Future Mobile Satellite Systems: Issues in Timing Synchronisation and Resource Allocation Techniques

A thesis submitted for the degree of
Doctor of Philosophy
in Electronic and Electrical Engineering

2002

Aikaterini Koutsaftiki

Department of Electrical and Electronic Engineering
University College London

ProQuest Number: 10014989

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

Acknowledgements

Στην γιαγιά μου,

Άννα

Abstract

In an era where the personal communication character of the user is the main technology focus, the convergence of mobile satellite and mobile terrestrial systems seems not too distant. The extend of the integration of the satellite-based and the terrestrial-based systems introduces major architectural challenges.

The first chapter of this thesis presents a general overview of the mobile satellite systems, presents the organisation of the thesis and details the main contributions.

In the second chapter some background information relevant to the research is presented. Specifically, aspects of GSM and GPRS are presented. Additionally, after a brief description of S-PCSs, the subject of satellite and GSM integration is addressed.

In the third chapter issues regarding timing synchronisation, when GSM is integrated with a satellite system, are examined, and solutions to mitigate parts of the problems identified are proposed.

In the fourth chapter the GPRS system is examined in the satellite environment. The research is focused in the GPRS's RLC/MAC protocol, which is mainly responsible for the allocation of resources amongst the users.

Dynamic Channel Allocation (DCA) techniques, are expected to improve dramatically the efficiency of the resources allocation issue, by allowing a more intelligent use of the limited spectrum available for mobile communications. In the fifth chapter, novel DCA algorithms are introduced and are examined for purposes of resources allocation for mobile satellite systems.

In the sixth chapter heuristic algorithms are used in combination with the proposed DCA schemes in order to offer a further optimisation in the systems examined. Simulated annealing, is incorporated to the real-time system scenarios, in contrast to its conventional use for non-dynamic systems and major benefits to the systems are introduced.

The final chapter brings the thesis to a conclusion by summarising the main results and proposing areas where further work may be fruitful.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

Until recently satellite communications were mainly considered in the context of providing point to point trunk telephony or one way broadcasting. However, in the beginning of the nineties satellite technology evolution allowed the introduction of more 'user-oriented' satellite services to be proposed. Services similar to GSM were included in this portfolio supported by these systems. These Satellite Personal Communication Services systems (S-PCSs) were first proposed in 1990 by the US-based Iridium consortium led by Motorola. In an era where the personal communication character of the user is the main technology focus, the convergence of mobile satellite and mobile terrestrial systems seems not too distant. Specifically, satellites are expected to play a very important role in the environment where the user is expecting to receive seamless services, irrespective of location or environment. The extension of the integration of the satellite-based and the terrestrial-based systems

14

is one of the most important areas of research at the moment, which is generating interest also because of the architectural challenges that it raises. However, such a convergence can be shown to be extremely beneficial due to the complementary characteristics of a terrestrial and a satellite system. Satellite systems can offer the very important characteristic of the enlargement, the service extension and the completion on a global basis, while the terrestrial part can cater for the service of the more concentrated traffic volumes. In this broad context, this research addresses issues of such integration, specifically in the timing synchronisation and resource allocation areas, and additionally focuses on enhancements of such algorithms for satellite mobile systems of the future.

1.2 Organisation of the Thesis

In chapter 2 a review of the technologies relevant to this thesis is given. Specifically, the aspects of GSM and GPRS that are relevant to this research are presented and additional information is provided about S-PCS systems.

In chapter 3 timing synchronisation problems in GSM and satellite integration are addressed, and solutions to this issue are introduced.

In chapter 4 the behaviour of the RLC/MAC layer of the GPRS protocol operating over a MEO satellite system is studied. Several models are tested for the uplink domain of GPRS and innovative schemes to improve system performance are introduced.

In chapter 5 an overview of the heuristic techniques that are used for optimisation

purposes in the area of satellite resource allocation is given. The application of different techniques is examined for a system that is time invariant.

In chapter 6, algorithms that were proved to be efficient in chapter 5 are adapted and incorporated into a time variant system to examine if they offer a performance improvement in a dynamic environment as well.

Finally, in chapter 7 the thesis is concluded and suggestions for further work are given.

1.3   Summary of main contributions

This work logically consists of two parts. In the first part, we address and propose solutions for the timing and synchronisation issues around GSM, which would also apply to GPRS. We then study the performance of the RLC/MAC layer in allocating resources in the radio access system of such networks. The second part of this work focuses on a novel approach to satellite mobile networks, and resource allocation optimisation techniques for a fully dynamic system are studied.

The main contributions offered by this dissertation may be summarised as follows:

- Introduction of solutions in order to mitigate synchronisation problems in a GSM and S-PCS integrated system. The algorithms were tested for their efficiency in both a 'fixed-satellite cell' and a 'fixed-earth cell' satellite system environment.

- Detailed development of a simulation model/platform where the examination

of the GPRS protocol behaviour in the satellite environment was effected.

- Identification of areas that further research could be focused on to enhance

  such a behaviour was additionally included. Further to that, specific schemes

  focusing on these areas were introduced to offer system improvement.

- Development of novel heuristic algorithms which lead to the design of satellite

  networks with improved utilisation of resources.

- Application of novel simulated annealing algorithms, traditionally used only

  for static systems, to time variant systems for resource management purposes.

  An adjustment of such an application so that system performance is improved

  while reducing any further drawbacks, such as signalling overload or processing

  complexity, is addressed.

As a result of the work done the following papers were published:

- A. Koutsaftiki, C. Matrakidis, P. Lane, M. Kokkos and I. Macnamara, "Proposal for spotbeam pattern for GSM-based satellite personal communications", *Electronics Letters*, 18th February 1999, Vol.35 No.4, pp 279-280.

- A. Koutsaftiki, P. Lane and J. J. O' Reilly, "Aspects of the GPRS RLC/MAC protocol over a MEO satellite system", *IEEE Radio and Wireless Conference (RAWCON'99)*, Denver, 1-4 August.

- A. Koutsaftiki, C. Matrakidis, P. Lane, J. J. O'Reilly, "Improved dynamic channel allocation for satellite systems", *IEEE Vehicular Technology Conference 2000 (VTC'00) - Spring*, Tokyo, 15-18 May.

- A. Koutsaftiki, P. Lane and J. J. O' Reilly, "Proposals on UMTS System Configuration", *IEEE International Conference on Universal Personal Communications, (ICUPC' 97)*, San Diego, 12-16 October.

Following this introduction chapter, the next chapter presents an overview of GSM and GPRS.

Chapter 2

General Background

2.1   Introduction

In this chapter some background information relevant to the research is presented. Specifically, GSM and GPRS are presented, with more detailed information on the radio access part. Additionally, after a brief introduction of satellite networks and two important S-PCS systems, we address the subject of satellite and GSM integration. The chapter is concluded with the presentation of 'fixed-earth cell' satellite systems, as this is a topic of importance to the subsequent research.

2.2   GSM

In 1982, the Group Special Mobile (GSM) was formed by CEPT, in order to prevent the technological fragmentation of the first generation analogue cellular systems. This program was aimed at the implementation of a second generation

system, and its major participants were Germany and France. The main target of this group was to conduct research, in order to create a system that was expected to:

- achieve good spectral efficiency, and a larger subscriber capacity

- support international roaming, and offer widespread availability

- offer good speech quality

- have low service and terminal costs

- support new services and achieve compatibility with ISDN

- be adaptable to traffic density

- provide services to both vehicle and portable stations

In 1986 the research group decided that the system should use TDMA technology. In 1987 the European Council recommended that a Pan-European system should be introduced in the EC market.

The final outcome of the whole process was GSM, which was introduced in 1991. This was a digital system that was designed to occupy the 890-915 MHz band for uplink (MS to BS) transmission and the 935-960 MHz band for downlink (BS to MS) use, in each European country. GSM was a huge success, and in 1993, it was already in service in 22 countries. However, GSM is no longer solely a European standard, since as of July 1997 more than 120 countries located in South and North America, Asia and Australia had adopted it or one of its derivatives. Its derivatives

are GSM1800 and PCS1900 (in North America). This amounted to a total 44 million

customers worldwide.

2.2.1 GSM services

Consistent with ITU definitions, GSM can support teleservices, bearer services

and supplementary services (figure 2.1, [1]). These services are classified like ISDN

services as they were patterned after the services offered by ISDN for fixed terminals

tied to telephone lines.
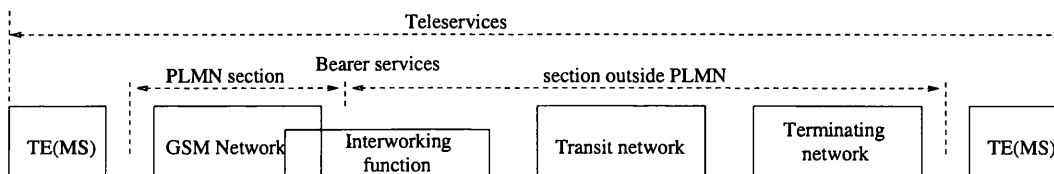
Figure 2.1: GSM Bearer services and teleservices

- **Bearer Services**: GSM can support data throughput ranging from 300 bits/sec

  upto 9600 bits/sec to the Integrated Services Digital Network (ISDN), to the

  Circuit Switched Public Data Network (CSPDN) or to the Packet Switched

  Public Data Network (PSPDN). Data is supported by using the X.25 or the

  X.23 protocols and the three lower OSI layers. It can be supported in a trans-

  parent and non-transparent circuit-switched modes. The transparent mode

  employs a Forward Error Correction (FEC) for good reception quality, while

  the non-transparent mode activates the Radio Link Protocol (RLP) which uses

  ARQ to achieve a satisfactory reception quality.

- **Teleservices**: GSM can support digitised voice to POTS, PSTN and ISDN on an international basis. An emergency call facility is also available on a national basis. There are also facsimile, videotext, teletext services and a Short Message Service (SMS). The latter is something like an e-mail service and can support messages of upto 160 bytes. SMS messages are transmitted in a store-and-forward fashion and can be stored in the SIM card. SMS can also be used in a broadcast mode and can employ an ARQ scheme to ensure reliability.

- **Supplementary Services**: Some examples of these additional services are call waiting, call forwarding, barring of incoming or outgoing calls in either the home environment or while roaming, caller identification, charging services such as freephone and reverse charging, and multi-party conversation.

2.2.2   The GSM Network [3, 4, 5]

The general structure of the Public Land Mobile Network (PLMN) section for a GSM network is presented in fig 2.2. As we can see, it is divided into a number of subsystems.

- **The MS Subsystem**: The Mobile Station Subsystem (MS) consists of the Mobile Terminal (MT) and the Subscriber Identity Module (SIM) card. The SIM card is a smart card which can be inserted into any GSM MT and pro-vides personal user mobility as opposed to simple equipment mobility. The MS and its user are characterised by three main personal identifiers: the
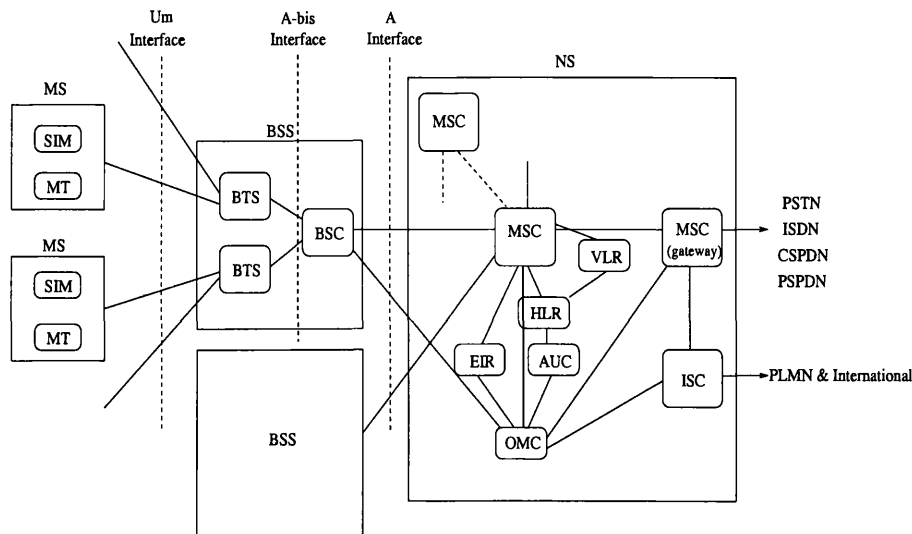
Figure 2.2: GSM Network Architecture

International Mobile Station Equipment Identity (IMEI), the International Mobile Subscriber Identity (IMSI) and the Mobile Subscriber ISDN Number (MSISDN). The IMEI is a unique international mobile equipment serial number. The IMSI is stored in the SIM and is a unique user identifier. The MSISDN is the actual user telephone number and is different from the IMSI (GSM was the first mobile system to make the distinction) for confidentiality reasons. The MS uses the Um radio interface to connect to the BSS.

