunnel Endpoint Identifier (TEID). For a particular MS, a number of GTP-U tunnels can exist, one per PDP context (i.e., session). These GTP-U tunnels are created during the PDP context activation procedure, or during the PDP context modification procedures.

The GTP header also includes a sequence number field. This is used if in-sequence delivery of user packets is desired between the SGSN and the GGSN. In such a case, the transmitting GSN associates a sequence number with each transmitted packet, such that the sequence number is carried as part of the GTP header. The receiving GSN entity then performs the reordering of the received packets as per the sequence number information within the GTP header.

For more details on the GTP-U protocol and its associated procedures, the reader is referred to [3GPP TS 29.060].

8.4 Interworking with External Packet Networks

The final segment of the end-to-end user plane transmission path extends from the GGSN within the GPRS packet domain to the external PDN. This segment is also called the Gi interface (see Fig. 8.6). The external PDN is typically an IP network. In this case, both the IPv4 and the IPv6 protocols are supported over the Gi interface. For the external PDN, the GGSN is seen as a router that routes IP packets. The important difference between a normal IP router and the routing functionality of the GGSN is that the IP packets are tunneled to the MS and not routed as normal IP datagrams.

The access to an external PDN may involve multiple other functions, namely, user authentication and authorization, end-to-end encryption and allocation of a dynamic PDP address belonging to the PLMN/Intranet/ISP addressing space, besides some others. In respect of these functions, the GPRS packet domain works in one of the two possible operating modes as follows:

- **Transparent access to Internet/Intranet.** In this operating mode, the MS is given a PDP address from the addressing space of the GPRS network operator. The PDP address may either be allocated to the MS at subscription time (statically assigned PDP address), or it may be allocated at the time of PDP context activation (dynamically assigned PDP address). This PDP address is then used for packet forwarding between the GGSN and the Internet/Intranet.

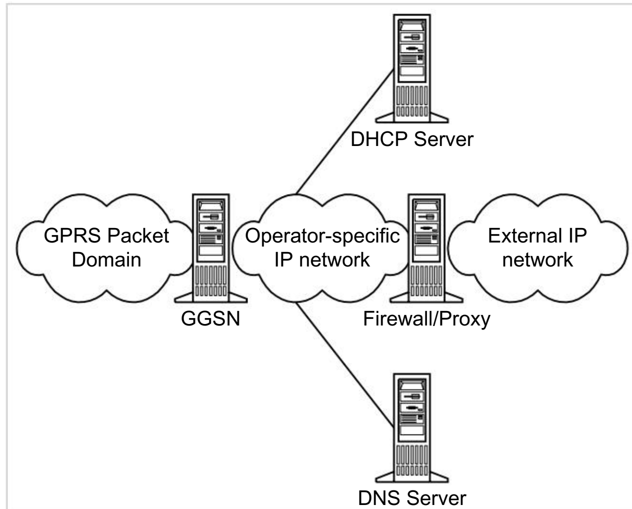
  In the transparent access mode, no authentication and authorization procedure is separately required to be supported by the GPRS domain. Thus, the MS need not send any authentication request at the time of PDP context activation, and the GGSN need not participate in any authentication or authorization procedure for the MS. In other words, the transparent access mode can be seen to provide a form of basic ISP service for Internet access. Figure 8.7 depicts an example configuration for transparent access to an Internet/Intranet.

  When the transparent access mode is used to access an Intranet, certain functions, such as end-to-end security, may be provided using protocols that are specific to the Intranet. In such cases, the MS as well as the IP node within the Intranet may implement certain Intranet- specific protocols, which reside above the IP layer. One example of this scenario is the use of the IPSec protocol for providing an end-to-end secure communication channel between the MS and the Intranet. Note that the GPRS domain is transparent to the use of the IPSec protocol, and thus operating in the transparent access mode.

- **Nontransparent Access to Intranet/ISP.** In this operating mode, the MS is given a PDP address from the addressing space of the ISP/Intranet. As in the case of transparent access mode, the PDP address may either be allocated to the MS at subscription time (statically assigned PDP address), or it may be allocated at the time of PDP context activation (dynamically assigned PDP address). This PDP address is then used for packet forwarding between the GGSN and the ISP/Intranet.

Figure 8.7

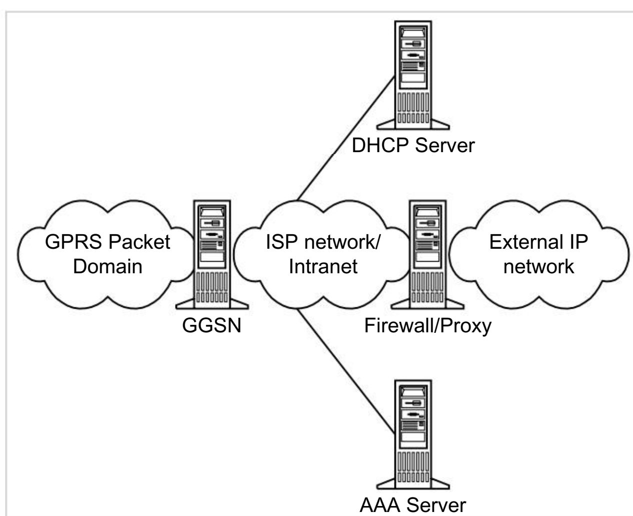
Transparent Access to Internet/Intranet



In the nontransparent access mode, specific servers within the ISP/Intranet implement the authentication and authorization procedures as well as PDP address allocation functions. AAA/DHCP servers within the Intranet/ISP can be used to perform such specific functions. It is thus required that an MS sends an authentication request at the time of PDP context activation, and the GGSN then requests user authentication from an AAA server within the Intranet/ISP. Similarly, other protocol configuration options can be retrieved from relevant servers within the Intranet/ISP at the time of PDP context activation. Figure 8.8 depicts an example configuration for nontransparent access to an Intranet/ISP network.

Figure 8.8

Nontransparent Access to Intranet/ISP Network



For more details on the interworking procedures used between the GPRS packet domain and the external PDN, the reader is referred to [3GPP TS 29.061].

Conclusion

The GPRS user plane provides a data transmission path to the external network according to the QoS parameters agreed during PDP context establishment. The complete end-to-end path can be seen as a combination of multiple smaller segments, each having its own associated user plane protocol stack. The first segment extends from the MS to the SGSN and includes LLC and SNDCP as the primary protocols, besides RLC/MAC and BSSGP. The segment between the MS and the SGSN was described in Section 8.2, along with a description of the key associated user plane protocols.

The next segment extends from the SGSN till the GGSN, and here the GTP-U is the prime user plane protocol. The GTP-U can also be used as the user plane protocol when interworking between two different GPRS packet domain networks. This second segment of the end-to-end path, along with the GTP-U description, was provided in Section 8.3.

The final leg of the end-to-end path extends from the GGSN to the external packet data network. Two modes of operation are possible for access to the external network, namely transparent access and nontransparent access. These modes provide different mechanisms for interworking of GPRS packet domain with an external Intranet/ISP/Internet network. The last segment of the end-to-end transmission path was discussed in Section 8.4.

Further Reading

The end-to-end user plane transmission path consists of different segments and their associated user plane protocols. Each of these user plane protocols are discussed in separate 3GPP specifications, which can be referred for more details on the relevant topic.

Most user plane protocols used within the segment from the MS to the SGSN are defined in the 3GPP 44 series and 48 series specifications. In the 44 series specifications, [3GPP TS 44.060] describes RLC/MAC, [3GPP TS 44.064] describes LLC, and [3GPP TS 44.065] describes the SNDCP protocol. In the 48 series specifications, [3GPP TS 48.016] describes the NS layer, while [3GPP TS 48.018] describes the BSSGP protocol.

Within the GPRS packet domain, GTP-U is the prime protocol, and it is described in [3GPP TS 29.060]. Interworking between GPRS packet domain networks and external PDNs is detailed in [3GPP TS 29.061].

Chapter 9

Enhanced Data Rates for Global Evolution

9.1 Introduction

Enhanced Data Rates for Global Evolution (EDGE) is a high-speed mobile data standard that can be introduced in the GSM and the GPRS system. EDGE is considered as a 2.5G standard, a transition between 2G and 3G. When applied on the GSM system, it works on top of the High-Speed Circuit Switched Data (HSCSD) and is known as the Enhanced Circuit Switched Data (ECSD). This allows data transfers of rates higher than those allowed in the current GSM data rates (2.4, 4.8, 9.6, and 14.4 kbps). The higher data rates are 28.8, 32, and 43.2 kbps per timeslot.

The evolution of GPRS toward EDGE is known as Enhanced GPRS (EGPRS). It is also known as EDGE classic. This allows data rates of up to 475 kbps for a receiver-supporting Rx on eight timeslots. EGPRS led to changes in the radio interface due to the introduction of a new modulation scheme called the Eight-state Phase-Shift Keying (8-PSK). The major impact of the modulation scheme is on the Base Station Subsystem (BSS) and on the MS.

This chapter looks at some of the key aspects of EDGE and how they relate to the concepts of GPRS covered so far in this book. The first sections deal with the modifications and the capabilities of the network and the MS required in order to make use of the EDGE services. The physical layer and the concepts of the air interface are then discussed. A detailed description is provided about the new modulation and coding schemes that have been introduced as a part of EGPRS. The next section introduces the concept of Link Adaptation (LA). The measurements that are made by the mobile and the mechanisms that are undertaken for LA are discussed.

The last section discusses the impact on the Radio Link Control (RLC) and Medium Access Control (MAC) procedures such as the Temporary Block Flow (TBF) establishment, the acknowledgment of the data blocks, the polling procedure, and so on. The structure of the RLC data blocks and the different fields for EGPRS are also discussed.

9.2 Impact of EDGE on Network Architecture

EDGE can be introduced in both the GSM and the GPRS networks to provide higher data rates. Introduction of EDGE does not require any significant changes in the core network architecture. One example of a minor change would be the MS-MSC/VLR signaling where the requested data rates (i.e., EDGE rates) are exchanged.

The major changes due to EDGE deployment come in the radio interface on BSS side (including BSC and Base Transceiver Station (BTS)) and the MS side. At the BSS, changes are required in the hardware of the Transceiver (TRX) units installed in the BTS. The TRXs must be capable of supporting the EDGE data rates. These units need to be installed in the existing BTS of any network in order to support the features of EDGE. The BTS software also needs upgrade to handle various aspects including support for additional channel coding schemes, MCS1 to 9. On the Packet Control Unit (PCU) side, there are software enhancements in the RLC/MAC layer for EDGE.

Just as some changes are needed at the BSS, the MS also needs to have some special capabilities to support the EGPRS features to support high data rates introduced by EDGE. The high data rates are achieved by using a new modulation scheme 8-PSK in addition to the existing GSM modulation scheme, Gaussian Minimum Shift Keying (GMSK). Also new Modulation and Coding Schemes (MCS), namely MCS1–MCS9, are introduced on the air interface. While MCS1–MCS4 use the GMSK modulation, the coding schemes MCS5–MCS9 use the 8-PSK modulation scheme. This leads to some special requirements from the MS. The 8-PSK modulation has high spectral efficiency as a result of which the mobile-baseband part needs to have more complex algorithms for the equalization, coding, and decoding processes, including the support of the new coding schemes introduced for the data transfer.

Like GPRS, the data transfers in EGPRS also take place in the form of a TBF. The concept of the TBF is the same as in the case of GPRS. The data transfer from the MS to the network constitutes an uplink TBF, while the transfer of data from the network to the MS is known as the downlink TBF. It is possible that more than one timeslot be allocated for a TBF. The allocation is governed by the multislot class of the mobile. The EGPRS multislot class is different from the GPRS multislot class. The EDGE capable mobile is characterized by both the classes: one purely EGPRS multislot class and one purely GPRS multislot class. The MS supports the EGPRS multislot class in the GPRS mode as well as in the EGPRS mode. The multislot classes are characterized by the following parameters:

- the maximum number of Tx timeslots allowed
- the maximum number of Rx timeslots allowed
- the maximum number of combined Rx + Tx timeslots allowed

These parameters are the same as that used for the GPRS multislot classes. The multislot classes are listed in the Table 9.1. The EDGE capable mobiles have also been divided into two classes on the basis of the modulation schemes that are supported:

TABLE 9.1

Multislot Classes for Mobile Station

Multislot	*Maximum Number of Time Slots*			*Type*
Class	*Rx*	*Tx*	*Sum*	
1	1	1	2	1
2	2	1	3	1
3	2	2	3	1
4	3	1	4	1
5	2	2	4	1
6	3	2	4	1
7	3	3	4	1
8	4	1	5	1
9	3	2	5	1
10	4	2	5	1
11	4	3	5	1
12	4	4	5	1
13	3	3	N/A	2
14	4	4	N/A	2
15	5	5	N/A	2
16	6	6	N/A	2
17	7	7	N/A	2
18	8	8	N/A	2
19	6	2	N/A	1
20	6	3	N/A	1
21	6	4	N/A	1
22	6	4	N/A	1
23	6	6	N/A	1
24	8	2	N/A	1
25	8	3	N/A	1
26	8	4	N/A	1
27	8	4	N/A	1
28	8	6	N/A	1
29	8	8	N/A	1
30	5	1	6	1
31	5	2	6	1
32	5	3	6	1
33	5	4	6	1
34	5	5	6	1
35	5	1	6	1
36	5	2	6	1
37	5	3	6	1
38	5	4	6	1
39	5	5	6	1
40	6	1	7	1
41	6	2	7	1
42	6	3	7	1
43	6	4	7	1
44	6	5	7	1
45	6	6	7	1

Note 1: Mobiles of Type 1 are not required to receive and transmit at the same time.