- **The BS Subsystem**: The Base Station Subsystem (BSS) consists of the Base Transceiver Stations (BTSs) and the Base Station Controllers (BSCs). The interface between the BTS and the BSC is the A-bis interface, which can be different for different manufacturers. Thus, in a GSM network, the BSCs and the BTSs are usually obtained from the same manufacturer. The BTS handles the radio link protocols within a given cell. The BSC controls several BTSs and is responsible for the channel set-up, hand-overs etc. There are also

the Speech Transcoders (STs), which are responsible for adapting the low rate

radio channel, to a 64 kbps channel, which is used in the fixed network. The

interface between the BSCs and the MSCs (in the Network Subsystem) is the

A interface which uses the Signalling Connection Control Part (SCCP), part

of the SS7 protocol stack, for signalling.

- **The Network Subsystem**: Finally, the Network Subsystem (NS) consists

of the several Mobile Switching Centres (MSCs) and Visitor Location Regis-

ters (VLRs), the Home Location Register (HLR), the Operations and Main-

tenance Controller (OMC), the AUthentication Centre (AUC) and the Equip-

ment Identity Register (EIR). The signalling between all these NS units uses

SS7 signalling.

The MSCs perform the basic switching function within the NSS and they co-

ordinate the setting up of calls to and from GSM users. They are the elements

that provide the subscribers with teleservices, bearer services, etc. They con-

trol functionalities such as registration, authentication, location updating and

hand-over. Additionally, some of the MSCs in the network act as gateways to

the PSTN, ISDN, CSPDN and PSPDN. The main difference between an MSC

and an ISDN switch is that the MSC is additionally responsible for resource

administration and user mobility functions. The VLRs are units which, nowa-

days, are implemented along with MSCs. The VLR temporarily stores the

roaming data along with the IMSI of each user controlled by the MSC con-

nected to the specific VLR. The VLR stores more precise location information

for the user than the HLR, as well as a copy of the remaining subscriber related information. The HLR contains the registration subscriber data along with information about the services that can be provided to him/her. It also contains location information as to which MSC is currently serving the user. HLRs are usually duplicated for reliability purposes. The OMC is responsible for the management and integration functions, such as security functions, network configuration and operations and administration functions. In order to provide higher security to the user, the EIR and AUC units are included in a GSM network. The EIR contains all the IMEIs of the legally registered MTs. The AUC stores data and confidential keys.

2.2.3 GSM Radio Interface [2, 6]

As has already been mentioned, the GSM system uses 2x25 MHz bands for its radio interface. Since a 100 kHz frequency band is excluded at each edge of the spectrum (i.e. 890-890.1 MHz and 959.9-960 MHz), 124 carriers are used with a 200 kHz spacing (RF carrier bandwidth). The multiple access technique used is a combination of FDMA and TDMA. That means that every one of the 124 RF channels, is divided into 8 physical channels (timeslots), each one occupied by a single user or a signalling channel. The RF channel (TDMA frame) data rate is 270.833 kbps, which means that the actual effective user bit rate is 270.833/8=33.854 kbps. The TDMA frame is transmitted in 4.615 ms. This means that each user is transmitting in bursts (timeslots) each lasting 0.577 ms. In this 0.577 ms, 156.25 bits are transmitted. In GSM there are four types of bursts:

- the normal burst

- the F burst used in the Frequency Correction Control Channel. This is a simpler burst because all the 148 bits are zero and, hence, a pure sine wave is produced.

- the S burst used in the Synchronisation Control Channel. Both the S and F bursts are used for synchronisation.

- the access burst used in Random Access Channel (RACH).

The structure of the bursts, the TDMA frame and the multiframe are shown in fig 2.3.
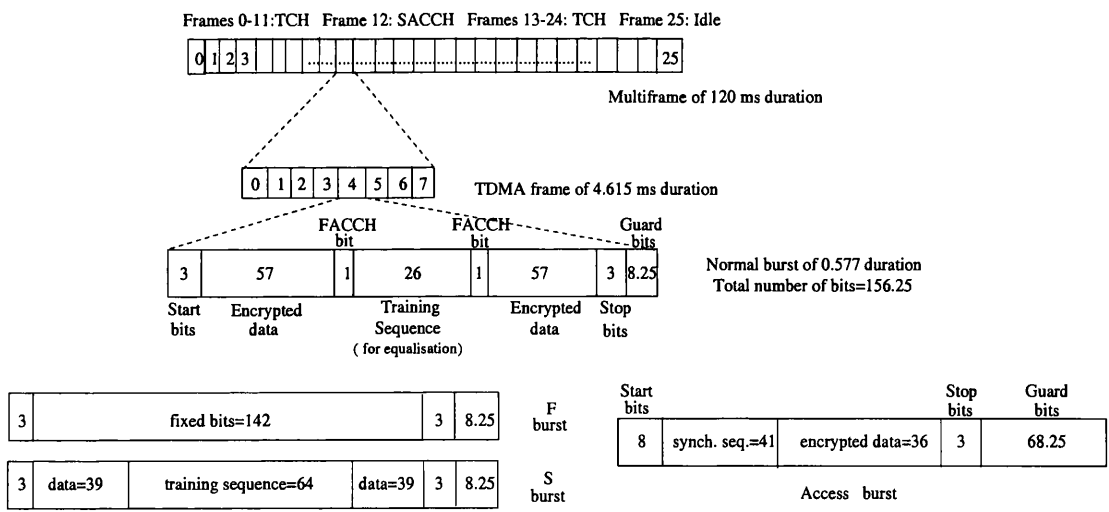
Figure 2.3: Multiframes, TDMA frames and bursts in GSM

There are two types of channels used in the radio interface:

1. *Traffic channels* (TCHs) which can be either full-rate traffic channels (TCH/F) with 22.8 kbps net bit rate or half-rate traffic channels (by using half-rate

coders) (TCH/H) with 11.4 kbps net bit rate. The traffic channels are divided into:

- Speech channels that can be either full-rate or half-rate traffic channels.

- Data channels which can support a variety of data rates (300 bits/sec to 9.6 kbps) in both half and full rate traffic modes.

2. *Control channels* (CCHs) which are responsible for signalling and synchronisation procedures between the MS and the BSC. These will be presented separately in the next section.

Control Channels

There are three types of control channels:

1. *Broadcasting Control Channels* (BCHs), which are only downlink channels used in the idle mode. These are:

   - Broadcast Control Channel (BCCH), which is used for broadcasting information.

   - Frequency Correction Channel (FCCH), using the TDMA frame slot 0.

   - Synchronisation Channel (SCH). There is only one FCCH and one SCH channel in a GSM network, using the TDMA frame slot 0.

2. *Common Control Channels* (CCH), which are used during call set-up. These are:

- Access Grant Control Channel (AGCH) (for downlink use), which allo-
  cates an SDCCH to the mobile, after receiving a RACH.

- Paging Channel (PCH) (for downlink use), which alerts the MS of an
  incoming call.

- Random Access Channel (RACH) (for uplink use), which is used by the
  mobile to request access to the GSM network by using a slotted ALOHA
  scheme.

3. *Dedicated Control Channels* (DCH), which are for both downlink and uplink
   use. These are:

- Stand-alone Dedicated Control Channel (SDCCH) (782 bits/sec), which
  is responsible for main signalling processes such as location updating,
  registration etc. It is mainly used in the idle mode.

- Slow Associated Control Channel (SACCH) (391 bits/sec), which is re-
  sponsible for sending measurements of the power received from other
  BTSs for handover purposes. It is transmitted every 120 ms and is mainly
  used during the call time.

- Fast Associated Control Channel (FACCH) (9.2 kbps), which is used for
  urgent procedures and very fast signalling in call establishment or during
  a hand-over.

Encoding and modulation

The GSM system transmits digitised speech and data. The whole speech encoding and modulation system is shown in figure 2.4. The speech encoder is a Regular Pulse Excited Linear Predictive Coder (RPE-LPC). Encoded speech comes out of the full-rate speech encoder with a 13 kbps rate. The data channel and control channel rate is 9.6 kbps and 9.2 kbps respectively, after encoding. After the channel FEC encoding, all channel rates are 22.8 kbps. Then the data are interleaved. Interleaving is used to tackle a burst of errors by scrambling the data. During ciphering the data encryption is taking place.
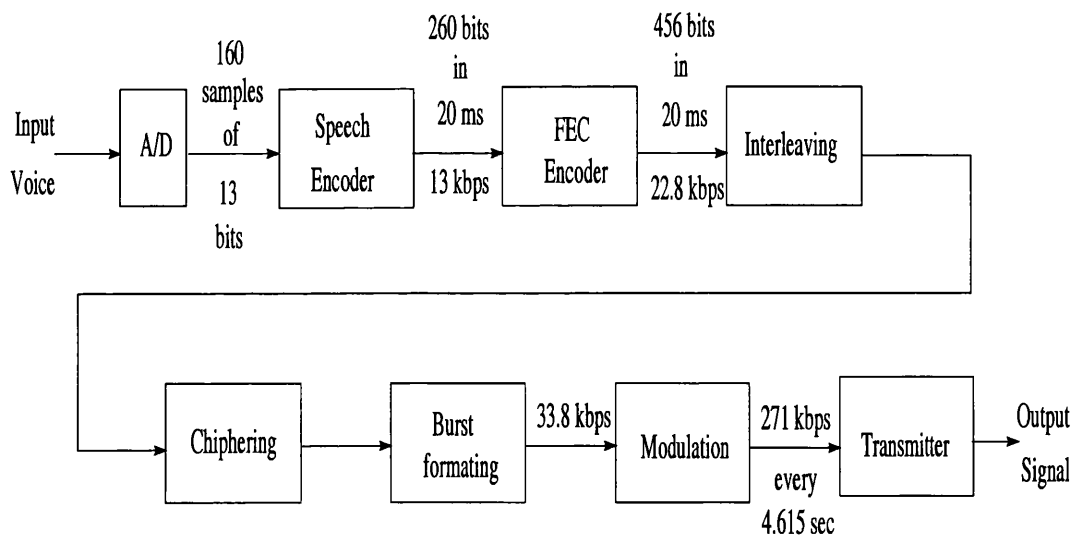
Figure 2.4: Flow diagram of encoding and modulating a full-rate TCH

The modulation scheme that is used in GSM is the binary BT=0.3 Gaussian Minimum Shift Keying (GMSK). This particular technique helps minimising the bandwidth occupied by the modulation spectrum and hence, gives good spectral efficiency.

2.2.4   GSM protocol structure

GSM is OSI compatible and its specifications define the three lower layers of the OSI model. Figure 2.5 shows an overview of the signalling architecture.

The first layer, the physical layer, is responsible for modulation, channel coding, the structure of the frames, and error correction techniques.

The second layer, the data link layer, specifies the Link Access protocol on Dm channels (LAPDm), which is a derivative of LAPD adapted to the air interface. LAPDm can work in both unacknowledged and acknowledged modes of operation. LAPDm is responsible for the protected transfer of signalling messages over the air interface. Segmentation and reassembly of layer 3 messages is also provided. It is also in control of the distribution and random access procedures.

The third layer is responsible for Radio Resource (RR) Management, Mobility Management (MM) and Connection Management (CM) procedures. RR is defined as the set of functions that relate to the assignment, allocation and distribution of the resources. MM is defined as the set of functions that relate to location updating, registration as well as security procedures. CM is defined as the set of functions that are related to call management, such as establishment, control and termination functions and provides the network with both supplementary service and short message service management.

The Message Transfer Part (MTP) provides message transfer in the SS7 part of the network. A derivative of MTP is MTP' which contains the three lower OSI layers and provides protected transport of the signalling messages across the A interface.

For signalling between the BSC and the MSC, the Base Station System Application Part (BSSAP) and the Signalling Connection Control Part (SCCP) layers are used.

2.3 GPRS

Data transmission with a satisfactory quality of service is the next immediate target on which the mobile industry is focused. In this context, the General Packet Radio Service (GPRS) system is a new GSM extension, which is expected to offer packet switching services with rates upto 171.20 kbps.

Until recently, telephony was the main mobile application and applications such as facsimile and SMS were available for data transmission. The demand for access to a wider range of data applications while on the move is dramatically increasing. Examples of such data applications include access to the world wide web, electronic mail, file transfer protocol, telematics, fleet management, etc. These applications are of an extremely bursty nature, and provision over traditional circuit-switched systems would lead to an undesirable waste of resources.

GPRS is a new GSM-based packet switching service which will be used for the transfer of bursty type of data. The main applications of GPRS will be internet (email, web, telnet) and intranet types of applications. GPRS will optimise the use of GSM resources by allocating them only when needed and sharing them efficiently amongst different users.

In the next sections a general overview of the important aspects of GPRS will be provided, with an emphasis on the user access part.
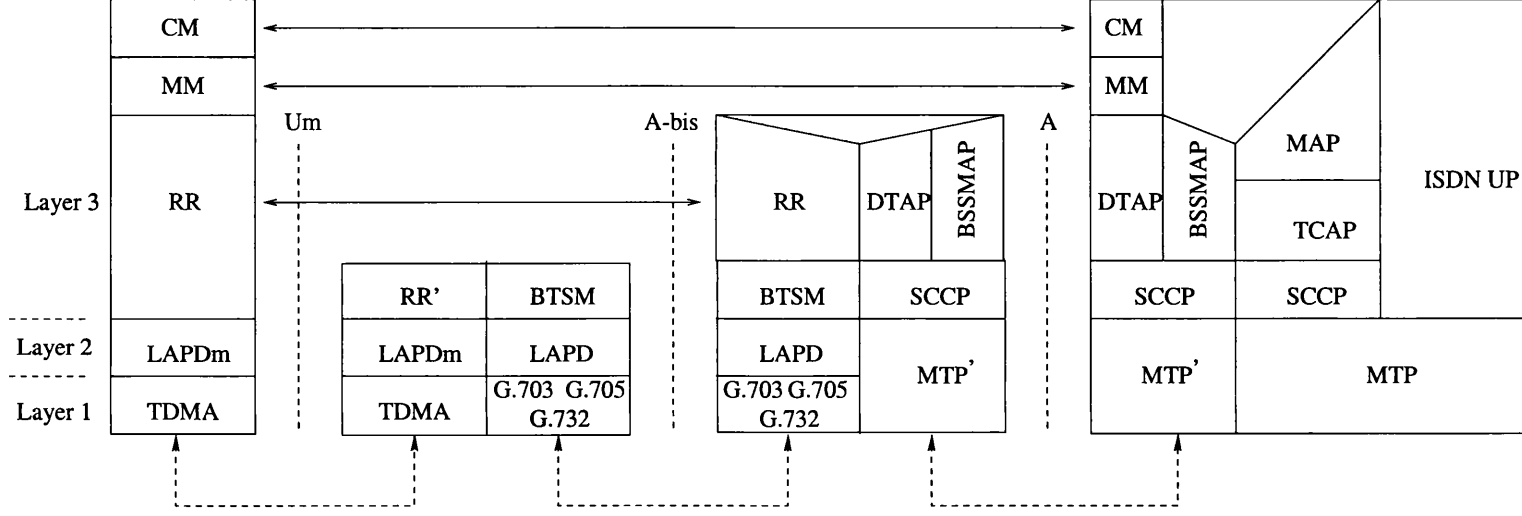
Figure 2.5: GSM signalling Protocol Architecture

2.3.1   GPRS network

The fact that GPRS is a packet switching technique introduces major changes to the GSM network (figure 2.6) and additions to the GSM's air interface (figure 2.2). There are two GPRS Support Nodes (GSNs), specifically the Gateway GSN (GGSN) and the Serving GSN (SGSN), which are the two main new entities in the GSM/GPRS network ([7]). These two nodes are responsible for packet transfer in the PLMN. The SGSN is in charge of distributing and sending the packets to the routing area that it covers. The GGSN acts as a gateway between the GPRS network and external PSPDN networks. It also coordinates the routing of the protocol data units (PDUs) to the SGSN serving the MS area. Specifically, the routing of PDUs between different GSNs is performed using a tunnelling process. In this process, the PDUs are IP encapsulated in the initial GSN, routed through the GPRS backbone via the G_n and G_p interface, to the terminating GSN where the encapsulation is removed. The Gp interface connects GGSNs and SGSNs that are located in different PLMNs. The Gn interface connects all the GSNs that are located in the same PLMN. Finally, the Gi interface connects the gateway GGSN to the external data networks and can use the IP or X.25 protocol.

An additional major modification in the traditional GSM network is that the HLR now also includes GPRS subscriber information which is kept in a new database called the GPRS Register (GR). In the GR the IMSI is mapped to different Packet Data Protocol (PDP) network addresses.
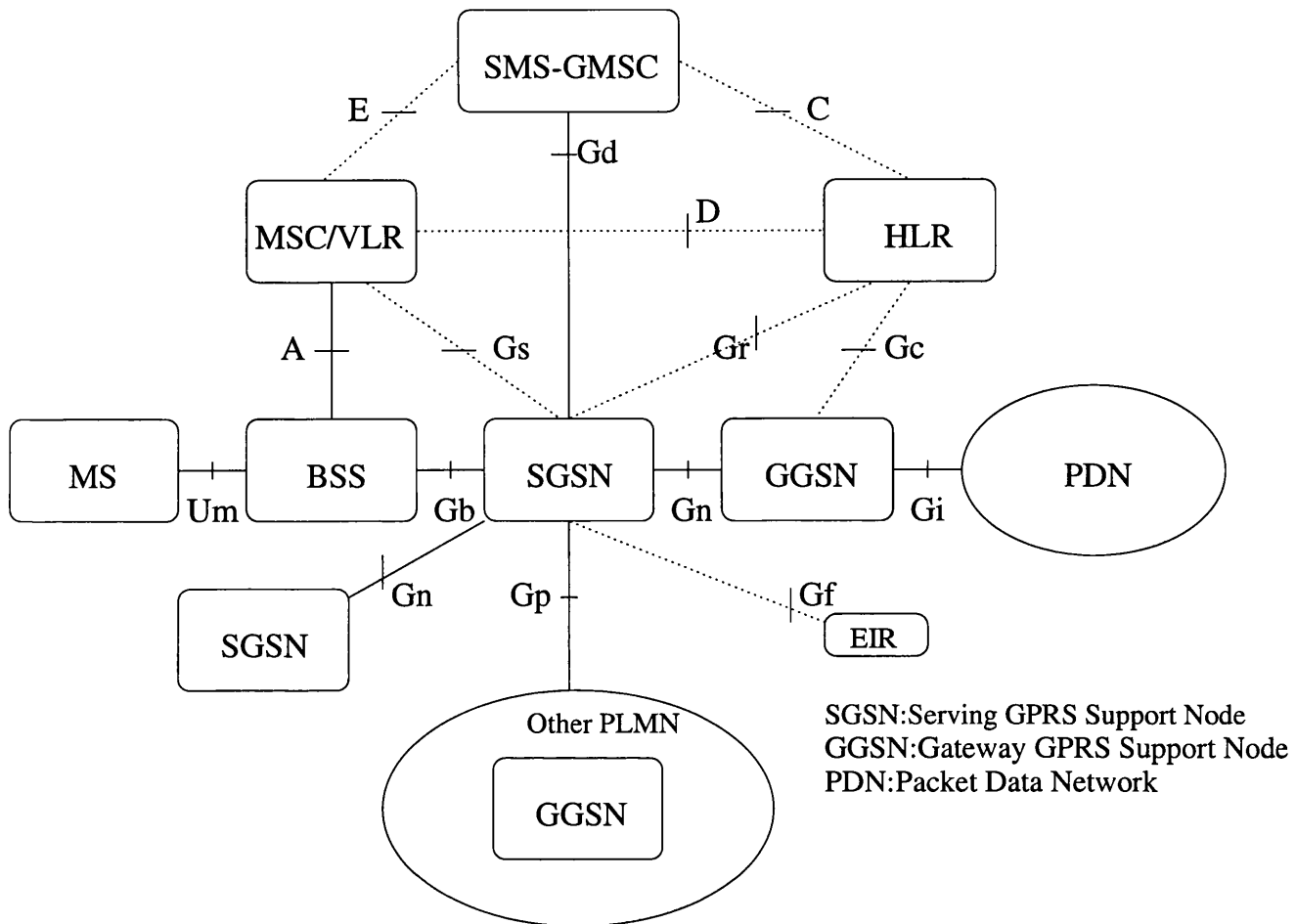
Figure 2.6: The GPRS network

2.3.2 GPRS protocol stack

The GPRS transmission plane is presented in figure 2.7. The Data Link Layer (DLL) is divided in the Radio Link Control/Medium Access Control (RLC/MAC) and the Logical Link Control (LLC) layers. The RLC/MAC layer ([9], [40]) is the layer on which this research is focused and will be explained in detail in the following section. The LLC layer acts as a logical link between the MS and the SGSN, and it is based on GSM's LAPDm. The Subnetwork Dependent Convergence Protocol (SNDCP) lies above the LLC and is responsible for multiplexing as well as encryption, segmentation and compression of network layer messages. The BSS

GPRS Protocol (BSSGP) is derived from GSM's BSSMAP, and conveys routing and QoS information between the SGSN and the BSS. The tunnelling process, referred to earlier, is performed by the GPRS Tunnelling Protocol (GTP) layer running over TCP/IP. Below TCP/IP, protocols such as ATM or Ethernet can be used.

RLC/MAC layer

The Radio Link Control (RLC) layer's main functions are the segmentation and reassembly of LLC frames into RLC data blocks, as well as the Backward Error Correction (BEC) procedure, which enables the selective retransmission of uncorrectable codewords. The Medium Access Control (MAC) layer is responsible for allocating the available resources and also multiplexing the data and the control signalling on both uplink and downlink directions. Mobile-originated channel access uses the slotted-aloha scheme exactly as in GSM. The MAC layer is also responsible for the contention resolution of these access attempts. For the mobile-terminated access attempts, the MAC layer is responsible for their scheduling, as well as for the queueing of the packets.

The two most important features that the RLC/MAC layer introduces to GPRS are the use of Uplink State Flags (USFs) and the selective ARQ mechanism, used in GPRS acknowledged mode. The USFs are 3-bit words that are included in the beginning of each RLC block and allow the multiplexing of data from upto 7 mobiles (the 8th USF value is reserved for signalling purposes) in the same timeslot. Each USF value reserves one timeslot for a particular mobile. The USF always points to the next uplink radio block. There are 7 USF values used for each timeslot. In figure

IP/X.25
SNDCP
LLC
RLC
MAC
Physical Link
Physical RF

MS

LLC	relay
RLC	BSSGP
MAC	Frame Relay
Physical Link	Physical
Physical RF	Layer

BSS

SNDCP	GTP
LLC	TCP/UDP
BSSGP	IP
Frame Relay	L2
Physical	Physical
Layer	Layer

SGSN

IP/X.25
GTP
TCP/UDP
IP
L2
Physical
Layer

GGSN

Figure 2.7: The GPRS transmission plane
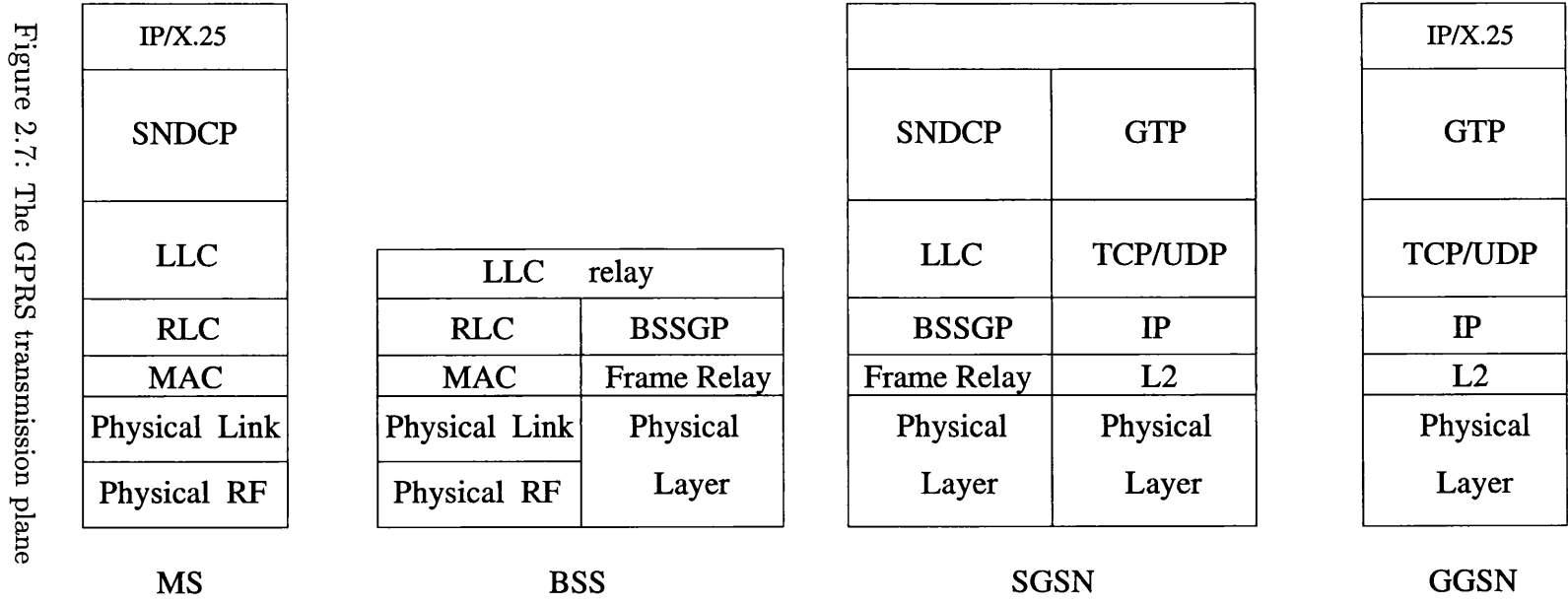
2.8 the A and B mobiles are sharing timeslots one and two by using the R2 and R1 USF values respectively. A mobile can be assigned with more than one timeslot and USF combinations, and therefore can use more than one data channel for a packet transfer. Additionally, it can receive data which is totally independent from that which it is transmitting. Finally, in GPRS the downlink and uplink data transfer can be assymetrical.