Note 2: Mobiles of Type 2 are required to receive and transmit at the same time.

- Type 1 MS supports both the GMSK and the 8-PSK modulation scheme in the downlink direction but only the GMSK modulation in the uplink direction
- Type 2 MS supports 8-PSK in both the uplink and the downlink directions.

The mobiles of the first class support MCS1–MCS9 coding schemes in the downlink direction, whereas only MCS1–MCS4 are supported in the uplink direction. On the other hand, the mobiles of the second class support all the coding schemes MCS1–MCS9 in both the uplink and the downlink directions. These class definitions for the mobile can be easily explained if the throughput requirements for the EGPRS services are considered. The most common services using EGPRS are Web browsing and video broadcasts. These services require high data rates in the downlink direction for the data transfer. But the transfers in the uplink direction mainly constitute the signaling and the commands, which does not demand high data rate. Owing to these requirements, the constraints of support of 8-PSK have not been put in both the directions.

9.3 EDGE Modulation and Coding Schemes

The EGPRS radio interface is the same as that of GPRS. The concept of the multiframes (i.e., 52-frame multiframes), the physical channels (i.e., PDCH), the logical channels (PBCCH, PCCCH, PRACH, PDTCH, PTCCH, and PACCH), and their mapping on the physical channels is the same as in the case of GPRS. The only difference is in terms of a new modulation scheme and the introduction of new coding schemes. The modulation and the coding schemes are discussed in detail next.

The major advantage of EGPRS is the higher throughputs as compared to GPRS. This is mainly due to the new modulation scheme used for the data transfer. While GPRS uses the GMSK modulation which is associated with limited spectrum efficiency, EGPRS uses the 8-PSK modulation. 8-PSK has the eight-state constellation and thus allows every symbol to be represented with three bits. This allows the bit rate to be three times higher than the GMSK modulation.

In addition to the modulation scheme, EGPRS also uses new coding schemes MCS1–MCS9 that provide higher data rates and better mechanisms for link adaptation. Table 9.2 lists the maximum throughput that can be obtained from the different EGPRS coding schemes.

As depicted in Table 9.2, the coding schemes MCS1–MCS4 use the GMSK modulation, whereas MCS5–MCS9 use the 8-PSK modulation. It is possible for the network to support EGPRS and not 8-PSK modulation. However, the benefits of EGPRS-enhanced throughput would not be available in such a network. Thus, the MCS1–MCS4 coding schemes would be supported, while the schemes MCS5–MCS9 would not be supported. The only advantage of EGPRS without 8-PSK would be improved signaling procedures in the MAC layer, as discussed in the later sections of this chapter.

The coding scheme to be used for any data transfer is determined by the network depending on the radio conditions. The coding schemes have been designed in order to provide better performance and higher throughput as compared to the GPRS coding schemes.

TABLE 9.2

EGPRS Modulation Schemes and Associated Attributes

Modulation and Coding Scheme	*Modulation*	*Code Rate*	*Max. Throughput (kbps)*	*Raw Data within Radio Block*	*RLC Blocks/ Radio Block*	*Family*
MCS1	GMSK	0.37	8.8	178	1	C
MCS2	GMSK	0.49	11.2	226	1	B
MCS3	GMSK	0.53	14.8	298	1	A
MCS4	GMSK	0.66	17.6	354	1	C
MCS5	8-PSK	0.76	22.4	450	1	B
MCS6	8-PSK	0.85	29.6	594	1	A
MCS7	8-PSK	0.92	44.8	2 x 450	2	B
MCS8	8-PSK	1.0	54.4	2 x 546	2	A
MCS9	8-PSK	1.0	59.2	2 x 594	2	A

9.4 CHANNEL CODING

In GPRS, the higher layer PDUs are transmitted in the form of RLC/MAC data blocks which are different in uplink and downlink directions. Similarly, in EGPRS too, the downlink data are transferred using EGPRS downlink data block formats and the uplink data are transferred using EGPRS uplink data blocks. The block structures for GPRS and EGPRS are, however, different. These blocks are transmitted over four consecutive radio bursts on four Time Division Multiple Access (TDMA) frames of a given Packet Data Channel (i.e., PDCH). The EGPRS blocks are composed of a header part and two data parts (where second part is conditional). There are different data block formats for the different coding schemes. These are discussed in more detail in the subsequent subsections.

It is important to note that the multiplexing of GPRS and EGPRS MS is possible on the same timeslot. It is thus possible that a GPRS data block carries an Uplink State Flag (USF) meant for an MS that is involved in an EGPRS data transfer. Also, it is possible that the USF in an EGPRS data block is meant for a mobile that is involved in a GPRS data transfer. It should thus be possible for a GPRS mobile and an EGPRS mobile to decode the USF in the EGPRS and the GPRS data blocks, respectively. The channel coding is hence done to take care of this condition.

9.4.1 Channel Coding for MCS1–MCS4

As discussed earlier, the USF of a GRPS mobile may be encoded in an EGPRS data block. These blocks can be transmitted using the 8-PSK or the GMSK modulation. If 8-PSK modulation is used, the GPRS mobile would not be able to decode the USF carried in the block. However, if the GMSK modulation is used, the GPRS mobile can decode the block and hence the USF field. The EGPRS coding schemes, MCS1–MCS4, use the GMSK modulation and hence it is possible that these blocks may carry the USF for a GPRS mobile. The coding of the MCS1–MCS4 data blocks is thus done in the same way as the GPRS coding schemes, CS1–CS4.

For MCS1–MCS3, the USF is first precoded with a block code and then coded with the data with a convolution code. In case of MCS4, the USF is precoded with the block code only. In all these coding schemes, the same stealing flag as the CS4 coding scheme is chosen. It is thus easier for a GPRS mobile to decode the USF encoded in an EGPRS data block using coding scheme MCS1–MCS4. Such a mobile would be able to detect the stealing flag and hence decode the USF from the EGPRS data block, but it would not be able to decode the data encoded using the MCS.

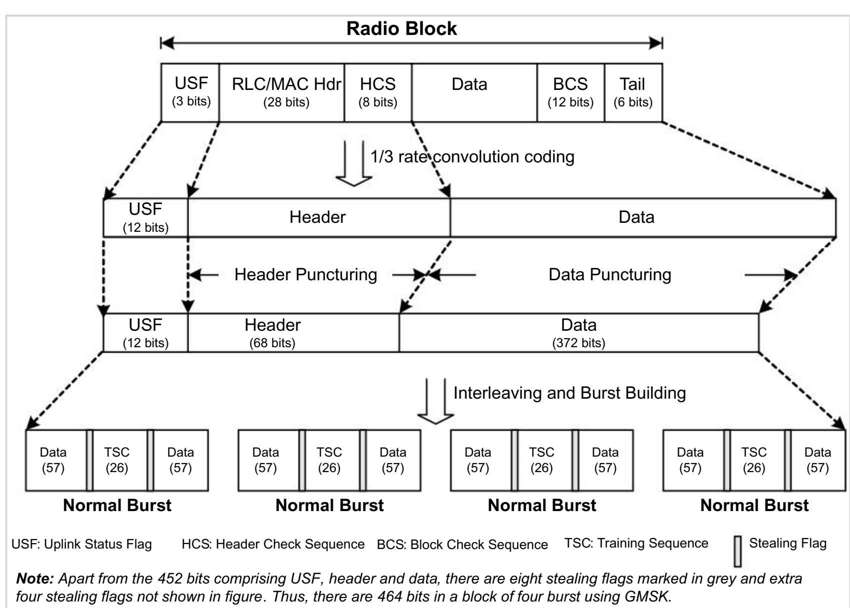
The data block consists of the header part and the data part. The header part also contains the USF. The data part is variable depending on the MCS that is being used. The data part of the EGPRS coding schemes is listed in Table 9.2.

The encoding of the data block is depicted in Fig. 9.1. As represented in the figure, the process of encoding starts by adding a Header Check Sequence (HCS) and a Block Check Sequence (BCS) to the header and the data part. The 3-bit USF is then precoded to 12 bits.

The header part and the HCS are then encoded using cyclical encoding with 1/3 convolutional coding. The data part with the BCS is encoded by adding the tail bits (6 bits) and then encoding with the same rate 1/3 convolutional code. Both the header part and the data part are punctured. The header is lightly punctured, while the data part is punctured variably depending on the puncturing scheme that is used.

FIGURE 9.1

Channel Coding for MCS1–MCS4 Data Blocks



There are many puncturing schemes available for the same coding scheme. The transmitter can use any scheme. The scheme that has been used by the sender is indicated in the Coding and Puncturing Scheme (CPS) field of the RLC/MAC data block. The use of this field is explained in more detail when the RLC data blocks are discussed.

After the puncturing, the data block size is 452 bits. A stealing bit pattern of 12 bits is added to the block. Among these 12 bits, 8 bits are adjacent to the training sequence and are used to indicate the coding scheme that is being used. Four other bits called extra stealing flags are spread across the four normal bursts and are defined for future use and set to 0. Here, it may be noted that the extra stealing flags are used only for the MCS1–MCS4 that use GMSK modulation and are not used for higher MCS schemes. Another difference that is shortly observed is that in MCS1–MCS4, the stealing flags are placed adjacent to the training sequence, whereas in higher coding schemes, they are placed to the right.

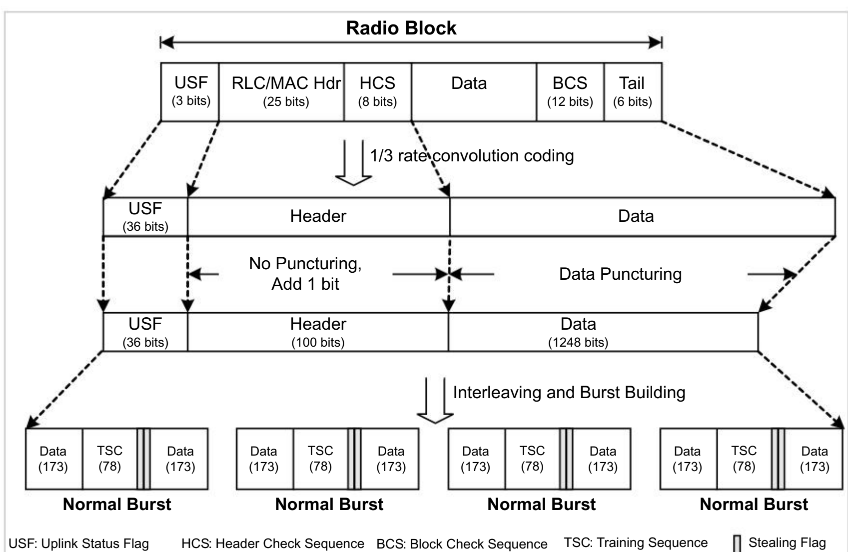
9.4.2 Channel Coding for MCS5–MCS6

The encoding of the MCS5 and MCS6 data blocks follows almost the same principles as the MCS1–MCS4 data blocks. The encoding is described in Fig. 9.2.The header part and the data part are used for the encoding. Both the coding schemes share the same structure for the RLC/MAC header. The sizes of the data part of the blocks, however, differ for both the coding schemes as listed in Table 9.2.

The encoding of the USF in the blocks differs from the MCS1 to MCS4. In this case, the USF is precoded with 36 bits. Also, the data block header is not punctured after coding in this encoding. The fundamental difference in the encoding of the MCS5 and the MCS6 data blocks

FIGURE 9.2

Channel Encoding for MCS5 and MCS6 Data Blocks



Note: *Apart from the 1384 bits comprising USF, header and data, there are eight stealing flags marked in grey (there are no extra stealing flags). Thus, there are 1392 bits in a block of four bursts using eight PSK.*

when compared to the MCS1–MCS4 data blocks is that these blocks use the 8-PSK modulation. The use of this modulation scheme results in three times higher number of encoded bits since each symbol is represented by a 3-bit code. The number of bits is 1392. These bits are then interleaved onto four bursts that are transmitted on the air interface.