An RLC block is created according to the process presented in figure 2.9. Specifically, a network PDU (N-PDU) is segmented into LLC frames, each with a maximum size of 1600 octets. An LLC frame is then segmented into RLC blocks after convolutional encoding has been performed. The size of each RLC block is 456 bits and is transmitted over the air interface through four normal GSM bursts (presented in section 2.2.3).

When an LLC frame is segmented into RLC blocks, these blocks carry a Temporary Frame Identity (TFI) which is unique amongst concurrent frame transfer sequences to a cell. The TFI provides frame numbering so that retransmitting of the erroneous or lost blocks is performed by the ARQ scheme. The ARQ mechanism uses a 64 block transmit window size and the ACK/NACKs are sent periodically from the receiving side. Since this scheme is selective ARQ, the sending side then retransmits only the erroneous RLC data blocks. The reception of error-free blocks results in the further sliding of the window. When all the blocks belonging to the same LLC frame are received correctly, they are then forwarded to the LLC layer. In GPRS unacknowledged mode the transfer of RLC data blocks does not include any retransmissions. The RLC blocks are numbered in the same way to enable the

TDMA fr TS No	1-3	4-7	8-11	12-15	16-19	20-23	24-27	28-31	32-35	36-39	40-43	44-47	48-51
Downl. 0	F A	F	chres AB F	F	F	F	F	F	F	F	F	F	F B
Upl.													
Downl. TS1	R1	R1	R1	R2	R2 A1	R1 A3	R1 B2	R1	ACK-B R1	R2	R1 A3	R1	ACK-A R1
Upl.													
Downl. TS2	R1	R1	R1	R2	R1 A2	R2 B1	R1 A4	R1	NACK-A R1	R1	R1	R1	R1
Upl.													

Figure 2.8: Example of USF sharing between two mobiles
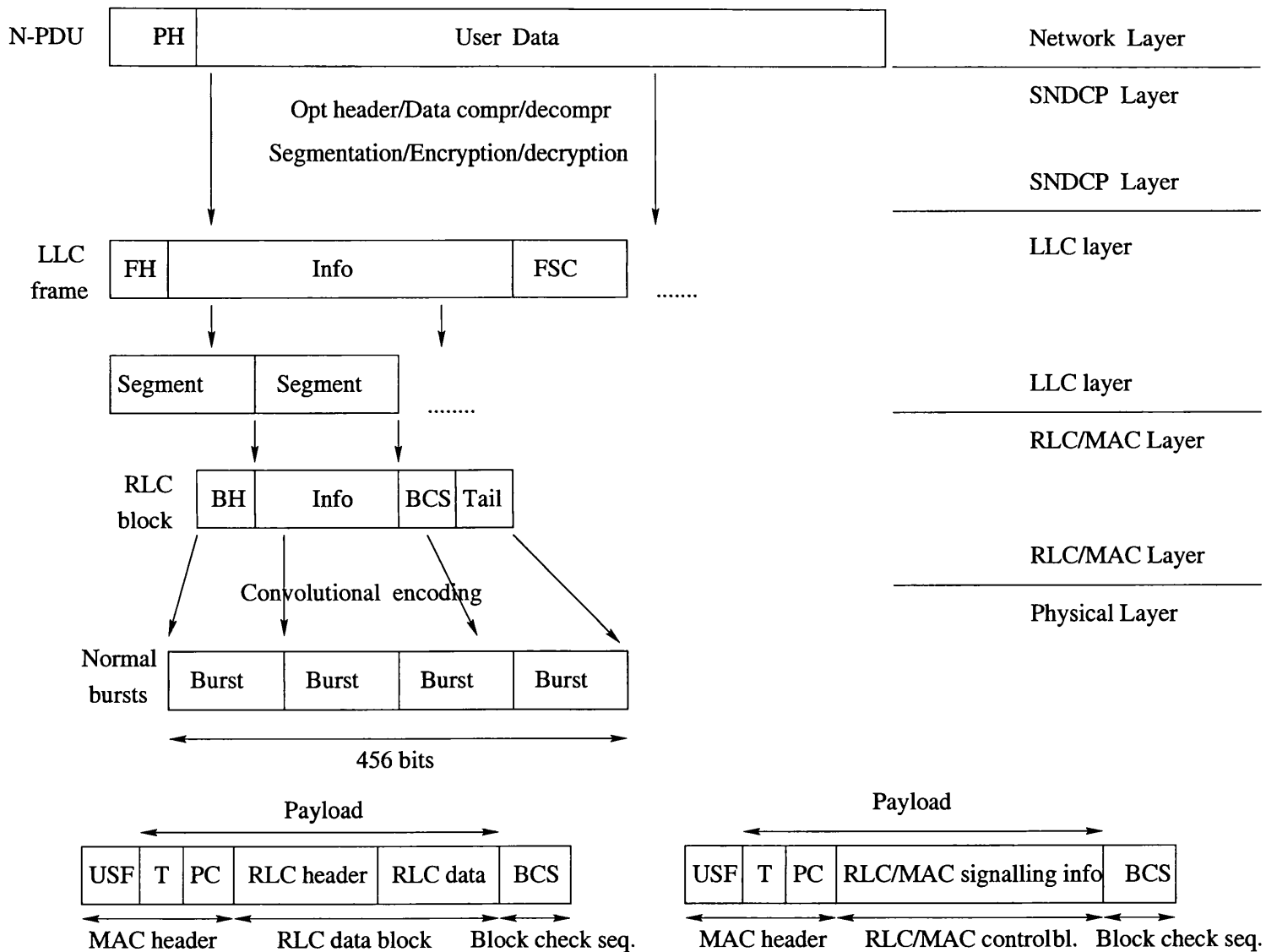
Figure 2.9: Packet transformation data flow

LLC reassembly process. The receiving side sends ACKs/NACKs only for control

signalling purposes, but no list of correct and erroneous blocks is included in these

messages.

2.3.3 The GPRS air interface

The physical channels which are dedicated to GPRS operation are called Packet

Data Channels (PDCHs). At least one PDCH acts as a master and accommodates

all the control signalling channels which are called Packet Common Control Chan-

nels (PCCCHs). Slave channels usually carry Packet Traffic Channels (PTCHs)

which consist of user data and dedicated signalling. The GPRS logical channels are

presented in figure 2.10 and have the same functionalities as the corresponding GSM

channels. The Packet Notification Channel (PNCH) is a multicast notification to

a group of mobiles and is a form of resource assignment in order to notify the user

that the PRACH burst has been received correctly. The PACCH channel carries

acknowledgements, power control information as well as resource assignment and

reassignment messages.

Two examples of the use of PDCHs, an MS initiated packet transfer and a

network initiated packet transfer, are illustrated in figure 2.11. A user can choose

between a one-phase or two-phase access to the network. In the two-phase access

approach, which is an optional one, the user provides the network with a more

complete description of the requested resources.

In GPRS, four different coding schemes, CS1-CS4 are used [9]. CS1-CS3 use a

1/2, 2/3 and 3/4 rate convolutional coding scheme respectively, while CS4 is un-
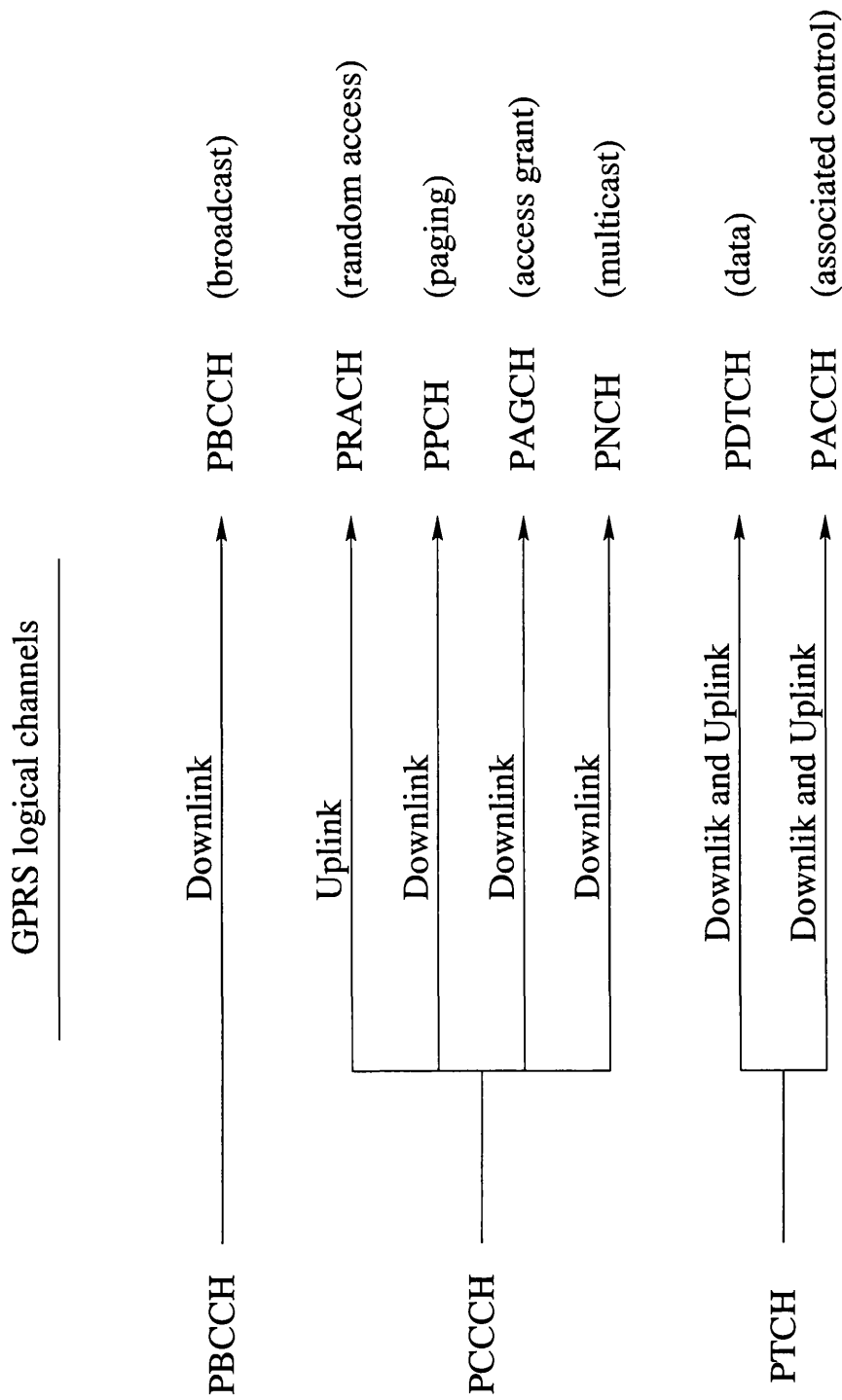
Figure 2.10: The GPRS air interface

coded. Some important details about each coding scheme are presented in table 2.1.