9.4.3 Channel Coding for MCS7–MCS9

The data blocks of coding schemes MCS7–MCS9 carry two RLC data blocks. The encoding is hence done on the header part of the data block, the data part of the first RLC data block, and the data part of the second RLC data block. The USF is precoded with 36 bits. The check sequence of 8 bits is added to the header and 12 bits for each data part. The encoding of the data block is done using 1/3 convolutional coding and then the bits are punctured. The procedure for the encoding is described in Fig. 9.3.

There are different puncturing schemes for the coding schemes in this case. These puncturing schemes can be used on both the data parts of the data block. The scheme that is used is indicated in the header of the data block in the CPS field.

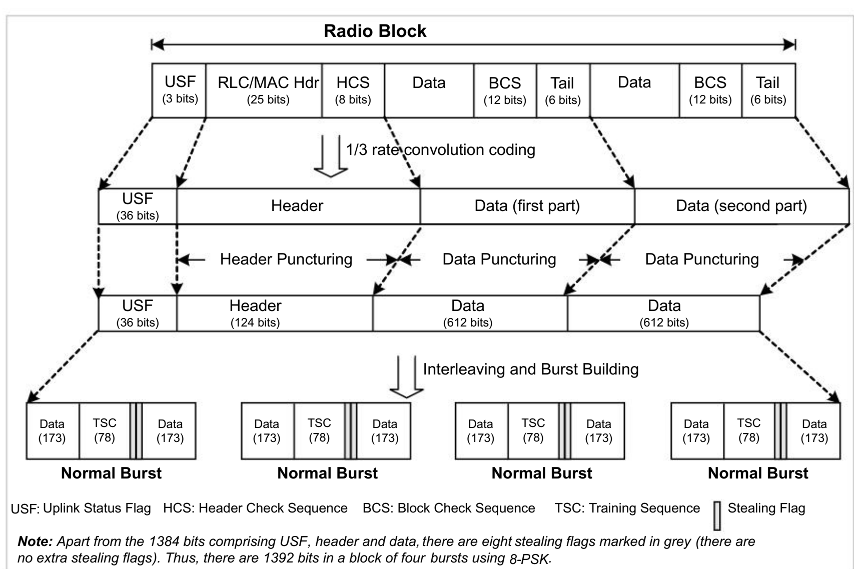
The encoded bits are interleaved and transmitted over the air interface. In case of MCS7 coding scheme, each encoded data part gets mapped to four bursts on the air interface. In the case of MCS8 and MCS9, each data part is interleaved on two bursts.

9.4.4 Channel Encoding for Control Channels

In case of GPRS, the coding scheme CS1 is used on the control channels. In the case of EGPRS also, the same coding scheme is used for all the signaling procedures. In the case of Random

FIGURE 9.3

Channel Coding for MCS7, MCS8, and MCS9



Access Channel (RACH), however, two new training sequences have been defined for sending the channel request messages. These are explained in more detail in the Section 9.7.

9.4.5 Concept of Coding Family

The MCS coding schemes have been divided into four families: A, A', B, and C. Each family consists of a set of the MCS and is associated with a data unit of a fixed size. When a data transfer is initiated with a MCS, the radio blocks transmitted as part of the data transfer consist of the data units of its family. Different MCS belonging to the same family transmit different number of data units in a single radio block and hence different data rates are achieved.

Table 9.3 lists the families, the MCS belonging to each family, and the size of the data unit associated with each family. As depicted here, the MCS1 and MCS4 belong to the family C and transmit data units of the size 22 bytes. Similarly, MCS7, MCS5, and MCS2 belong to the family B and transmit the data units of the size 28 bytes. Also, MCS8 and MCS9 belong to the family A' and A, respectively. As observed in the Ttable 9.3, the MCS3 and MCS6 belong to both family A and A'. The size of the data units for A and A' are 37 and 34 bytes, respectively.

TABLE 9.3

MCS Families and the Basic Data Units

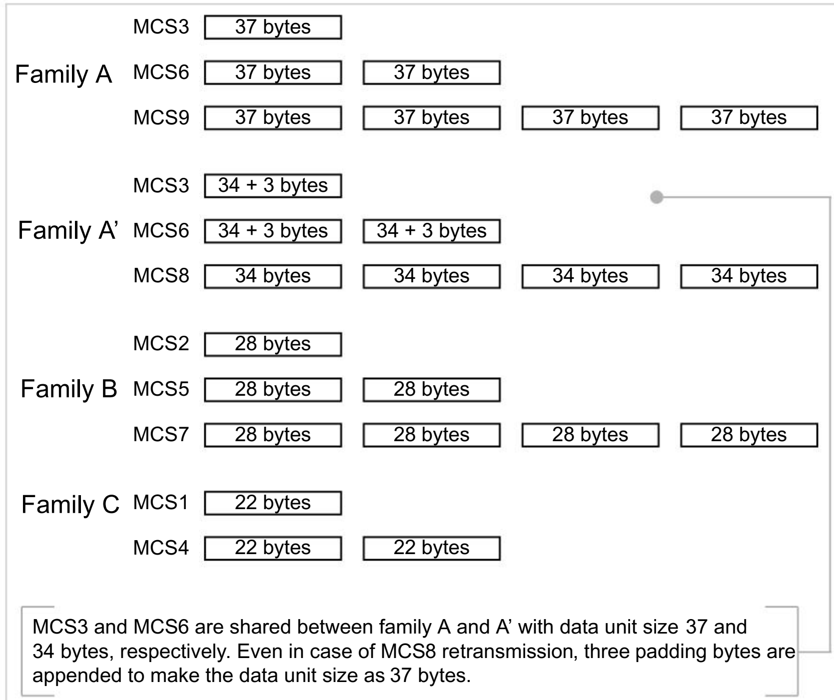
Family	*MCS Belonging to the Family*	*Size of Each Data Unit (Bytes)*
A	MCS9, MCS6, MCS3	37
A'	MCS8, MCS6, MCS3	34
B	MCS7, MCS5, MCS2	28
C	MCS4, MCS1	22

In case of retransmissions, a radio block may be transmitted using an MCS belonging to the same family as the original coding scheme. In other words, if an MCS9 block needs to be retransmitted, the radio blocks can be transmitted using MCS3 and MCS6. Similarly, the retransmission of an MCS8 block can be done using MCS3 and MCS6. From this table, it is clear that the family A and A' have different size of the data unit. When MCS3 or MCS6 are used for the retransmission of a MCS8 block, three padding bytes are appended to the data part of the radio block.

Another point to consider is the data rate during the retransmissions. When a radio block is sent using the MCS9, it consists of four data units of the size 37 bytes. However, when this block has to be retransmitted, the MCS3 or MCS6 can be used. If the MCS6 is used for the retransmissions, two radio blocks are required in order to transmit the four data units. This, hence, leads to a reduction in the data rate by two times. Similarly, if the retransmissions are done using MCS3, four radio blocks are needed to transfer the data units, leading to a four-time reduction in the data rate. Figure 9.4 depicts this concept clearly.

FIGURE 9.4

MCS Family and the Data Units for Each Coding Scheme



From this figure, it is clear that the coding schemes belonging to the same family transmit different number of data units and hence different data rates are achieved for the coding schemes.

9.5 RLC/MAC Block Structure

The introduction of new coding schemes and the increased throughputs in EGPRS necessitated some changes in the RLC/MAC layer. The structure of the RLC data blocks has been modified. Some changes have also been introduced in the procedures for TBF establishment in EGPRS mode, which are discussed in the later sections of the chapter.

9.5.1 EGPRS RLC Data Blocks

The changes to the radio interface in EGPRS necessitated changes in the RLC data block. The data block consists of the RLC/MAC header followed by RLC data parts depending on the coding schemes being used for the data transfer. If MCS7, MCS8, or MCS9 coding scheme is used, the data block consists of two RLC data parts, whereas for all other coding schemes, only one RLC data part is used. The structure of the data block is as depicted in Fig. 9.5

FIGURE 9.5

Structure of RLC Data Blocks

RLC/MAC Header	RLC Data Part 1 (Mandatory)	RLC Data Part 2 (Only for MCS7, MCS8, and MCS9)

The header part of the data block is important for the decoding of the rest of the radio block and also for the incremental redundancy feature. This header is hence encoded using a different coding scheme. The data parts present within the data block are coded independently using the defined coding scheme. Both the parts can be encoded using the different coding schemes but at the same rate. The control information is transmitted in the same control blocks as those used in case of GPRS.

There are nine different coding schemes, and three different headers are defined for these coding schemes:

- Header Type 1, defined for MCS7–MCS9
- Header Type 2, defined for MCS5 and MCS6
- Header Type 3, defined for MCS1–MCS4

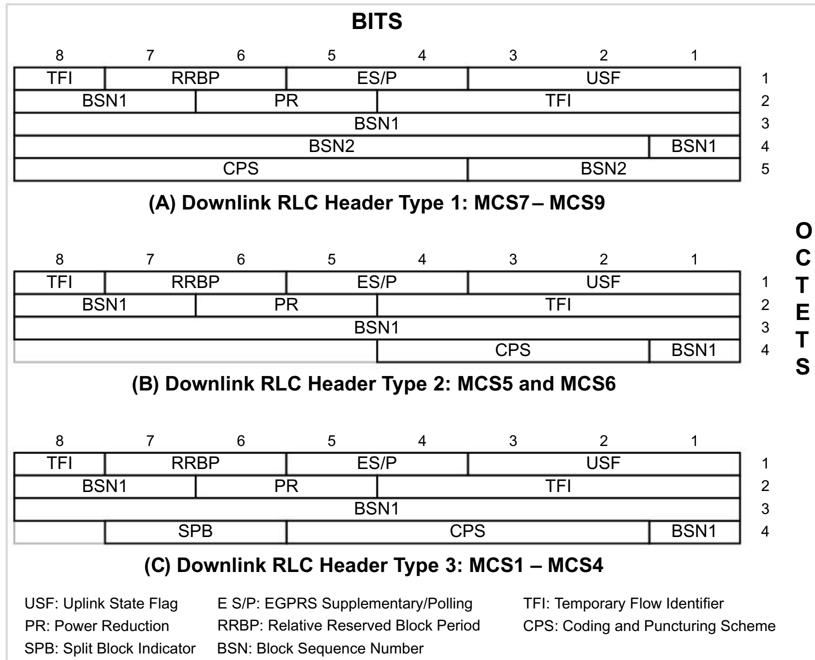
The headers are different in the uplink and the downlink directions. The downlink headers are discussed first (see Fig. 9.6).

The parameters in the downlink header are defined as follows:

- **Temporary Flow Identifier (TFI).** This identifies the TBF to which the data blocks belong. Note that this 5-bit field is spread over octets 1 and 2.
- **Relative Reserved Block Period (RRBP).** The network uses this field to schedule the transmission of a packet control ack message or PACCH block from the mobile. This parameter indicates the position of the uplink radio block on which the mobile should transmit the message.
- **Block Sequence Number.** The data transfer takes place in terms of data blocks. These data blocks are numbered in sequence, known as the Block Sequence Number (BSN). In EGPRS, the BSN is a 11-bit field and has the range 0–2047. In case two RLC data blocks are being carried in the data block, the BSN of the second block BSN2 is encoded relative to the BSN of the first block BSN1. The BSN2 field is applicable for header 1 carrying MCS7–MCS9 as only they carry multiple blocks in one message.
- **EGPRS Supplementary/Polling bit (E S/P).** This field indicates whether the RRBP field is valid or not. It also determines the contents of the next uplink control block. In case this block is supposed to carry the packet downlink ack/nack message, this field determines the contents of the message.
- **Uplink State Flag.** This field is used in the dynamic allocation of resources to determine which mobile would transmit in the next uplink radio block. It allows the multiplexing of multiple users on the same timeslot. The concept will be discussed in more detail in the subsequent sections.

FIGURE 9.6

Downlink RLC Header Type 1, 2, and 3



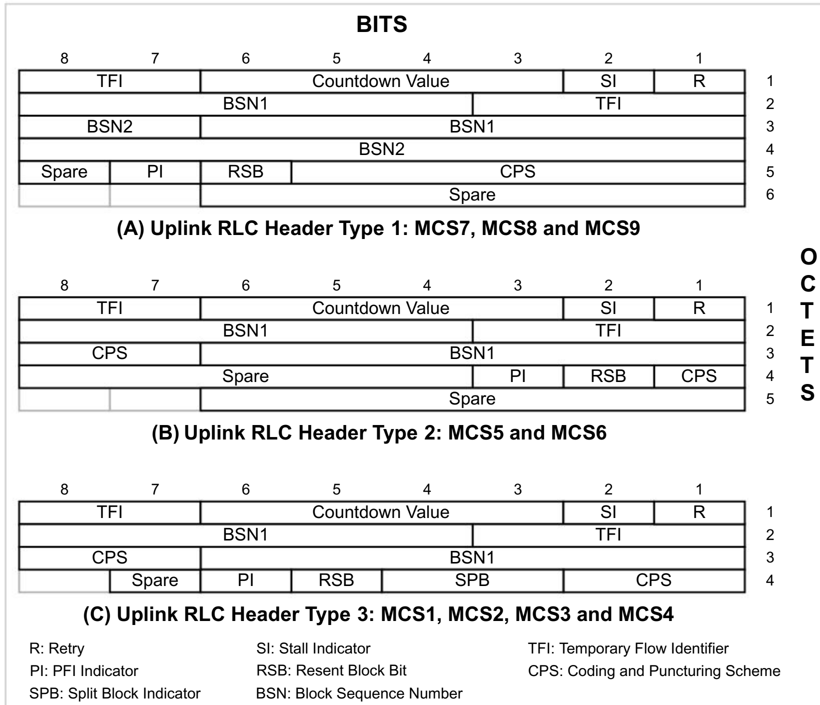
- **Power Reduction (PR).** This field is used in case of downlink power control to indicate the power level reduction of the current RLC data block.
- **Coding and Puncturing Scheme.** This field is used to indicate the kind of coding scheme and puncturing scheme used for the data block.
- **Split Block Indicator (SPB).** This field is used only in header type 3. The purpose is to indicate that a particular data block is segmented and retransmitted using two data blocks. This field also indicates whether the block is the first part or the second part of the block.

The RLC/MAC headers in the uplink direction are discussed in the rest of this section (see Fig. 9.7). Some parameters in the uplink header are the same as in the downlink header. These are hence not discussed again. The most important parameters in the uplink direction are those that provide information about the data transfer from the MS. These and other parameters that are encoded in the uplink header are as follows:

- **Countdown Value (CV).** This parameter allows the network to calculate the remaining RLC data blocks to be transferred in the ongoing data transfer. Depending on the network parameters, the mobile calculates this value and starts sending the countdown of the data blocks accordingly. A value 0 for this field indicates the last data block of the uplink data transfer to the network.
- **PFI Indicator (PI).** This parameter indicates the presence or the absence of the PFI field in the data block.

FIGURE 9.7

Uplink RLC Header Type 1, 2 and 3



- **Stall Indicator (SI).** The RLC "transmit window" stores the acknowledgment information of the data blocks transmitted from the sending side. In case acknowledgments are not received from the receiving side, the sending side can still send the data blocks till the transmit Window Size (WS) is the same as the WS for the data transfer. In such a case, the window is said to be "stalled" that is indicated by the SI value of "1." The value "0" indicates that the window is not stalled and hence the data transfer can continue.
- **Retry (R).** This parameter is used by the mobile station to indicate whether the channel request message has been sent once or more in the most recent access requests.
- **Resent Block Bit (RSB).** This single-bit field indicates whether the RLC data block contained within the EGPRS data block has been transmitted before or not. The value "0" indicates that the blocks are being transmitted for the first time, while value "1" indicates retransmissions.
- **Split Block Indicator (SPB).** This field is used only in header type 3. It indicates whether the first part or the second part of the block is being retransmitted.

9.6 LINK ADAPTATION

The purpose of link adaptation is to select the coding scheme for a data transfer according to the radio conditions. If the radio conditions are good, a coding scheme with high data rate is selected leading to a higher throughput. On the other hand, if the radio conditions are bad, a coding scheme with low data rates is selected, leading to a low throughput. The radio conditions are hence continuously monitored by the network in order to make decisions about the coding scheme to be used.

In case of EGPRS, the link adaptation uses the concept of families that was discussed in the previous sections. This concept is used for the segmentation of the data blocks during the retransmissions. Another concept that is used is the transfer of the RLC data blocks using the selective type Automatic Repeat Request (ARQ) mechanism or by a selective type II hybrid ARQ mechanism within one TBF. These mechanisms can be used for a data transfer that operates in the RLC acknowledged mode. The link adaptation is implemented in the PCU.

The following sections discuss the procedures for making the measurements on the radio interface. These measurements enable the analysis of the radio conditions and hence the network can decide on the coding scheme to be used. The next section discusses the link adaptation mechanism. The last section introduces the concept of incremental redundancy.

9.6.1 Measurements for Link Adaptation

During a data transfer, the radio blocks transmitted on the air interface may not be decoded properly by the receiving end. This may be due to bad radio interface conditions. In this case, the receiving side may send a negative acknowledgment for the radio block or may not send any acknowledgments also. In both the cases, the sending side would retransmit the data block (in the second case, the retransmission would take place after the sending window gets stalled). However, in order to ensure the successful reception of the retransmitted data block, the transmission should be done in accordance with the prevalent radio conditions. This means that the coding scheme to be used for the retransmission should be such that the block can be successfully decoded at the receiving end.

In GPRS, if a block has to be retransmitted, the same coding scheme is used as in the initial transmission of the radio block. Thus, it is possible that the receiver might not be able to decode the block and hence the TBF may be released if this coding rate is not appropriate for the radio conditions.

In order to solve this problem, the coding scheme for a data transfer should be decided carefully. This decision is based on the radio conditions that are determined from the measurement reports. The measurements are performed by the MS in case of a downlink data transfer and by the BTS in case of an uplink data transfer. The measurement reports are analyzed and the decision on the suitable coding scheme is made.

In the case of GPRS, the decision is based on the parameter receive quality, RxQual. This parameter ranges between values 0–7 that correspond to a range of Bit Error Rate (BER) values. These values are calculated on a block basis (four bursts) and then averaged. The RxQual then corresponds to this averaged value of BER. However, this value is more suitable for the speech services and not the packet data transfers.

In EGPRS, a new parameter has been introduced to determine the radio conditions and hence decide the coding scheme. This parameter is the Bit Error Probability (BEP). It provides more information about the radio conditions. This parameter and its measurement are the topics of discussion in the present section.

The BEP is determined on a burst-by-burst basis and is then used to determine two important parameters, the CV_BEP and the MEAN_BEP. The CV_BEP indicates the quality from one burst to the other and thus reflects the effect of frequency hopping and the interleaving loss/gain. The MEAN_BEP indicates the C/I (carrier/interference) ratio and the velocity. It is evaluated by averaging the BEP values calculated on all the four bursts of a radio block.

Table 9.4 lists the mapping between the possible MEAN_BEP values and the range of the actual BEP values for both the GMSK and the 8-PSK modulation.

TABLE 9.4

Mapping Between the MEAN_BEP and the Actual BEP Values

Reported MEAN_BEP	*Range of log 10 (Actual BEP) for GMSK*	*Range of log 10 (Actual BEP) for 8-PSK*
MEAN_BEP_0	> -0.60	> -0.60
MEAN_BEP_1	–0.70 to –0.60	–0.64 to –0.60
MEAN_BEP_2	–0.80 to –0.70	–0.68 to –0.64
MEAN_BEP_3	–0.90 to –0.80	–0.72 to –0.68
MEAN_BEP_4	–1.00 to –0.90	–0.76 to –0.72
MEAN_BEP_5	–1.10 to –1.00	–0.80 to –0.76
MEAN_BEP_6	–1.20 to –1.10	–0.84 to –0.80
MEAN_BEP_7	–1.30 to –1.20	–0.88 to –0.84
MEAN_BEP_8	–1.40 to –1.30	–0.92 to –0.88
MEAN_BEP_9	–1.50 to –1.40	–0.96 to –0.92
MEAN_BEP_10	–1.60 to –1.50	–1.00 to –0.96
MEAN_BEP_11	–1.70 to –1.60	–1.04 to –1.00
MEAN_BEP_12	–1.80 to –1.70	–1.08 to –1.04
MEAN_BEP_13	–1.90 to –1.80	–1.12 to –1.08
MEAN_BEP_14	–2.00 to –1.90	–1.16 to –1.12
MEAN_BEP_15	–2.10 to –2.00	–1.20 to –1.16
MEAN_BEP_16	–2.20 to –2.10	–1.36 to –1.20
MEAN_BEP_17	–2.30 to –2.20	–1.52 to –1.36
MEAN_BEP_18	–2.40 to –2.30	–1.68 to –1.52
MEAN_BEP_19	–2.50 to –2.40	–1.84 to –1.68
MEAN_BEP_20	–2.60 to –2.50	–2.00 to –1.84
MEAN_BEP_21	–2.70 to –2.60	–2.16 to –2.00
MEAN_BEP_22	–2.80 to –2.70	–2.32 to –2.16
MEAN_BEP_23	–2.90 to –2.80	–2.48 to –2.32
MEAN_BEP_24	–3.00 to –2.90	–2.64 to –2.48
MEAN_BEP_25	–3.10 to –3.00	–2.80 to –2.64
MEAN_BEP_26	–3.20 to –3.10	–2.96 to –2.80
MEAN_BEP_27	–3.30 to –3.20	–3.12 to –2.96
MEAN_BEP_28	–3.40 to –3.30	–3.28 to –3.12
MEAN_BEP_29	–3.50 to –3.40	–3.44 to –3.28
MEAN_BEP_30	–3.60 to –3.50	–3.60 to –3.44
MEAN_BEP_31	< –3.60	< – 3.60

The MS makes the measurements for the downlink data transfers and sends the MEAN_BEP and the CV_BEP values for the blocks that it has received since the last measurement was sent. The values may be for GMSK and/or 8-PSK depending on which modulation is used for the transmission. Hence, the mobile sends the parameters GMSK_MEAN_BEP, GMSK_CV_BEP for the GMSK modulation and the 8-PSK_MEAN_BEP and the 8-PSK_CV_BEP for the 8-PSK modulation. Additionally, the MS also reports the measurement on a per timeslot basis, i.e., the MEAN_BEP_TNx for all the timeslots that have been allocated by the network.

The CV_BEP, coefficient of the channel quality, is calculated from the mean BEP using the following formula:

$$\mathrm{CV\ (BEP)} = \frac{\mathrm{std\ (BEP)}}{\mathrm{mean\ (BEP)}}$$

where std (BEP) is the standard deviation of the BEP calculated within a radio block. It is calculated using the following formula:

$$\mathrm{std\ (BEP)} = \sqrt{\frac{1}{3} \cdot \sum_{i=1}^{4} (\mathrm{BEP}_i - \mathrm{mean\ (BEP)})^2}$$

Table 9.5 lists the mapping for the coefficient of variation of the channel quality. This table holds good for both the GMSK and the 8-PSK modulation.

Table 9.5

Mapping Between the Reported and the Actual Values of CV_BEP

Reported CV_BEP	*Range of CV_BEP*
CV_BEP_0	2.00 > CV_BEP > 1.75
CV_BEP_1	1.75 > CV_BEP > 1.50
CV_BEP_2	1.50 > CV_BEP > 1.25
CV_BEP_3	1.25 > CV_BEP > 1.00
CV_BEP_4	1.00 > CV_BEP > 0.75
CV_BEP_5	0.75 > CV_BEP > 0.50
CV_BEP_6	0.50 > CV_BEP > 0.25
CV_BEP_7	0.25 > CV_BEP > 0.00

9.6.2 Link Adaptation Mechanism

The aim of the link adaptation is to use the measurements made on the radio interface in order to decide the coding scheme to be used for a data transfer. The concept of MCS families helps in adapting the data rates to the radio interface conditions. The MCS families allow the retransmission of a higher coding rate block using lower coding rates. The retransmissions occur in the RLC acknowledged mode of operation. If a transmitted block is not decoded correctly, the block has to be retransmitted by the sending side. During retransmission, a lower data rate can be chosen. The concept should be clear from the Fig. 9.8.