From table 2.1 we observe that GPRS can offer significantly higher data rates

Figure 2.11: Two cases of packet transfer: MS initiated and network initiated

Left diagram:

MS — Network

Message	Channel
Channel Request →	PRACH/RACH
← Immediate assignment	PAGCH/AGCH
Resource Request →	PACCH (opt)
← Resource Assignment	PACCH (opt)
Data →	PDTCH
Data →	PDTCH
Data →	PDTCH
Data →	PDTCH
← Temp ACK	PACCH
Data →	PDTCH
Data →	PDTCH
← Final ACK	PACCH

Right diagram:

MS — Network

Message	Channel
← Paging Request	PPCH/PCH
Channel Request →	PRACH/RACH
← Immediate assignment	PRACH/RACH
Paging Response →	PAGCH/AGCH
← Resource Assignment	PACCH/PAGCH/AGCH
← Data	PDTCH
← Data	PDTCH
← Data	PDTCH
Temp ACK →	PACCH
← Data	PDTCH
← Data	PDTCH
Final ACK →	PACCH

Coding scheme	Convolutional code rate	USF Payload	block (bits)	Data bit rate (kbps)
CS1	1/2	3	181	9.05
CS2	2/3	6	268	13.4
CS3	3/4	6	314	15.6
CS4	1	12	428	21.4

Table 2.1: GPRS coding schemes

per timeslot than GSM, which attains a maximum data rate of 9.6 kbps as discussed earlier. Considering also the fact that several timeslots can be used by the same user for data transmission, one can conclude that very high data rates can be achieved with GPRS. There is of course a trade off between the high coding rates and the total delay caused by the retransmissions of erroneous blocks, and the scheme that provides the most balanced solution must be adopted for each different radio environment. USFs provide the user with information that has to be decoded a lot faster than the rest of the block in order to enable him/her to prepare the data transmission in the next block time. Therefore, in schemes CS2-CS4 a 12-bit USF codeword is created. In schemes CS2 and CS3 a 6-bit USF block is created, and then convolutional coding is applied to the whole block without puncturing the first 12 bits.

2.4 S-PCSs

Personal Communication Networks (PCNs) are expected to offer personal communication services through a single number, anywhere, anytime, to a hand-held terminal. In order for the 'anywhere' prerequisite to be achieved, satellites should be integrated in such a system in order to provide service extension in areas or re-

gions where the density of users is too low to justify the financial investment needed to commission a fixed infrastructure. Hence, satellite Personal Communication Systems (S-PCSs) are becoming an indispensable part of PCNs.

In the following sections we will first generally present the basic components and characteristics of a satellite system. Then, we will focus on a description of S-PCN systems and will focus on two major representatives, Iridium and ICO[1]. Since a large percentage of this research is focused on issues of GSM/GPRS when applied over a satellite system, a small analysis on the GSM and satellite integration will be included as well. Finally, a more detailed presentation of the differences between 'earth-fixed cell' and 'satellite-fixed cell' systems will complete the background on S-PCSs.

2.4.1 General satellite system description

A satellite system can be modular and can consist of several independent components. Generally, it is considered to be made up of the satellite and the terrestrial components (figure 2.12). The terrestrial component consists of core networks such as the PSTN and ISDN, which provide services to the end users. The satellite component can be divided into the ground and space sections. In the ground segment, two types of terminals can exist, namely gateways and MSs (user terminals). The MSs are of variable sizes ranging from handheld to transportable. The commercial applications included in such a system can be classified as Fixed Satellite Services

---

[1]Please note that the work presented with respect to these two systems took place in the beginning of 1998

(FSS), Direct Broadcast Satellite (DBS), and Mobile Satellite Services (MSS).
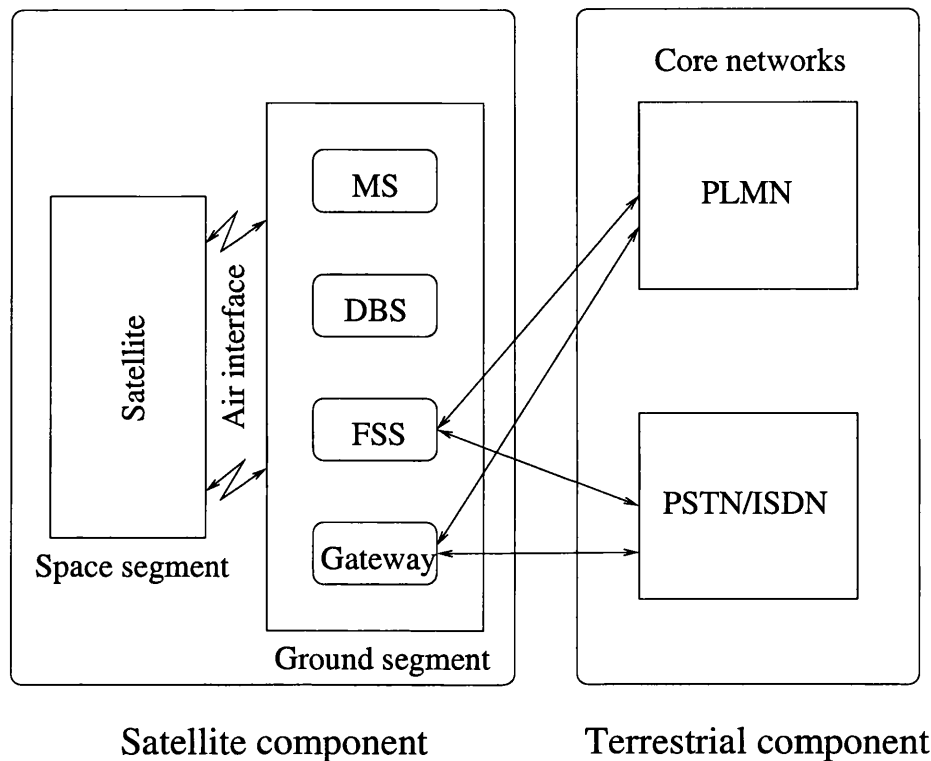
Figure 2.12: General components of a satellite system

The space segment is primarily made up of the satellites. A satellite can simply be a transmitter, a receiver and an antenna arrangement and is then called a 'bent-pipe' satellite. Alternatively, it can also include some On-Board Processing (OBP) capabilities which can vary from just signal improvement to routing and distributing messages via InterSatellite Links (ISLs).

The area on the earth's surface where the transmissions from a satellite can be received is called the satellite footprint. There are two kinds of satellite footprints: 'satellite-fixed cell' and 'earth-fixed cell' and they will be presented in more detail in section 2.4.4. If the satellite antenna allows targeting of transmissions into a fraction of the footprint area (e.g. a phased-array antenna) then the footprint area can be

divided into satellite cells covered by satellite spotbeams. S-PCSs use this cellular

concept in order to increase the capacity of the system. A big advantage of satellite

frequency reuse is that, in contrast to a terrestrial cellular system, the capacity that

is allocated to a specific cell is not permanent and can change if a beam is steered

to another direction in order to offer extra capacity to another spotbeam area.

The orbits that are mainly available for the satellite systems are the following four

ones: Low Earth Orbit (LEO), Medium Earth Orbit (MEO), GeoStationary Orbit

(GSO) and Highly Elliptical Orbit (HEO). Their main parameters are summarised

in table 2.2.

Parameter	LEO	MEO	GSO	HEO
altitude (km)	700-1000	≈10000	35786	perigee:1000 apogee:39400
period	≈100 min	≈350 min	1 day	0.5 day
inclination (degrees)	0-90	0-90	0	60-70
elevation angle	low	moderate to high	high	moderate to high
number of satellites	>40	>10	3-4	3-4
user terminals	small	small	large	moderate
propagation delay (ms)	5-10	50-80	250	50-250
handover	frequent	infrequent	none	infrequent

Table 2.2: Different orbit parameters [11]

2.4.2 Major S-PCSs

The parameters of the most important S-PCSs are summarised in table 2.3. In

the following sections Iridium and ICO are presented in more detail due to the fact

that they are often referred to the subsequent research.

Parameter	Iridium	ICO	Globalstar
orbit	LEO	MEO	LEO
number of satellites	66(+6)	10(+2)	48(+8)
satellite visibility time (min)	11	115	16
footprint diameter (km)	4700	12900	5850
max data rate (kbps)	2.4	9.6	9.6
beams per satellite	48	163	16
circuits per satellite	1100	4500	≈3000
access method	TDMA/FDMA	TDMA/FDMA	CDMA

Table 2.3: Major S-PCSs system parameters [12]

Iridium

The Iridium system consists of a constellation of 66 LEO satellites (figures 2.13 & 2.14) in polar circular orbits, at 420 nautical miles above sea level. Each satellite generates 48 beams in a hexagonal shape (fig 2.14). These beams overlap and hence, provide global coverage. The frequent handovers due to the LEO satellites are compensated with ISLs between the satellites. In Iridium, a 4 timeslot TDMA/TDD (over the L-band) scheme is used with a Quadrature Phase Shift Keying (QPSK) modulation technique. (Fixed Earth Stations) FESs use the S-band and the C-band is not used. The bandwidth needed in the L-band is 10.5 MHz.

Iridium interworks with GSM and other PCN/PCS networks with dual standard phones or through cellular cassettes installation to satellite phones. It also enables roaming around the world while maintaining the single phone number and single bill options. One disadvantage of Iridium is that it offers speech, facsimile, paging and data services at a 2.4 kbps rate.
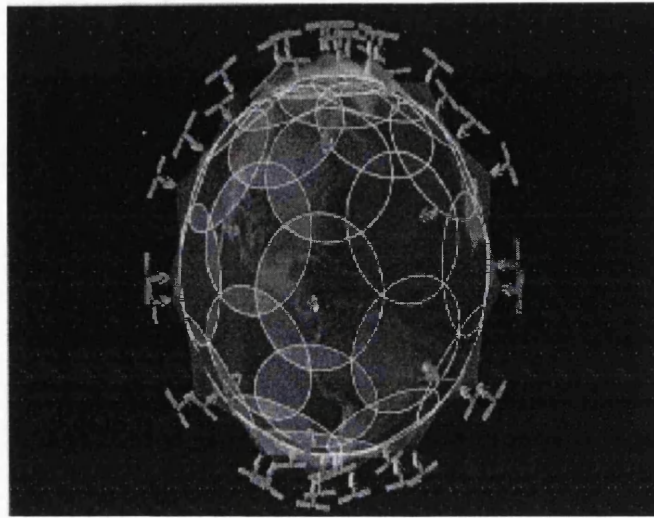
Figure 2.13: An illustration of the Iridium constellation with the solar cell arrays pointing in every way
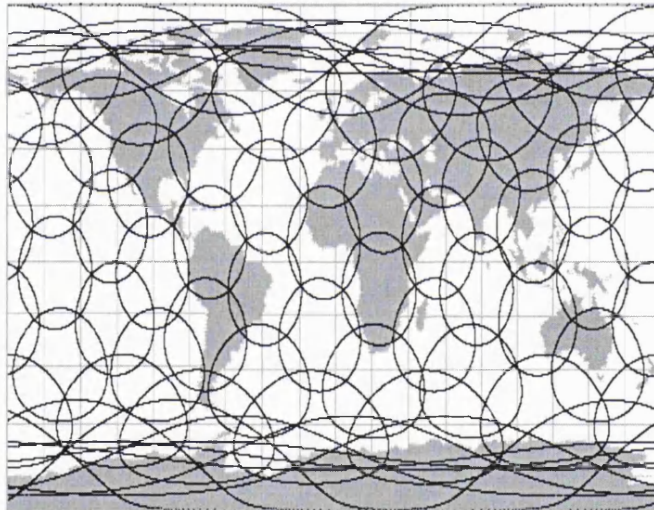
Figure 2.14: The coverage areas of Iridium satellites constellation

ICO

ICO will be a constellation of 10 satellites in MEO, 10,390 km above the earth's surface. This constellation will be arranged in two planes of five satellites each, plus one spare in each plane. In 2003 ICO is expected to offer its first commercial satellite services, after a period of extensive customer testing in 2002. Customers will be able

to choose interworking with their preferable terrestrial cellular network at purchase, by choosing one of the dual or triple mode handsets available. Cellular customers that will roam to the satellite networks will receive their satellite bill incorporated with the terrestrial one.