FIGURE 9.8

Retransmission of RLC Data Block Containing Two RLC Data Parts

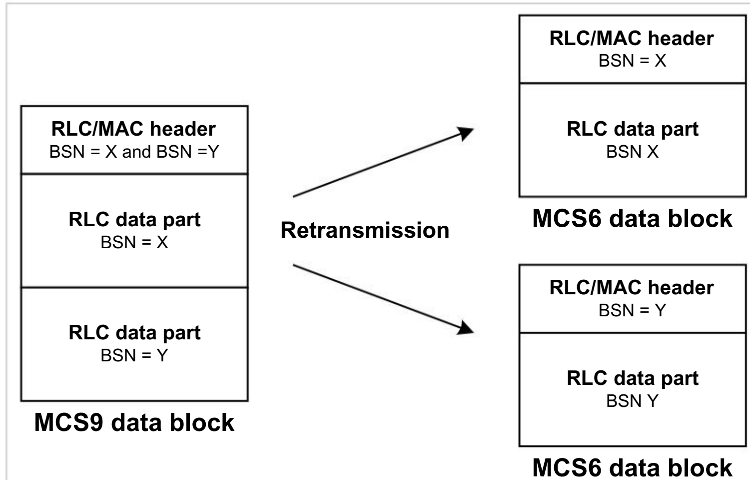


Figure 9.8 depicts the case when an MCS9 data block is retransmitted. MCS9 belongs to family A, which also consists of MCS6 and MCS3 coding schemes. The retransmission of the MCS9 block is thus possible by splitting the radio block into two data blocks of MCS6 or four data blocks of MCS3. An MCS9 data block consists of two RLC data blocks of different BSN values, X and Y, as depicted in the figure. When this block is split into two MCS6 data blocks, one of these data blocks consists of the first BSN value X and the other data block consists of the second BSN value Y. Thus, the RLC data blocks carried in the radio block have been split into two different radio blocks each consisting of a single RLC block.

As mentioned earlier, it is also possible to transmit the MCS9 block using four blocks of MCS3 coding scheme. In this case, each RLC data block in the MCS9 block would be split into two MCS3 data blocks. The situation in this case would be the same as the example considered next.

It is also possible that a block that carries a single RLC block may be split into two radio blocks. For example, such a case may happen when an MCS5 data block has to be retransmitted. This MCS5 block may be split into two MCS2 data blocks. In this case, both the MCS2 blocks would carry the same value of the BSN, as depicted in Fig. 9.9. The order of the blocks is determined by the Split Block Indicator field, SPB, in the RLC data block. This mechanism of splitting the radio block into multiple data blocks of lower data rate is known as "segmentation."

9.6.3 Incremental Redundancy

The principle of IR involves the retransmissions of the data block till the block is correctly decoded by the receiver. After the receiver fails to decode a particular data block when it has been transmitted for the first time, it stores the soft bits at the output of the demodulator. These soft bits are used along with the bits that are received on the retransmissions of the data block. The bits in the retransmission may be completely different from the bits of the first

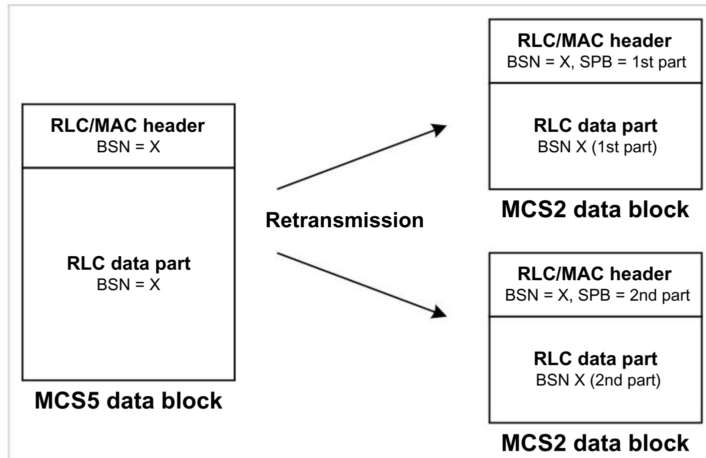


Figure 9.9

Retransmission of RLC Data Block Containing Single RLC Data Parts

transmission. The retransmission bits are determined by the puncturing scheme that is used. If the puncturing schemes in the two transmissions are disjoint, then completely different bits are transmitted in the transmissions. Thus, the number of bits that are transmitted after the second retransmission is twice as high.

On the second transmission, the receiver would use the soft decisions that had been made after the first transmission and would also use the bits transmitted in the second transmission. This helps in greater efficiency during decoding.

The biggest problem with the IR is the memory requirement for storing the soft decisions after every retransmission. The memory available in the mobile for IR usage has to be carefully determined. The IR is optional at the network end, but it is mandatory at the MS.

9.7 MAC Procedures in EDGE: Uplink TBF Establishment

With the introduction of EGPRS, there were some modifications to the procedures for TBF management. This section and the next section discuss the procedures for the TBF establishment and release.

The procedure for an uplink TBF establishment involves the access request from the mobile requesting resources from the network and the assignment message from the network allocating the resources to the mobile. A mobile can request for the establishment of an EGPRS TBF only if EGPRS is supported in the serving cell. The information about the EGPRS capability of the cell is broadcast in the GPRS cell options in the System Information 13.

A new access request message has been defined for initiating the EGPRS data transfer, the EGPRS packet channel request. This message can be sent on the RACH or the Packet Random Access Channel (PRACH), in case the packet channel have been defined in the serving cell. However, the support for this message by the BTS is optional and is indicated in the GPRS cell options in the System Information 13. The message has the same format as the GPRS 11-bit packet channel request message and provides the same information.

It is possible for an MS to specify its EGPRS capabilities while sending the EGPRS packet channel request message using two training sequences. These training sequences have been defined for EGPRS in addition to the old ones that already exist. One of the training sequences indicates that the mobile is EGPRS capable and supports 8-PSK in both the uplink and the downlink direction, whereas the second training sequence indicates that the MS is EGPRS capable and supports 8-PSK in the downlink direction only. The new training sequences can be used along with the old values on the RACH and the PRACH. These training sequence values allow the network to differentiate between the channel request, packet channel request, and the EGPRS packet channel request message.

In case the BTS does not support the EGPRS packet channel request message, the mobile would not be able to specify its EGPRS capabilities in the access message. In this case, the mobile may send a channel request message on the RACH or the packet channel request message on the PRACH. The EGPRS capabilities are specified in the packet resource request message that is sent as part of the two-phase access procedure.

Just as in GPRS, if a TBF has to be established in RLC unacknowledged mode, the two-phase access procedure is mandatory. For a one-phase access procedure, the RLC mode is acknowledged mode by default.

9.7.1 One-phase Access Procedure for TBF Establishment in EGPRS Mode

Just as in case of GPRS, it is possible to establish the EGPRS TBF using the one-phase access as well as the two-phase access procedure. As discussed earlier, the MS can specify its EGPRS capabilities while sending the access request to the network. Hence, if the EGPRS packet channel request message is supported in the serving cell, the MS can establish a TBF using the one-phase access procedure. In this case, the MS initiates the TBF establishment by sending an EGPRS packet channel request to the network on the CCCH or the PCCCH, if the latter is present in the serving cell. The one-phase TBF establishment procedure on the CCCH and the PCCCH is discussed in the following sections.

9.7.1.1 One-phase Access Procedure on the CCCH

If the packet channels are not defined in the serving cell, the mobile sends the EGPRS packet channel request message on the RACH. The message contents of the EGPRS packet channel request are discussed next. The message contents are also listed in Table 9.6.

- **Establishment cause.** This field indicates the reason for the establishment of the TBF. The values encoded for the establishment cause are the same as the 11-bit packet channel request. The allowed establishment causes and the encoded values are listed in the Table 9.6.

- **Priority.** This field indicates the radio priority of the TBF being established.
- **Multislot class.** This field indicates the multislot class of the mobile, which determines the MS capabilities to transmit and receive on different combinations of the PDCH. The EGPRS multislot class of a mobile can be different from the GPRS multislot class.

Number of blocks. Some random bits are also sent in the access message to reduce the chances of two mobiles sending the same contents of the access message and hence leading to a contention. The number of random bits in the message depends on the establishment cause. The exact number of random bits available with different establishment causes is given in Table 9.6.

Table 9.6

Contents of the EGPRS Packet Channel Request Message

Establishment Cause	*Prefix*	*Multislot Class*	*Priority*	*Random Bits*
One-phase access request	0	5 bits	2 bits	3
Two-phase access request	110000	N	2 bits	3
Signaling	110011	N	N	5
One-phase access request in RLC unack mode	110101	N	N	5
Dedicated channel request	110110	N	N	5
Emergency call	110111	N	N	5

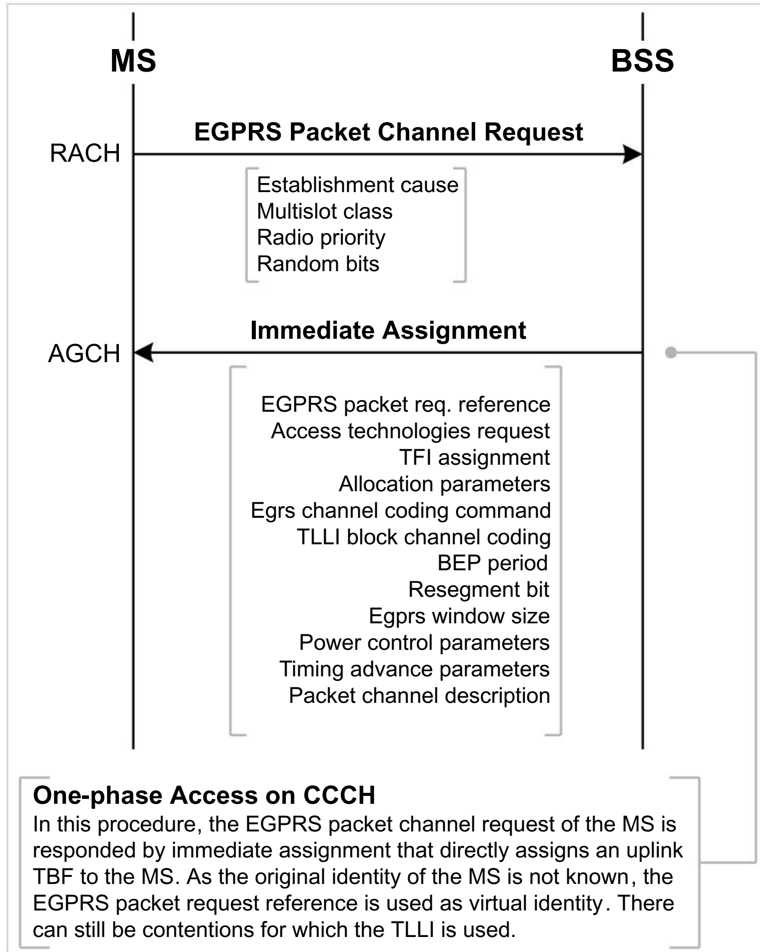
On receiving the access request, the network allocates the radio resources to the mobile. The allocation is done using the immediate assignment message that is transmitted on the AGCH in the downlink direction. The message exchange is depicted in Fig. 9.10.

The mobile for which the assignment is intended is identified by the request reference that is encoded in the assignment message. The request reference consists of the packet channel request contents and the frame number on which the access message was received by the network. If the Random Access (RA) in the request reference field of the message indicates the presence of an extended RA, the mobile uses the extended RA information to identify the immediate assignment message is in response to the EGPRS packet channel request. If on the other hand, an extended RA value is not present, it is assumed that the assignment message does not correspond to an EGPRS packet channel request. This assignment message contains the packet uplink assignment message in the IA rest octets field. The following information is transmitted in this message:

- **Access technologies request.** This field indicates the different access technologies that are requested from the MS. These technologies are listed as per the priorities such that the most important type comes first, then the less important, and so on. The MS also replies using the same order. Out of the three access types in GSM 900, GSM P, GSM E, and GSM R, the network can request any one of the access types.
- **TFI assignment.** This field indicates the TFI assigned for the TBF. This value is used to identify the uplink TBF.

FIGURE 9.10

Uplink TBF Establishment using One-phase Access Procedure on CCCH



- **Allocation parameters.** Different allocation schemes can be used to allocate resources for the TBF. Depending on the scheme, different allocation parameters are encoded in the assignment message. In case the resources have been allocated using the dynamic allocation scheme, the USF is sent to the MS. On the other hand if the fixed allocation is being used, the allocation bitmap is transmitted in the message.
- **EGPRS channel coding command.** This field indicates the EGPRS modulation and the coding scheme that is to be used for the data transfer.
- **TLLI block channel coding.** During the data transfer, there is a contention resolution period during which the data blocks are transmitted with the Temporary Logical Link Identifier (TLLI) of the MS encoded in the blocks. This field indicates the coding scheme that should be used to encode such data blocks. This coding scheme may be MCS1 or it may be the same as the EGPS channel coding command.

- **BEP period 2.** This field contains a constant value that is used for filtering the channel quality measurements in EGPRS.
- **Resegment bit.** This field indicates whether the retransmitted RLC data blocks would be resegmented as per the commanded MCS or not. A value "0" indicates that the blocks shall not be segmented, whereas the value "1" indicates that the data blocks would be resegmented as per the commanded coding scheme.
- **EGPRS WS.** This field indicates the WS to be used for the EGPRS data transfer. The network sets the WS according to the number of timeslots that are allocated to the TBF. The WS can be set independently for each TBF in the uplink and the downlink direction. A minimum and maximum WS is defined for every possible multislot capability. At the MS side, the maximum WS corresponding to its multislot capability is supported. On the other hand, on the network side, it is preferable that the maximum window size be allocated to the MS for a TBF. However due to optimization of the users and the resources, a smaller window may be selected. It should be noted that the WS once allocated to a TBF can be increased during the data transfer, but it cannot be reduced. This is important in order to avoid dropping of data blocks from the window.
- **Power control parameters.** These parameters are used by the MS to determine the Tx power level.
- **Timing advance parameters.** This field may contain the initial timing advance, calculated on the basis of the access request received from the MS or it may include the timing advance index and the timeslot to be used in case the continuous timing advance procedure is being used.
- **Packet channel description.** This field provides information about the Training Sequence Code (TSC) and the timeslot number to be used. Also, this field contains details of the hopping mode and the frequencies to be used.

Due to the size constraints of the immediate assignment message, only one timeslot can be allocated to the mobile irrespective of its multislot class.

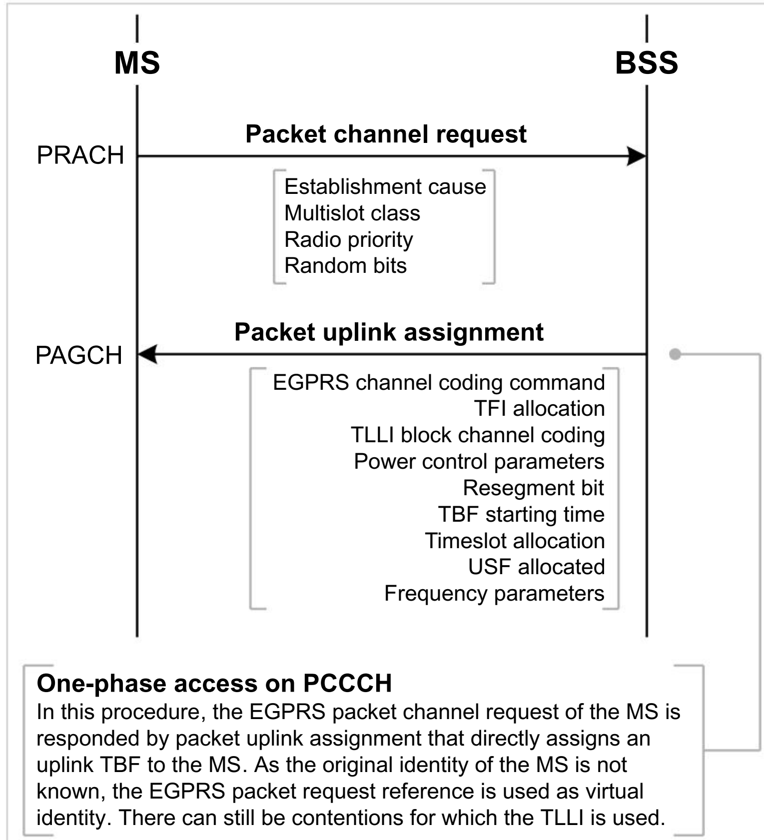
9.7.1.2 One-phase Access Procedure on the PCCCH

When the packet channels are defined in the serving cell and the EGPRS packet channel request message is supported, the mobile initiates the TBF establishment by sending the EGPRS packet channel request on the PRACH (PCCCH). The contents of the message are the same as discussed in the last section. On receiving the access message, the network allocates resources to the mobile in the packet uplink assignment message that is transmitted on the PAGCH. The signaling for the procedure is depicted in the following Fig. 9.11. Please note that the contents of the assignment message in the figure are as per the dynamic allocation scheme.

The contents of the assignment message are the same as discussed in the previous section. However in this case, more than one timeslot can be allocated to the mobile. The timeslot allocation is done depending on the EGPRS multislot class of the mobile. The assignment message contains the USF for each of the allocated timeslot in case dynamic allocation of

FIGURE 9.11

Uplink TBF Establishment using One-phase Access Procedure on PCCCH



resources is done. On the other hand if the fixed allocation is used, the timeslot allocation, the timeslot for the downlink control, and the allocation bitmap are transmitted in the assignment message.

9.7.2 Two-phase Access Procedure for TBF Establishment in EGPRS Mode

The two-phase access procedure is used in case the EGPRS packet channel request is not supported by the network. Even when this message is supported by the network, the procedure is used when the TBF has to be established in unacknowledged mode. The network may also force the two-phase access procedure for the TBF establishment by sending a single-block allocation in response to the access request. This is done when the network wants additional information about the MS capabilities in order to optimize the resource utilization.

9.7.2.1 Two-phase Access Procedure when the EGPRS Packet Channel Request Is not Supported

In this case, the mobile sends the channel request on the RACH or the packet channel request on the PRACH in order to request for the TBF establishment. These access request messages do not provide any information about the EGPRS capabilities of the mobile to the network. In response to the access request, the network allocates a single block to the mobile in the immediate assignment message on the CCCH or the packet uplink assignment on the PCCCH. The MS sends its EGPRS capabilities on the allocated block in the packet resource request message. The network then allocates the resources to the mobile according to the EGPRS capabilities of the mobile. The data transfer then progresses on these assigned resources.

9.7.2.2 Two-phase Access Procedure when the EGPRS Packet Channel Request Is Supported

When the EGPRS packet channel request message is supported in the serving cell, the two-phase access procedure is requested by the mobile only when the RLC mode of the TBF is unacknowledged mode. The signaling message exchange for this TBF establishment is depicted in Fig. 9.12.

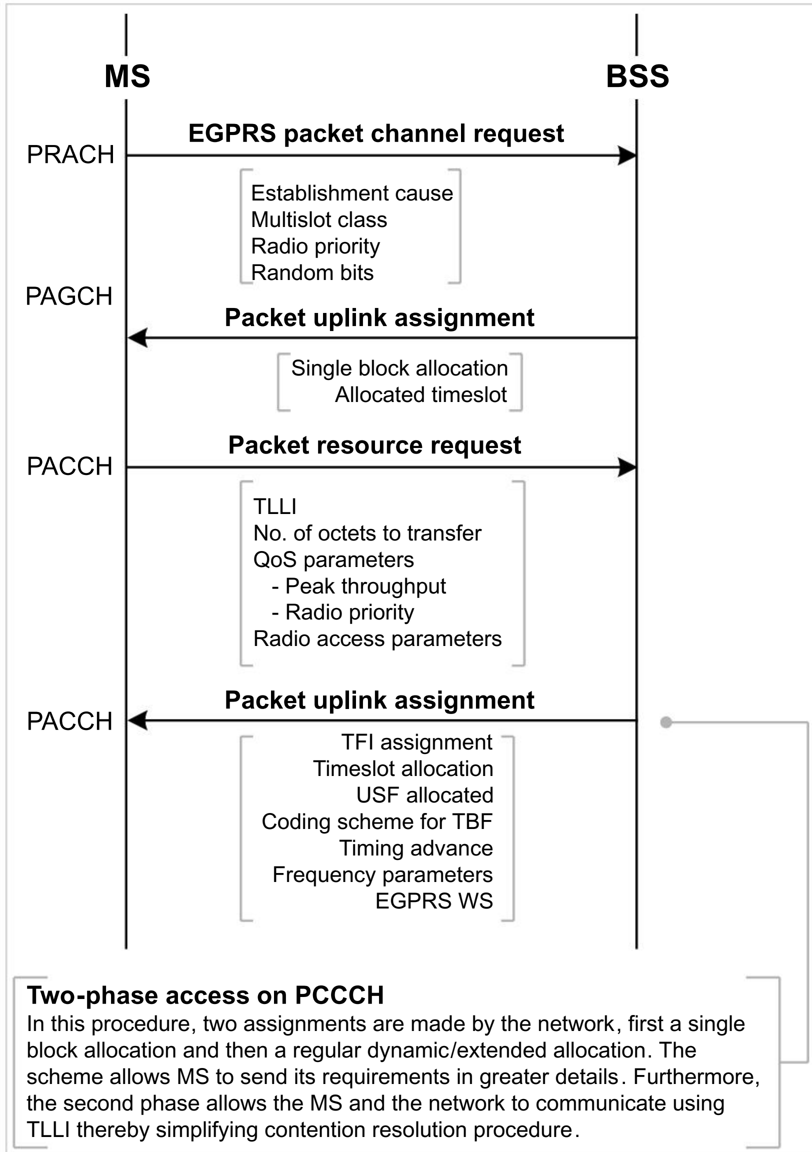
The mobile initiates the TBF establishment by sending the channel request on the RACH or the packet channel request on the PRACH. The establishment cause in these messages would indicate a two-phase access request. In response to the access request, the network allocates uplink resources in the immediate assignment message or the packet uplink assignment message. The network may allocate one or two uplink resources in order to receive the information about the mobile capabilities. In the first block that is allocated by the network, the mobile sends the packet resource request message. In case the second block has also been allocated by the network, the mobile sends the additional MS radio access capabilities message on this block.

The packet resource request message consists of the following information:

- **Access type.** This field indicates the reason for requesting the access. The reason may be two-phase access, MM procedure, cell update, or page response.
- **TLLI.** This field identifies the MS.
- **MS radio access capability.** This field provides information about the radio capabilities of the mobile to the network. The information provided is the band of frequency that the mobile supports.
- **Channel request description.** This field is used by the MS to request the uplink resources. The mobile provides information about the peak throughput, radio priority, RLC mode, and the number of RLC octets to be transferred. This information is used by the network to allocate appropriate resources.
- **Change mark.** The mobile stores the PSI2_Change_Mark or the SI13_Change_Mark (in case PBCCH is not defined). This allows the mobile to store information about the consistent set of the SI message that the mobile is using.

FIGURE 9.12

Uplink TBF Establishment using Two-phase Access Procedure



- **C value.** The field represents measurements made by the MS.
- **Signal variance, SIGN_VAR.** This field represents the signal variance calculated by the mobile.
- **Interference Level, I_LEVEL.** These are represented as the I_Level measured for all the timeslots. The interference levels are spread across 16 possible values that map to certain range of C_Value.

9.7.3 Downlink TBF Establishment

The procedure for the establishment of the downlink TBF in the EGPRS mode is the same as the procedure in the GPRS mode. The immediate assignment message is transmitted by the network on the CCCH. If the packet channels are defined in the serving cell, then the packet downlink assignment message is transmitted on the PCCCH. In case of the EGPRS, some more parameters are present in the message. These parameters are

- **EGPRS WS.** This field indicates the WS to be used for the EGPRS data transfer. The network decides the WS on the basis of the number of timeslots that are allocated.
- **Link quality measurement mode.** This field determines the behavior of the MS for the measurement reporting. Based on this parameter, the mobile may report the interference measurements or BEP measurement on each allocated timeslot. Also, this determines whether the mobile reports only the BEP measurements or the interference measurements as well.
- **BEP period 2.** This field contains a constant that helps in filtering the channel quality measurements for EGPRS.
- **Packet extended timing advance.** This field provides information about the timing advance to be applied by the mobile. This information is in addition to the timing advance information provided by the network in the packet timing advance IE.

Once the MS receives the downlink assignment message, it starts monitoring the allocated PDCH for the data blocks. The data blocks are identified by the downlink TFI value that is encoded in the blocks.

9.8 MAC Procedures in EDGE: TBF Release

The procedures for the TBF release are similar to the scenarios for GPRS.

In the uplink, TBF is released once the MS indicates that all the data blocks have been transmitted as part of the TBF. This is indicated using the countdown procedure. In the uplink data blocks, the CV is set to 0 in order to indicate the last data block of the TBF. When the network receives such a data block, it sends the packet uplink ack/nack message with the final ack indication bit set to 1 if all the blocks of the TBF have been received correctly. Before releasing the TBF, the network polls the MS in order to ensure that the acknowledgment message has been received by the MS. The MS sends a packet control acknowledgment in response to the polling after which the network releases the uplink TBF.