The ground segment (ICONET) (figure 2.15) consists of 12 earth stations located around the world and connected with high capacity systems. These earth stations function both as switches to route traffic within the ICO network, as well as inter-connecting devices to ground-based fixed and mobile networks. They also include databases for mobility management purposes. ICO's global mobility management system is based on that of GSM. Therefore, HLRs and VLRs are included in the ICONET.
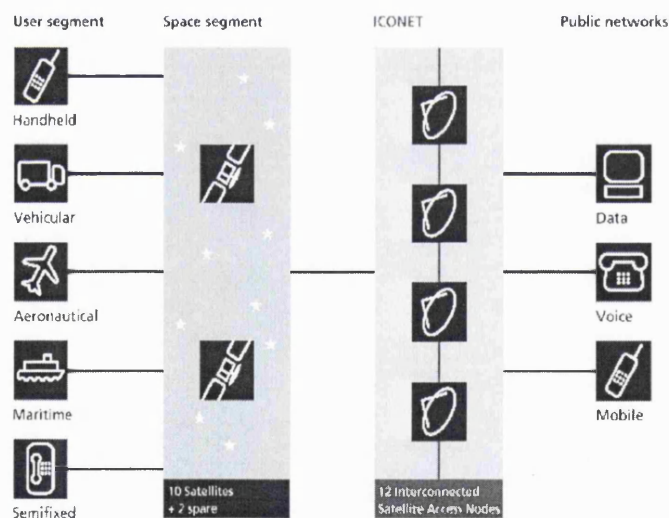
Figure 2.15: ICO system overview [13]

The satellite footprint consists of 163 service-link transmit and receive spot-beams. The satellite multiple access scheme is TDMA through which a minimum of 4,500 channels is expected to be supported per satellite. The ICO service links

will operate in the 2GHz band, while the feeder links will operate in the 5 GHz and

7 GHz bands.

2.4.3 GSM and S-PCSs integration

As already mentioned, future personal communication networks are expected

to offer the possibility of a worldwide seamless service with a single terminal and

through a single subscription. In order to achieve the former, S-PCSs are expected

to provide service extension to terrestrial communication systems, such as GSM. In

order to allow satellite systems to complement GSM, different levels of integration

between the two networks are possible. Five levels of integration have been defined

by the CCIR ("Comit Consultatif International des Radiocommunications") to act

as a basis for technical feasibility studies (final summary for COST 227 [24]). The

aim of COST 227 European action, "Integrated Space / Terrestrial Mobile Network-

s", was to 'consider feasible technical solutions able to provide, as far as possible,

the highest level of integration' [24]. The COST 227 was succeeded by COST 252

which, apart from examining the satellite and GSM integration, was also expected

to address its smooth migration to future generation mobile systems, such as Uni-

versal Mobile Telecommunication System (UMTS). The levels of integration defined

by the CCIR are described below:

- Geographical integration: In this case there is no real integration but there

  are two distinct systems based on different technologies and offering different

  services. This abstract level of integration can be perceived as having a satellite

system that is of a complementary character to the terrestrial one. Different

terminals and different subscriptions are needed.

- Service integration: In this scheme the interconnection takes place through the

  PSTN/ISDN. The user needs two subscriptions with separate numbers. There

  is no interaction between the systems as far as routing and mobility protocols

  are concerned. However, there is a common set of services that can be offered

  through the same mobile terminal.

- Network integration : In this level, there are network infrastructures common

  to the two systems. It should be possible to reach the user through a unique

  calling number. No compatibility constraints are imposed to the two different

  networks.

- Techniques integration: This level of integration is similar to the above. The

  main difference is that, in this case, terrestrial techniques, such as protocols,

  access schemes etc should be adopted by the satellite system in the closest

  format possible.

- System integration: In this case the satellite system can be considered to be

  an integral part of a unified system. With system integration there is interac-

  tion between GSM and S-PCSs even in a user connected mode. Therefore call

  continuity can be preserved through GSM to S-PCSs and vice versa through

  inter-segment handovers. These handovers are discussed in [15] and [16], and

  they were also included in the INSURED (Integrated Satellite-UMTS Real

Environment Demonstrator) ACTS European research program ([23]). IN-SURED uses the RACE SAINT[2] program outcomes as a starting point. The preliminary results of the INSURED Project have been implemented in a real environment demonstrator by integrating pre-operational satellites of a first generation mobile satellite system (IRIDIUM) with a GSM network.

In [19], the authors identify three deepest levels of integration, namely PLMN level integration, MSC level integration and BSS level integration. These are ex-amined with respect to their support of different service groups. It is shown that the largest set of services, namely mobile to mobile inter-network call, inter-network roaming and inter-network handover can be supported at either an MSC or BSS level of integration. In order to achieve an MSC level of integration, interface D has to be provided at least between the HLR of one system and the VLR of the other. Additionally, the G and the E interfaces between the two systems' VLRs and MSCs respectively, must be provided. In order to achieve a BSS level of integration, the A interface also has to be supported at least between the MSC of one system and the BSS or the Land Earth Station (LES) of the other. Having identified these two levels as being the highest levels of integration, they have been compared in [17] and [18] in terms of signalling load.

The choice of level of integration is a complex problem and depends on a number of different factors. For example, in [17], where only location updates are considered, the main conclusion is that BSS level of integration has a better performance in

---

[2]The SAINT project evaluates and identifies the requirements for the integration of satellites into future mobile networks. This applies in particular to the UMTS system in which satellite integration is expected to be mandatory [25].

terms of signalling load. In [18], call handling and inter-network handover are also included. In that case, the same conclusion can be drawn, except for cases where the ratio of the GSM coverage area to the LES service area is small and in cases where the user's average velocity is moderate.

Generally speaking, we can see that the trend in future mobile networks, 3rd generation and beyond, is to have an increasingly more independent radio access layer, such that different types of radio access technologies will connect through standardised radio independent interfaces to the core network (see [22] and [21]). In [22], the definition of the UMTS Radio Access Network (URAN) includes the UMTS Satellite Radio Access Network (USRAN) in addition to two terrestrial access, namely the UMTS Terrestrial Radio Access Network (UTRAN) and the Broadband Radio Access Network (BRAN). This could be considered as a kind of integration similar to the BSS type, discussed above for GSM.

2.4.4   'Satellite-fixed cell' and 'earth-fixed cell' satellite systems

Non-GEO satellite constellations can be divided into two categories, namely the traditional 'satellite-fixed cell' systems, and the 'earth-fixed cell' ones. Systems that belong to the first category consist of satellites that orbit around the earth along with the cells which their multibeam antenna forms on the earth's surface. The users of this system go through different types of handover. Specifically, there is the satellite handover, the beam handover and the channel handover. The first two

types of handover are the most important and frequent ones, and they occur because of the satellite's motion relative to the earth. Hence, a satellite user has to change cell (beam) and satellite during his usage time of the radio link very frequently in a MEO system and even more frequently in a LEO one. Because of the fact that this type handover happens very often in a satellite connection, the probability of failing to assign new resources has to be kept to very low levels in order to meet the desired QoS.
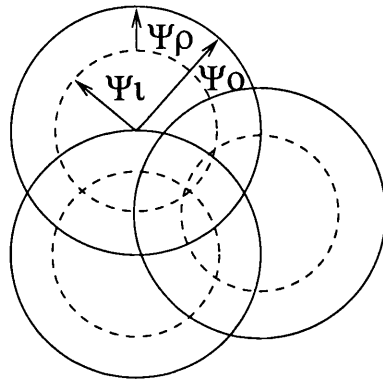
Figure 2.16: The footprint pattern for earth-fixed cell satellite systems

In order to overcome the undesired consequences of handovers, the concept of the 'earth-fixed cell' satellite system was introduced. In such a system the cells are fixed regions on the earth's surface and they do not move along with the satellite. Specifically, while in a 'satellite-fixed cell' system the cell always corresponds to the same beam and satellite, in an 'earth-fixed cell' system a cell corresponds to a specific beam and satellite only for a fixed amount of time. This is achieved with multibeam antennas that are steered in the opposite direction of the satellite motion for this specific time interval. At the end of this interval both a satellite handover and a beam handover take place and the beams are assigned to new cells.

Hence, the blocking probability due to satellite motion handover is minimised and the traffic density variations are localised in every cell and depend only on user mobility patterns. The only cause of an unexpected interbeam handover is user mobility, just like in a terrestrial system. However, in the existing satellite systems this type of handover is much less frequent than in terrestrial ones since the sizes of the satellite cells are much greater in comparison to the terrestrial ones.
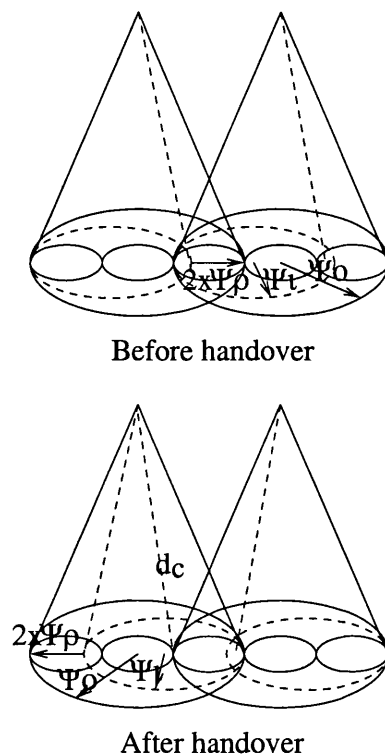
Figure 2.17: Handover between earth-fixed satellite cells

The satellite footprint pattern for an 'earth-fixed cell' satellite system is represented in figure 2.16. Each satellite footprint can be considered to consist of two parts. These parts are the satellite-fixed footprint, which is shown as the circle with the solid circumference and earth-centred angle Ψ_o, and the earth-fixed footprint, which is shown as the circle with the dashed circumference and earth-centred angle

Ψ_i ([20]), where

$$\Psi_o - \Psi_i = \Psi_\rho.$$

The earth-fixed footprint is the actual 'viable' part that is providing connection to the satellite users that it covers, and the satellite-fixed footprint is the area where the satellite can offer coverage given specific system design parameters such as height and elevation angle. In figure 2.17 the simultaneous satellite and beam handover (cell switching) is illustrated. The handover occurs for the cells that are positioned in the overlapping area of two contiguous satellite-fixed footprints. This overlapping between adjacent satellite-fixed footprints results in an increase of the constellation size compared to the one needed in a 'satellite-fixed cell' system with the same design parameters.

2.5   Summary

In this chapter the systems most important to our research have been presented, and some of their key characteristics have further been analysed. Specifically, an overview of the GSM system was presented with a more detailed treatment to the radio interface. Additionally, this chapter introduced the GPRS system and its impact on the existing design of the GSM network. A special emphasis was made on the RLC/MAC protocol which is the main focus of the GPRS research. Finally, a brief presentation of today's satellite networks and two illustrative S-PCNs were also included. Two very important 'future' satellite topics, namely GSM and satellite integration, and 'fixed-earth cell' satellite systems, were also addressed due to their

relevance to the work presented in the following chapters.

In the next chapter timing synchronisation issues in an integrated GSM and S-PCS environment are addressed. Both 'satellite-fixed cell' and 'earth-fixed cell' systems will be considered.

Chapter 3

Timing Synchronisation for GSM

over a satellite system

3.1 Introduction

As discussed in chapter 2, several papers have been published on GSM and S-PCNs integration issues. However, this integration is difficult to tackle due to the different nature of the satellite and terrestrial network environment. One problem area is the network-user synchronisation in a TDMA environment, due to the long propagation and differential path delays inherent in a satellite network. In this chapter, we specifically address the timing synchronisation problems caused by such a scenario.

In the next section, background theory on the synchronisation feature as it is achieved in a GSM network, is presented. In the third and fourth sections the introduction, analysis and results of our proposed schemes are presented for a 'satellite-

fixed cell' and 'earth-fixed cell' system respectively.

Finally, conclusions are presented in the last section.

3.2   Synchronisation and timing advance feature

As already mentioned (section 2.2.3), the FCCH and the SCH are used by the MS to acquire the synchronisation. Specifically, the SCH is providing the MS with all the available data needed for obtaining information on the limits between slots as well as the number of the timeslot. In order to find an SCH burst the user only has to find an FCCH burst, since an SCH occurs eight timeslots after the FCCH on the same frequency. An FCCH burst is easily recognisable as it is a pure unmodulated sine wave.

In the transmission mode, in order to enable the slots of one channel to bear the same timeslot number in both directions, the uplink slots are shifted from the downlink ones by three burst periods. Specifically, after receiving a channel with a particular timeslot number, a mobile station emits its data 3 burst periods later under the same timeslot number. However, the propagation delays involved in a cell, hinder the maintenance of an exact 3 burst period shift at both BSS and MS. This could result in the overlapping of bursts that are sent by users transmitting in adjacent slots.

Hence, the feature of Timing Advance (TA) is introduced in a GSM network. According to this feature, the mobile station is doing a shift of 3 burst periods minus the TA time period. The TA time period is compensating for the bidirectional

propagation delay. It is calculated at the base station and is then provided to the users through the associated control channel.