The downlink TBF release is simpler as compared to the uplink TBF release. This is because the network has information regarding the number of blocks to be transferred as part of the downlink TBF. In the last data block of the downlink TBF, the network sets the Final Block Indicator (FBI) to 1, indicating the last data block to the MS. After the transmission of this last data block, the network polls the MS for the packet downlink ack/nack message. If the MS has received all the blocks correctly, it sends the acknowledgment with the final ack indication set to 1. On receiving the acknowledgment, the network starts a timer and releases the TBF on the expiry of this timer.

9.9 Modifications to the RLC Procedures

As discussed before, the new coding schemes and hence the higher throughputs have led to some changes in the RLC/MAC data blocks. In addition to this, certain other parameters have also been affected. One such parameter is the WS. While in case of GPRS, the WS was fixed and had the value 64, in case of EGPRS, the size is variable and depends on the number of timeslots allocated for a TBF. The modified WS also led to change in the mechanism of acknowledging the radio blocks on the air interface.

9.9.1 Window Size

It is the function of the RLC layer to transfer data using the RLC data blocks. The RLC protocol may operate in the acknowledged or the unacknowledged mode. In the acknowledged mode, the transmitter and the receiver maintain information about the data blocks that have been transmitted. The receiver sends information about the successful decoding of the blocks. In case the blocks are not received successfully, a retransmission is requested from the transmitter. The status of the blocks is maintained using the concept of windows. In case all the blocks lying in the window are unacknowledged, the transmission cannot proceed with the higher sequence numbers. Or in other words, the window gets stalled. In the case of GPRS, there is a standard WS of 64.

However in the case of EGPRS, two RLC data blocks are sent in the same radio block. Hence, the number of transmitted blocks is twice that of GPRS. In this situation, the WS of 64 is insufficient. A window of this size would lead to the stalling of the RLC protocol. This led to the requirement for a window of larger size.

The WS is determined at the time of the TBF establishment in the assignment message. The size is encoded in a 5-bit field of the assignment message, the EGPRS WS. The values of the field represent the WS in the range 64–1024. The size can vary in steps of 32.

The size depends on the number of timeslots that are allocated for a particular TBF. For every number of allocated timeslots, a maximum value of the WS has been defined. It is mandatory for a mobile to support the maximum WS corresponding to its multislot class. The minimum WS to be supported is 64. Table 9.7 lists the range of the WS corresponding to the number of allocated timeslots.

It is possible to change the size during the TBF. However, it is not possible to reduce the WS for an ongoing TBF. This is because if the size is reduced, it would lead to a loss of

Table 9.7

Maximum and Minimum Window Size for the Allocated Timeslots

No. of allocated timeslots	1	2	3	4	5	6	7	8
Minimum WS	64	64	64	64	64	64	64	64
Maximum WS	192	256	384	512	640	768	896	1024

information about the data blocks that lie in the reduced window. For instance, consider a TBF for which the WS was initially 128, and during the data transfer the size was reduced to 64. In this case, the information of about 64 blocks would be lost. This is an unacceptable situation.

9.9.2 Acknowledgment Mechanism

As discussed earlier, the RLC protocol may work in the acknowledged or the unacknowledged mode. In the acknowledged mode, the receiver end sends acknowledgments about the data blocks that had been transmitted as part of the data transfer. The mobile sends the acknowledgments for the downlink TBF in the packet downlink ack/nack message. Similarly, the network acknowledges the data blocks in the uplink direction in the packet uplink ack/nack message. These messages consist of acknowledgment bitmaps that indicate whether a block has been successfully received or not.

Each bit provides information about a data block. A "0" indicates unsuccessful reception of the data block, whereas a "1" indicates successful reception. As already discussed, control messages are sent using the CS-1 coding scheme that allows a maximum message length of 23 bytes. If we consider the maximum WS 1024, WS of an EGPRS data transfer, then it is clear that this size of the message is not sufficient to provide information about the data blocks that are transmitted in an EGPRS data transfer.

This leads to the requirement of a new acknowledgment mechanism for the EGRPS data transfers. The requirement is to represent a sequence of 1s or 0s more efficiently. In order to do so, the acknowledgment bitmap has been compressed. The compression is done by encoding the sequence of 1s or 0s by a code that provides information about the number of successive 1s or 0s. The compressed acknowledgment bitmap is carried in the EGPRS ack/nack description IE of the acknowledgment message.

The description provides the sequence number of the blocks that constitute the start and end of the window whose information is being provided in the message. In addition to this, the length of the acknowledgment bitmap is also provided in the message. Also, it is important to know whether the first code represents the sequence of 1s or 0s. The bitmap starting color code provides this information. A value 1 indicates that the first code represents a sequence of 1s, and a value 0 indicates a sequence of 0s.

To understand this, consider a case when an MS is involved in a downlink data transfer, such that the network has transmitted 80 blocks. The MS has to send the acknowledgment of these blocks to the network, where the successful (bit 1) and the unsuccessful (bit 0) reception of the data blocks is represented as follows:

11111111111100000001111111111111111110000001111111111111111000000000000111111111

In other words, the sequence of successful and unsuccessful receptions is as follows:

Successful: 12 blocks represented by code word "001000"

Unsuccessful: 7 blocks represented by code word "00011"

Successful: 18 blocks represented by code word “0100111”

Unsuccessful: 6 blocks represented by code word “0010”

Successful: 15 blocks represented by code word “110101”

Unsuccessful: 12 blocks represented by code word “0000111”

Successful: 10 blocks represented by code word “00111”

The compressed acknowledgment bitmap that is transmitted by the mobile is as follows:

0010000001101001110010110101000011100111

The bitmap starting color code in this case is encoded as “1” indicating that the first coded value in the bitmap represents a sequence of 1s.

9.9.3 Polling Mechanism

It is possible that the receiver may not be able to acknowledge the entire window in a single message, even though the compression technique has been introduced. In case of the downlink data transfer, the network uses the polling mechanism to solve this problem.

The fundamental principle behind the polling procedure remains the same in EGRPS as in GPRS. The network polls the mobile to provide information about the transmitted blocks in a single block control message. A field RRBP indicates the radio block in which the mobile should respond with the acknowledgment message.

Some enhancements have been introduced in this procedure for EGPRS. The acknowledgment information is sent across multiple messages. The acknowledgment bitmaps carried in these messages are partial bitmaps. The bitmaps may be of the following types:

- **First Partial Bitmap (FPB).** This bitmap provides information about the blocks starting from the start of the window.
- **Next Partial Bitmap (NPB).** This bitmap would provide information about the blocks starting from the highest BSN that has been reported in the previous acknowledgment bitmap.

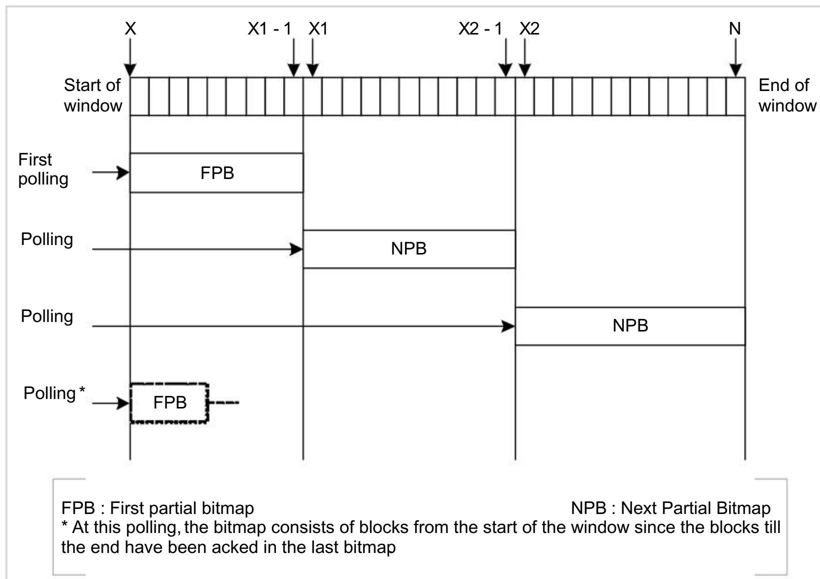
It is possible for the network to specify the exact starting point in the window from which the information about the acknowledgments is required. The mobile then sends the acknowledgment to the network. In addition to the acknowledgments, the packet downlink ack/nack message that is sent in the uplink may also contain the channel quality report. The network controls the contents of the message using the EGPRS Supplementary/Polling (E S/P) field that is encoded in the downlink data blocks. The following behavior is defined depending on the value of the E S/P field:

- A value 01 indicates that the EGPRS packet downlink ack/nack contains the FPB.
- A value 10 indicates that the EGPRS packet downlink ack/nack contains the NPB.
- A value 11 indicates that the EGPRS packet downlink ack/nack message contains the NPB, and if there is any space available in the block, then the measurement report is carried in the message.

Consider a case when the network has transmitted *N* blocks as part of a downlink data transfer. Let the window consist all the *N* blocks and the start of the window is from BSN "X." The polling mechanism can be used by the network to obtain the acknowledgment information from the WS, as depicted in Fig. 9.13.

FIGURE 9.13

Illustration of the Polling Mechanism



As part of the polling, the network requests the acknowledgments for the blocks from the start of the window. The mobile, hence, sends the FPB in the acknowledgment message. In this case, let us assume that the bitmap contained information till the sequence number X1–1. When the network polls the second time for the NPB, the mobile sends the information about blocks starting from X1 to a sequence number, say X2–1. On the next polling, the mobile sends the NPB providing information about the blocks from sequence number X2 to the end of the window. Once the entire window has been acknowledged, the next polling results in the mobile sending information from the start of the window in the acknowledgment message.

9.10 MULTIPLEXING OF GPRS AND EGPRS MOBILES

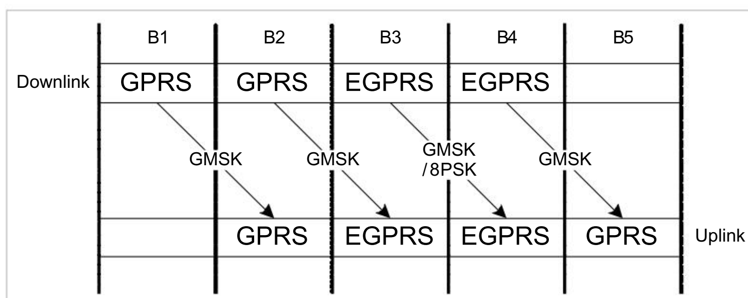
One of the main requirements of the EGPRS system is that it should be possible to multiplex a GPRS and an EGPRS mobile on the same physical channel. This multiplexing may be in the uplink or in the downlink direction.

The downlink direction does not cause any problem for the system since the transmission of the data blocks is done using the coding schemes applicable to the mobiles. In other words, the network would transmit the data blocks to a GPRS mobile using the GPRS coding schemes, whereas for an EGPRS mobile, the EGPRS coding schemes are used.

The problem, however, lies in the multiplexing of the uplink TBF between a GPRS and an EGPRS mobile. In the case of dynamic allocation of resources, the USF assigned to a GPRS mobile may be encoded in the header of a data block that is meant for the EGPRS mobile. The GPRS mobile, however, would not be capable of decoding a data block transmitted using the 8-PSK modulation. As a result, the network faces the restriction to transmit the USF in a data block transmitted using GMSK modulation when it wants to command a GPRS mobile to transmit on the next radio block. Figure 9.14 indicates the modulation scheme that is used in the downlink depending on whether the next radio block has to be used by a GPRS or an EGPRS user.

FIGURE 9.14

Multiplexing of GPRS and EGPRS in Uplink Using USF Granularity 1



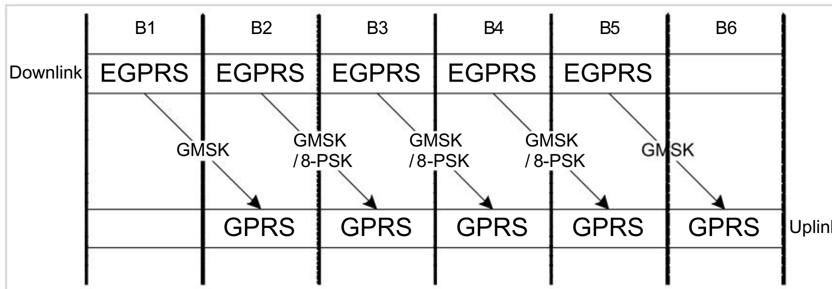
Similarly, when there are data transfers in the opposite directions on the same PDCH, there are constraints on the coding scheme to be used. For instance, consider that an EGPRS data transfer in the downlink and a GPRS data transfer in the uplink are multiplexed on the same physical channel. In this case also, the USF for a GPRS mobile cannot be carried in a downlink block transmitted using 8-PSK modulation. This, hence, poses a restriction on the coding schemes that can be used for the EGPRS data transfer. The GMSK coding schemes, MCS1–MCS4, would not provide a high data rate, and thus the performance would have to be compromised in this case.