However, in the registration phase the TA is not known to the user. In this initial part of communication, the user, using an access burst, is sending the Random Access CHannel (RACH) in GSM, or the Packet Random Access CHannel (PRACH) in GPRS, to request assignment of resources. Since the TA is not known at that point, the reception window has to be large enough to accommodate the (P)RACH transmission delay plus twice the maximum differential propagation delay from any two points in a cell. After the transmission of the (P)RACH burst, the base station can calculate the TA, which is constantly updated until the end of the call and is sent to the mobile to enable it to transmit the rest of the data correctly positioned within the system's timing window.

The maximum possible differential path imposes a maximum differential propagation delay between users in a cell. The following equation gives the maximum differential path x_{dp} that can be used in a mobile system, given some of the system characteristics:

$$x_{dp} = \frac{T_s - \frac{Pr_b}{Ch_r}}{2} c, \tag{3.1}$$

where T_s is the duration of the TDMA frame timeslot, Pr_b is the number of bits used for (P)RACH, Ch_r is the mobile channel rate and c is the speed of light.

According to equation 3.1, it has been calculated that in GSM the reception window can accommodate mobiles that have 35 km maximum differential path.

Proposals about the initial synchronisation, the staggering between uplink and downlink frames and the TA feature have been published in the past ([15],[31]). All proposed schemes suggest changes in the GSM access scheme in order to accommodate the satellite environment differential path delay. An alternative approach would be to achieve synchronisation without altering the GSM radio interface. This could be accomplished by subdividing the footprint in such a way that it fulfils the path length requirements. In the following sections, this approach of a footprint pattern will be analysed for two different types of satellite systems. The first system is a satellite-fixed cell type and the second one is an earth-fixed cell type.

3.3 Satellite-fixed cell pattern

In this type of satellite system, the footprint is divided into concentric rings in such a way that the differential path delay from any two points within the same ring is smaller than or equal to the differential path delay allowed for the use of RACH. The centre of the rings coincides with the centre of the footprint.

3.3.1 System analysis

The half angle of the satellite footprint is given by (fig 3.1, [32]) :

$$\psi = \frac{\pi}{2} - \epsilon_{min} - \arcsin(\frac{r}{r + h} \cos \epsilon_{min}), \qquad (3.2)$$

where ϵ_{min} and h are the elevation angle and the height of the satellite respectively and r is the earth's radius.
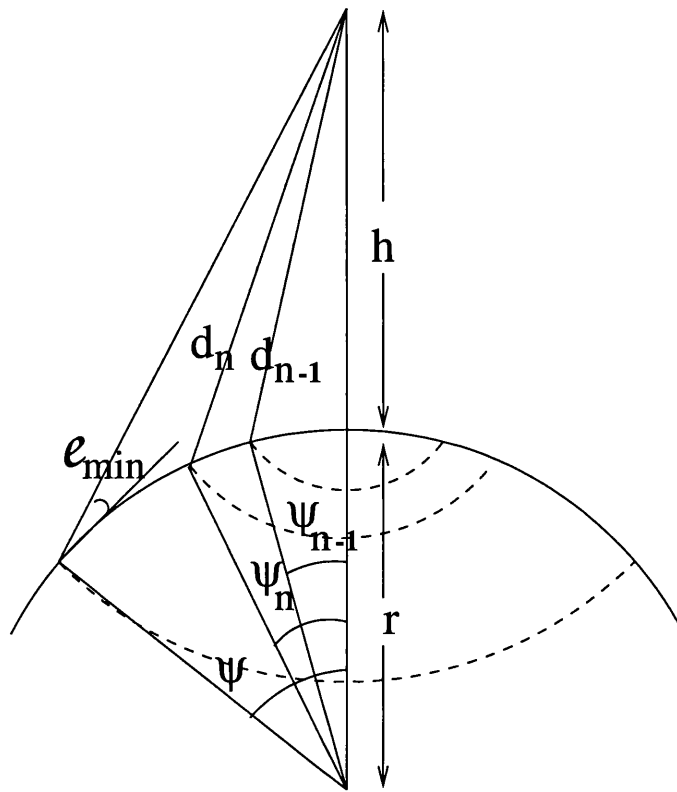
Figure 3.1: Geometrical considerations

The area ϵ of a spherical cap of radius r and height H is given by the following

formula:

$$\epsilon = 2\pi r H, \tag{3.3}$$

Therefore, the total area $\mathcal{E}_{tot}$ of the footprint and the ring area $\mathcal{E}$ are given by:

$$\mathcal{E}_{tot} = 2\pi r^2 (1 - \cos\psi), \quad \mathcal{E} = 2\pi r^2 (\cos\psi_{n-1} - \cos\psi_n), \tag{3.4}$$

By applying the law of cosines we calculate the distance from the outer circum-

ference of each ring to the satellite as:

$$d_n = \sqrt{r^2 + (r+h)^2 - 2r(r+h)\cos\psi_n}, \tag{3.5}$$

Two approaches are used for calculating the number of rings needed:

i In the first approach the rings are of equal area to ensure a uniformly spread user density pattern. In this approach every ring can be considered as a cell, or, for realisation purposes, every ring can be divided into a number of cells. This approach also facilitates a uniform frequency reuse pattern. To analyse this approach, we evaluate the area $\mathcal{E}_1$ of the inner circle from (3.4), where $\psi_o = 0$ $(n = 1)$. If $n = \left\lceil \frac{\mathcal{E}_{tot}}{\mathcal{E}_1} \right\rceil$ is the number of rings then the two following conditions must be satisfied:

1. $\mathcal{E}_i = \mathcal{E}_{i+1} = \mathcal{E}$ $\forall i \in \{1, 2, ...(n-1)\}$

2. $d_i = \sqrt{r^2 + (r+h)^2 - 2r(r+h)\cos\psi_{i-1}} + y$ $|$ $y \leq x$,

$$\psi_{i-1} = \begin{cases} \arccos\left(1 - \frac{\mathcal{E}_{i-1}}{2\pi r^2}\right), & i = 1 \\[2mm] \arccos\left(\cos(\psi_{i-2}) - \frac{\mathcal{E}_{i-1}}{2\pi r^2}\right) & \forall i \in \{2, ...n\} \end{cases}$$

Specifically, the first of the above conditions ensures that all the rings are of equal areas, while the second one ensures that the differential path from any two points in the ring is not more than the value allowed.

ii In the second approach the minimum number n_{min} of rings is used to divide the footprint into rings of unequal areas. In this case, the only constraint is that $d_n - d_{n-1} = x_{dp}$ and $d_{n_{max}} - d_{n_{max}-1} \leq x_{dp}$, if x_{dp} is the maximum differential path allowed.

Therefore the minimum number of rings n_{min} can be calculated as:

$$n_{min} = \left\lceil \frac{d_{n_{max}} - h}{x_{dp}} \right\rceil.$$ (3.6)

3.3.2 Results and discussion

We have calculated the number of rings for a pure GSM channel (x_{dp}=35 km) over the satellite using the set of equations from the first approach. The results are plotted in figure 3.2 as solid lines. The dashed lines in the same figure show the minimum number of rings needed, calculated using equation 3.6. Furthermore, two points for the ICO and Iridium type of constellations have been plotted in the above figure to enable a comparison to be made. It can be seen that for lower altitudes, by using the second scheme, we need significantly less rings than with the first one. The latter case is leading to the use of bigger rings and could lead to the use of bigger cells and fewer resources for the same spotbeam size. For a MEO constellation the results for both approaches converge.

elevation angle (deg)	10	20	30	40	50	60
α for height=800 km	2.88	2.15	1.69	1.40	1.2	1.08
α for height=10000 km	1.39	1.30	1.21	1.14	1.09	1.05

Table 3.1: The ratio $\alpha = \frac{\xi_o}{\xi_n}$ for different satellite altitudes and elevation angles

The reason why values converge when satellite height is increasing, especially for a small number of rings, is that for higher altitudes the areas of the inner and the outer rings become almost equal. Therefore, the fact that rings have equal area

in the first scheme does not have any impact on the results when compared to the second scheme. Specifically, if $\mathcal{E}_o$ and $\mathcal{E}_n$ are the areas of the inner and outer ring respectively, then the ratio $\alpha = \frac{\mathcal{E}_o}{\mathcal{E}_n}$ for different satellite altitudes and elevation angles is given in the table 3.1.

The following equation ([32]) gives the number of satellites N needed given the height and the elevation angle:

$$N = \left\lceil \frac{2\pi}{3\Psi} \right\rceil \times \left\lceil \frac{2\pi}{\sqrt{3}\Psi} \right\rceil, \tag{3.7}$$

where Ψ is the half-sided footprint centre angle calculated from equation 3.2.

Figure 3.3 shows the number of satellites needed (equation 3.7, [32]) as a function of orbit height, for the same values of n, as in figure 3.2. The number required is seen to be practicable if we compare the results of the proposals presented here to commercial systems such as Iridium and Teledesic.

In order to give some indicative values of the cell sizes that will be used in such a scheme, we calculated the distance between consecutive rings. Specifically, in figure 3.4 this distance is presented for two different sets of values for height and elevation angle. The dashed line shows the results for a LEO constellation of 800 km height and 10° elevation angle, while the solid one shows the results for a MEO one of 11000 km height and 30° elevation angle. It can be observed that, for example in the LEO one, the outer rings have a distance of 50 km between them, which leads to the use of cells with a maximum radius of 25 km.

In order to assess further the implications of our design, we have estimated
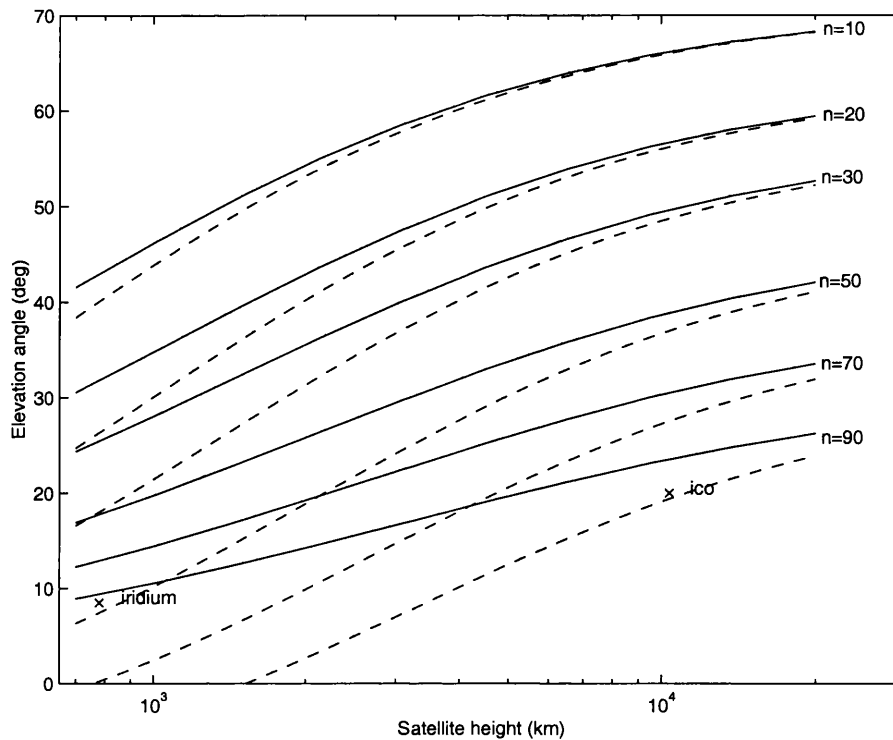
Figure 3.2: The number of rings used for x_{dp}=35 km

the number of cells that are required in such a scheme. Using the distance values

between successive rings as plotted in figure 3.4, it was possible to calculate the ring

area using equation 3.4. We assume that for a given ring of width D, the number

of cells required within that ring can be estimated by dividing the ring area by the

area of a reference circular area. The approach outlined below can be considered

as a simple theoretical estimation for the number of cells required. If one thinks of

hexagonal cell patterns, whose height is equal to D, we consider two cases: in the

first case, the reference circle represents the inscribed circle of the hexagonal cell,

while in the second approach, the reference circle represents the circumscribed circle

for that hexagon. The first case can be considered to yield an upper limit for the