The concept of USF_Granularity can be of some help in such cases. Using this concept, the mobile can be instructed to transmit over four consecutive data blocks on detecting its allocated USF in the downlink. The USF granularity is indicated to the mobile during the resource assignment for the uplink TBF. The network can transmit the allocated USF of the GPRS mobile using a GMSK downlink block and then transfer 8-PSK blocks over the next three radio blocks in the downlink. The concept is depicted in Fig. 9.15. This would result in one out of four downlink blocks to be sent using GMSK modulation and hence a better performance as compared to the previous case.

Another point to be noted is that the USF allocated to the GPRS mobile is transmitted only in the downlink block B1, by using GMSK modulation. The USF is not sent in the blocks B2, B3, and B4. These carry some unassigned values of USF.

FIGURE 9.15

Multiplexing of GPRS and EGPRS in Uplink Using USF Granularity 4



Conclusion

This chapter provides an overview of EDGE. The network architecture is the same as the GPRS. The user is provided higher data rates and better performance. This is mainly driven by the use of a new modulation scheme, the 8-PSK, that allows for higher data rates and also the new coding schemes, MCS1–MCS9. The coding schemes allow better link adaptation and the introduction of the concept of incremental redundancy.

There are some changes in the RLC/MAC layer functionality. The procedures for the TBF establishment are different as compared with GPRS. A new message, the EGPRS packet channel request, is introduced in this case. Other than this, the signaling for the TBF establishment is the same as in the case of GPRS. It is possible to use the CCCH and the PCCCH for the signaling. The format of the RLC data blocks has also been modified. There is a header part and a data part. There are different formats for the different coding schemes.

It is possible to send two RLC data blocks in a single radio block leading to twice the number of blocks that are transmitted. This has led to some changes in the RLC protocol. The WS for the acknowledgments has been modified. While in GPRS the size was fixed to 64, in case of EGPRS a variable WS is used. The number of timeslots that are allocated to a TBF determines the WS. The increased WS has also led to a modified acknowledgment mechanism.

Further Reading

The EGPRS air interface is the same as the GSM air interface. The concepts of the burst, frames, and the multiframes remain the same. The GSM air interface is described in the 45.abc series of GSM specifications. [3GPP TS 45.001] is an introductory specification that describes the TDMA/FDMA structure used in GSM and provides general information on frame hierarchy (including hyperframe, superframe, and multiframe), bursts, and frequency hopping.

[3GPP TS 45.002] is a more elaborate specification that describes various types of logical channels, the permitted channel combinations, and it also describes how the logical channels map to physical channels. It also specifies the digital structure of bursts.

[3GPP TS 45.003] specifies the channel-coding rules for different logical channels including rules for CRC, parity and tail bits computation, the interleaving rules, and rules of how channel-coded information is mapped to a burst.

[3GPP TS 45.004] is a very small specification that describes the 8-PSK modulation technique used in EGPRS.

[3GPP TS 45.008] is another important specification that deals with radio subsystem link control. It covers a host of important topics like power control techniques applicable for MS and BTS and measurements in idle and active mode among other topics. It covers the concepts of CV_BEP and the Mean_BEP.

[3GPP TS 44.060] describes the RLC/MAC layer and the signaling procedures for the TBF establishment and the data transfer. It provides complete information about the structure of the data blocks and describes the fields of the blocks.

[3GPP TS 43.064] is another important specification that provides an overview of the GPRS/EGPRS air interface and the RLC/MAC procedures.

References and Bibliography

[3G H. Holma] Harri Holma and Antti Toskala, "WCDMA for UMTS: Radio Access for Third Generation Mobile Communications," John Wiley & Sons.

[3G J. Korhonen] Juha Korhonen, "Introduction to 3G Mobile Communications," Artech House.

[3G S. Kasera] Sumit Kasera and Nishit Narang, "3G Networks: Architecture, Protocol and Procedures," Tata McGraw-Hill.

[3GPP TS 22.002] 3GPP TS Technical Specification 22.002, "Circuit Bearer Services (BS) supported by a Public Land Mobile Network (PLMN)."

[3GPP TS 22.003] 3GPP TS Technical Specification 22.003, "Circuit Teleservices supported by a Public Land Mobile Network (PLMN)."

[3GPP TS 22.060] 3GPP TS Technical Specification 22.060, "General Packet Radio Service (GPRS): Service description (Stage 1)."

[3GPP TS 23.002] 3GPP TS Technical Specification 23.002, "Network Architecture."

[3GPP TS 23.003] 3GPP TS Technical Specification 23.003, "Numbering, Addressing and Identification."

[3GPP TS 23.008] 3GPP TS Technical Specification 23.008, "Organisation of Subscriber Data."

[3GPP TS 23.012] 3GPP TS Technical Specification 23.012, "Location Management Procedures."

[3GPP TS 23.018] 3GPP TS Technical Specification 23.018, "Basic Call Handling; Technical Realization."

[3GPP TS 23.060] 3GPP TS Technical Specification 23.060, "General Packet Radio Service (GPRS): Service Description (Stage 2)."

[3GPP TS 24.002] 3GPP TS Technical Specification 24.002, "GSM-UMTS Public Land Mobile Network (PLMN) Access Reference Configuration."

[3GPP TS 24.007] 3GPP TS Technical Specification 24.007, "Mobile Radio Interface Signalling Layer 3; General Aspects."

[3GPP TS 24.008] 3GPP TS Technical Specification 24.008, "Mobile Radio Interface Layer 3 Specification; Core Network Protocols; Stage 3."

[3GPP TS 24.011] 3GPP TS Technical Specification 24.011, "Point-to-Point (PP) Short Message Service (SMS) Support on Mobile Radio Interface."

[3GPP TS 29.002] 3GPP TS Technical Specification 29.002, "Mobile Application Part (MAP) Specification."

[3GPP TS 29.016] 3GPP TS Technical Specification 29.016, "Serving GPRS Support Node SGSN—Visitors Location Register (VLR); Gs Interface Network Service Specification."

[3GPP TS 29.018] 3GPP TS Technical Specification 29.018, "Serving GPRS Support Node (SGSN)—Visitors Location Register (VLR); Gs Interface Layer 3 Specification."

[3GPP TS 29.060] 3GPP TS Technical Specification 29.060, "GPRS Tunnelling Protocol (GTP) Across the Gn and Gp Interface."

[3GPP TS 29.061] 3GPP TS Technical Specification 29.061, "Interworking Between the PLMN Supporting Packet Based Services and Packet Data Networks (PDN)."

[3GPP TS 42.017] 3GPP TS Technical Specification 42.017, "Subscriber Identity Module (SIM); Functional Characteristics."

[3GPP TS 43.020] 3GPP TS Technical Specification 43.020, "Security-related Network Functions."

[3GPP TS 43.022] 3GPP TS Technical Specification 43.022, "Functions Related to Mobile Station (MS) in Idle Mode and Group Receive Mode."

[3GPP TS 43.064] 3GPP TS Technical Specification 43.064, "Overall Description of the GPRS Radio Interface; Stage 2."

[3GPP TS 44.018] 3GPP TS Technical Specification 44.018, "Mobile Radio Interface Layer 3 Specification; Radio Resource Control (RRC) Protocol."

[3GPP TS 44.060] 3GPP TS Technical Specification 44.060, "MS-BSS Interface; Radio Link Control/Medium Access Control (RLC/MAC) Protocol."

[3GPP TS 44.064] 3GPP TS Technical Specification 44.064, "MS-SGSN Logical Link Control (LLC) Layer Specification."

[3GPP TS 44.065] 3GPP TS Technical Specification 44.065, "MS-SGSN; Subnetwork Dependent Convergence Protocol (SNDCP)."

[3GPP TS 45.001] 3GPP TS Technical Specification 45.001, "Physical Layer on the Radio Path; General Description."

[3GPP TS 45.002] 3GPP TS Technical Specification 45.002, "Multiplexing and Multiple Access on the Radio Path."

[3GPP TS 45.003] 3GPP TS Technical Specification 45.003, "Channel Coding."

[3GPP TS 45.004] 3GPP TS Technical Specification 45.004, "Modulation."

[3GPP TS 45.005] 3GPP TS Technical Specification 45.005, "Radio Transmission and Reception."

[3GPP TS 45.008] 3GPP TS Technical Specification 45.008, "Radio Subsystem Link Control."

[3GPP TS 45.010] 3GPP TS Technical Specification 45.010, "Radio Subsystem Synchronization."

[3GPP TS 46.001] 3GPP TS Technical Specification 46.001, "Full Rate Speech Processing Functions."

[3GPP TS 46.010] 3GPP TS Technical Specification 46.010, "Full-rate Speech Transcoding."

[3GPP TS 48.014] 3GPP TS Technical Specification 48.014, "BSS- SGSN Interface: Gb Interface Layer 1."

[3GPP TS 48.016] 3GPP TS Technical Specification 48.016, "BSS-SGSN Interface: Network Service."

[3GPP TS 48.018] 3GPP TS Technical Specification 48.018, "BSS-SGSN Interface: BSS GPRS Protocol (BSSGP)."

[3GPP TS 51.011] 3GPP TS Technical Specification 51.011, "Specification of the Subscriber Identity Module—Mobile Equipment (SIM-ME) Interface."

[GPRS C. Bettstetter] C. Bettstetter, H. J. Vogel, J. Eberspacher, "General Packet Radio Service GPRS: Architecture, Protocols, and Air Interface," IEEE Communications Survey, Third Quarter 1999, Vol. 2, No. 3.

[GPRS E Seurre] Emmanuel Seurre, Patrick Savelli, Pierre-Jean Pietri, "GPRS for Mobile Internet," Artech House.

[GSM G. Heine] Gunnar Heine, "GSM Networks: Protocols, Terminology, and Implementation," Artech House.

[GSM M. Mouly] M. Mouly and M. B. Pautet, "The GSM System for Mobile Communications," Published by the Authors.

[GSM N. Narang] N. Narang and S. Kasera, "2G Mobile Networks: GSM and HSCSD," Tata Mcgraw Hill.

[GSM T. Halonen] Timo Halone, Javier Romero, Juan Melero, "GSM, GPRS and EDGE Performance: Evolution Towards 3G/UMTS,"John Wiley, & Sons.

[ITU-T E.164] ITU-T Specification E.164, "The International Public Telecommunication Numbering Plan."

[ITU-T E.214] ITU-T Specification E.214, "Structure of the Land Mobile Global Title for the Signalling Connection Control Part (SCCP)."

[ITU-T Q.701] ITU-T Specification Q.701, "Functional Description of the Message Transfer Part (MTP) of SS7."

[ITU-T Q.702] ITU-T Specification Q.702, "SS7 MTP: Signalling Data Link."

[ITU-T Q.703] ITU-T Specification Q.703, "SS7 MTP: Signaling Link."

[ITU-T Q.704] ITU-T Specification Q.704, "SS7 MTP: Signaling Network Functions and Messages."

[ITU-T Q.711] ITU-T Specification Q.711, "Functional Description of the Signaling Connection Control Part."

[ITU-T Q.712] ITU-T Specification Q.712, "Definition and Function of Signaling Connection Control Part Messages."

[ITU-T Q.713] ITU-T Specification Q.713, "Signaling Connection Control Part Formats and Codes."

[ITU-T Q.714] ITU-T Specification Q.714, "Signaling Connection Control Part Procedures."

[ITU-T Q.715] ITU-T Specification Q.715, "Signalling Connection Control Part User Guide."

[ITU-T Q.716] ITU-T Specification Q.716, "Signalling System No. 7—Signalling Connection Control Part (SCCP) Performance."

[ITU-T Q.771] ITU-T Specification Q.771, "Functional Description of Transaction Capabilities."

[ITU-T Q.772] ITU-T Specification Q.772, "Transaction Capabilities Information Element Definitions."

[ITU-T Q.773] ITU-T Specification Q.773, "Transaction Capabilities Formats and Encoding."

[ITU-T Q.774] ITU-T Specification Q.774, "Transaction Capabilities Procedures."

[ITU-T Q.775] ITU-T Specification Q.775, "Guidelines for Using Transaction Capabilities."

Index*

* The index is made for full forms only. For indexing an abbreviation, first find the full form using the Abbreviation section and then refer the full form in this index. For example, to find GSM, find full form of GSM in Abbreviations section (which is Global System for Mobile communications), and then refer this index. For very long phrases, a part of abbreviations is used as index key word (e.g., 'GPRS Tunneling Protocol' is used instead of 'General Packet Radio Services Tunneling Protocol').