minimum number of cells required, while the second case yields a lower limit, as it
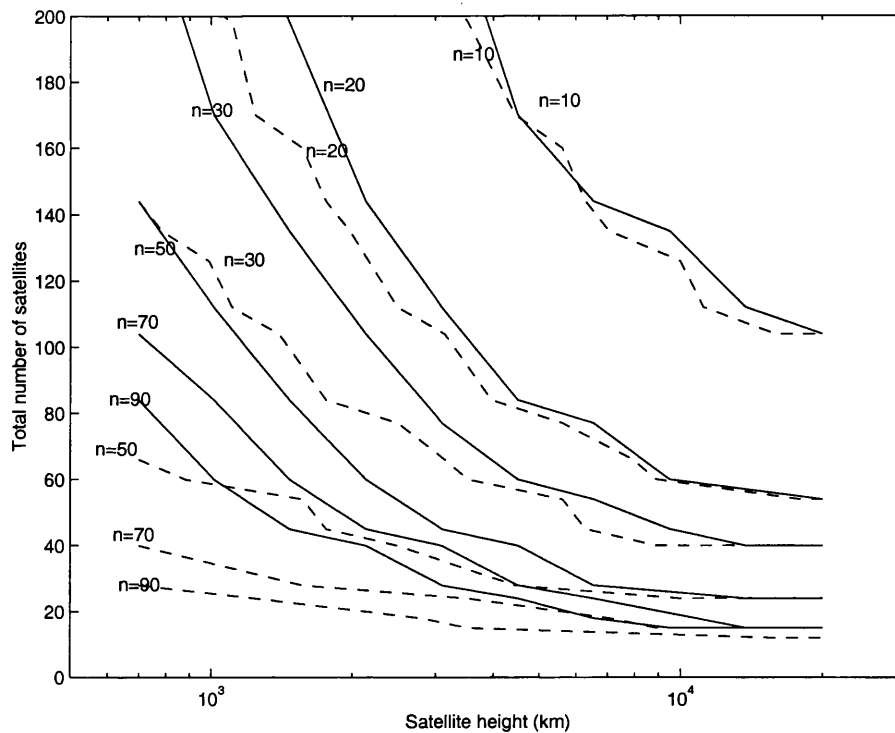
Figure 3.3: The number of satellites used for x_{dp}=35 km and various sets of n

takes into account some overlap between adjacent cells.

An important parameter to consider in any such design is the number of rings that are required within the footprint of the considered satellite system. For example, in the case of Iridium where 66 satellites were required, we can see in figure 3.2 that approximately 50 rings will be required with the second scheme. From our estimations it can be seen that this would require between 10000 and 14000 cells per satellite. On the other hand, with a Teledesic-type of constellation where 288 satellites, each one with a 700 km footprint radius [26], will be deployed, 7 rings are sufficient to cover the footprint. This translates to a minimum requirement of between 200 and 264 cells per satellite.

From figure 3.4, we can see that all cells from ring number 5 upto the outermost
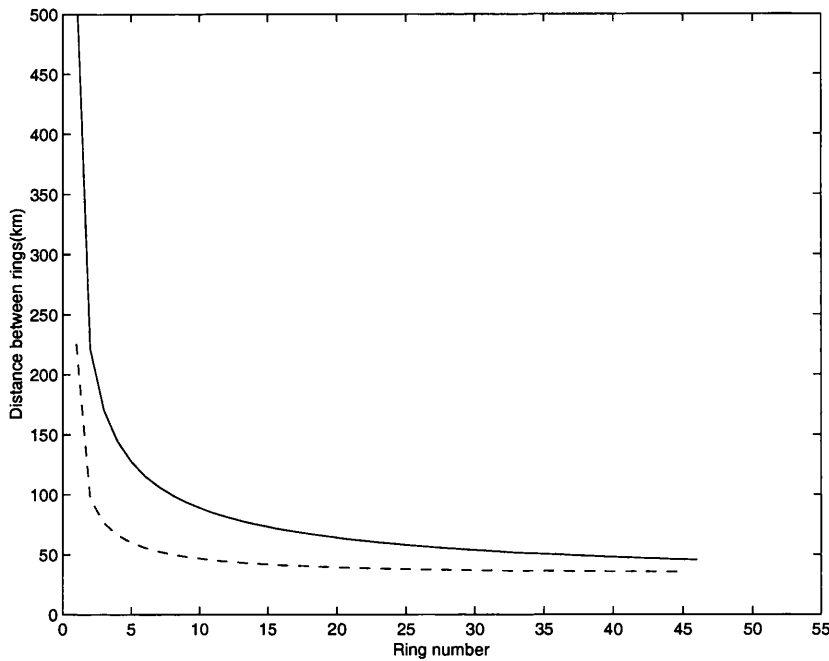
Figure 3.4: Distance between consecutive rings

ring for a LEO Iridium-like system, that follows the suggested design, would have a radius of approximately 25 km. Similarly, for a MEO ICO-like system the majority of cells would have a radius between 50 and 30 km. The cell radii mentioned above are considerably smaller than those of the most representative 2nd generation satellite systems. Specifically, the spotbeam diameter would be in the range of 700 km for Iridium with 48 beams per satellite, while for ICO each spacecraft will produce 163 beams with cell diameters of about 800 km [29]. The difference in size and number of cells obtained from our approach and those of conventional systems will increase the design complexity of our proposed satellite system. The most important design issues are described and briefly discussed below. Additionally, since these complexity design issues justify a work item in their own right, they are also mentioned in the final chapter as suggestions for further work.

I) Signalling load: It can be expected that the number of handovers during a call is inversely proportional to the cell radius. Therefore, a reduced cell size would lead to an increased amount of signalling traffic that will be associated with every user in the case of handover and location area updates. An initial solution to increase the cell size in the suggested system, in order to mitigate the effects on signalling load, would be to use a bigger timeslot (by combining two normal timeslots) for sending the PRACH message, thus increasing the differential path delay, for users that would be served by cells inscribed within the outer rings of the spotbeam. This leads to a reduction in the number of rings required, and a corresponding increase in the size of these cells. This has been discussed in [31]. Additionally, this has been adopted in GSM, where in some rural area cells, the x_{dp} can be increased by obtaining a huge guard time by using only every second burst [34]. In such cases, only the timeslots with an even number are used, but this scheme excludes timeslot number zero since this has to be reserved for BCCH. Of course, when adopting this scheme, a specific reception processing in the BTSs is required. Additionally, it should also be pointed out that this technique reduces the number of channels per MHz and this is the reason why this is applied to a rural area environment.

A more radical approach would be to subdivide the ring area into a smaller number of cells, which would not be of circular shape but of some kind of trapezium shape (see figure 3.6). This would satisfy the propagation delay criterion, since the user could move anywhere within a ring and still meet this criterion. Such an approach would certainly alleviate the problem to some

extent, especially when used in conjunction with the aforementioned solution. An optimal antenna configuration producing similar beamshapes, placed in a ring structure, has been presented by Kenneth N. Sherman using Genetic Algorithm methods in [27]. In this particular piece of work, the author is producing such cell layouts while optimising antenna gain requirements. Specifically, figure 3.5 is copied from the specific publication and is illustrating different cell pattern layouts for different number of rings and directivity requirements.

II) System capacity and architecture: Cell size is directly linked to system capacity. The total system capacity depends on the satellite power. On the other hand, the capacity per unit area (density) depends on spectrum reuse and total bandwidth. Both of the above parameters depend on the cell size. As mentioned above, in the emerging 2nd generation mobile satellite systems, the minimum diameter of a cell is several hundreds of kilometres. However, the trend is to reduce this in the next generation of satellite systems in order to achieve higher capacity densities. It is therefore necessary to strike the right balance between the power required per satellite and the capacity density achieved.

In our system the cell radius for an Iridium-like satellite spotbeam could be as small as 25 km, and for an ICO-like system between 30 and 50 km. As discussed above, it is likely that the diameter of a spotbeam for future systems can drop to 100 km by year 2005 [8]. Teledesic is also aiming at increasing capacity for satellite users by launching a constellation of 288 LEO, for which

it is implied in [26] that hopping spotbeams will cover areas of 5km radius. Another high capacity US-proposed satellite system, NetSat28, will use fixed beams of 112 km diameter approximately. Such systems would require large global capacity. Small cells will be introduced to achieve sufficiently high capacity densities but in order to reduce the power consumption per satellite a larger number of satellites is expected.

As discussed previously, when a system of the same magnitude as the Teledesic one is used for comparison purposes, we can attain a considerable improvement in the number of rings required, when compared to an Iridium-like system. Specifically, as mentioned for the second scheme, we can reduce this number from 50 to 7. This, of course, would not lead to increased cell sizes as the spatial difference between consecutive rings remains the same, but would certainly reduce the levels of power per satellite required since the number of cells per satellite decreases significantly. However, the average cell size would increase (look at figure 3.4) since the satellite would cover smaller footprint areas and therefore there would be fewer cells at the edges, and these are the smallest ones. Nevertheless, as pointed out in the previous discussion, regarding the cell size and signalling overhead, there could be other ways to adopt our technique with increased cell sizes.

It is worth noting that since our model applies to 2nd generation systems the economics might not justify the substantial investment that such systems, designed for the more profitable 3rd generation systems, would require. However, it could
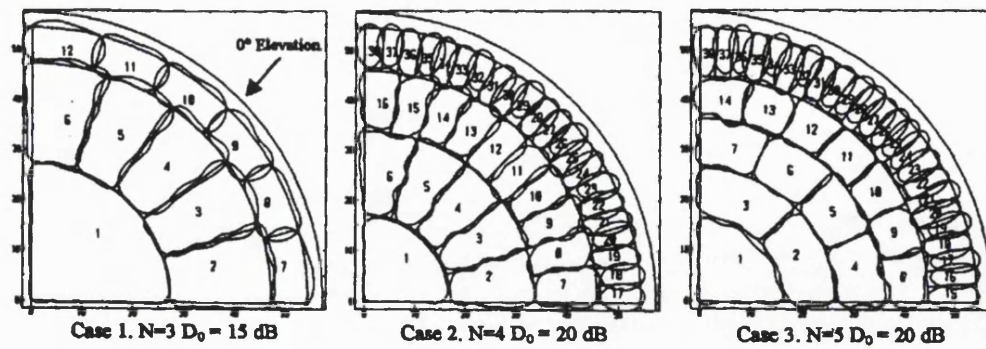
Figure 3.5: Different cell pattern layouts for different number of rings and directivity requirements [27]
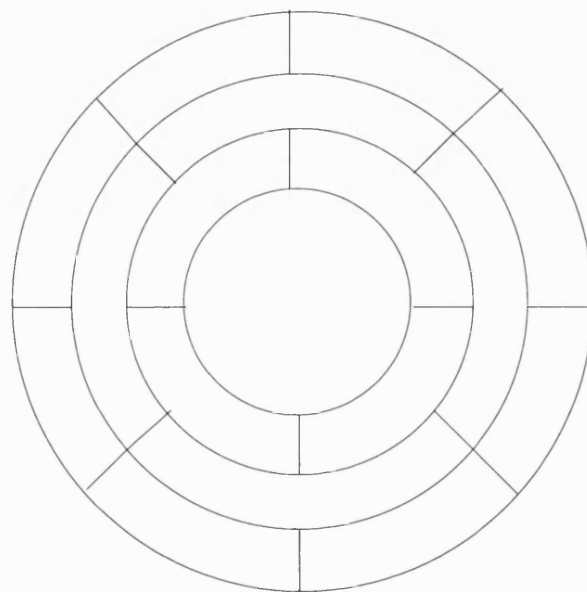
Figure 3.6: Example of a satellite footprint design using trapezium shaped cells

be claimed that the technology to do this will be soon available, although the economics of such systems will have to be carefully studied. Additionally, the principles behind the timing synchronisation methodology suggested here could be applied to other types of systems, and hence the technique could be viewed as providing an application-independent framework.

Hence, as mentioned above there are significant trade-offs associated with the

introduction of the design proposed in this chapter. A topic for further study could involve weighing the benefits and disadvantages of such a solution that is very sensitive to certain design parameters and economic factors, and is expected to be case specific.

Additionally, the approach of creating frequency reuse cellular patterns, while at the same time satisfying the timing constraint by introducing a design based on concentric rings, introduces some new challenges to the system designer. For example, the problem of fitting hexagonal cells, or patterns thereof, or even cells of novel shapes, such as trapeziums (figures 3.6, 3.5) into the reference circular area has not been extensively addressed in this work. Further investigation into this problem and the development of some design rules to address it, are topics that could be considered in further work.

Finally, another option that could be investigated would be altering the RACH bursts to match the frame structure and channel rate used in illustrative satellite systems. The required differential path would be calculated using equation 3.1. Using in this equation the T_s and Ch_r of Iridium and ICO systems ([33],[35]) and the $Pr_b = 88$ bits for GSM (P)RACH, we have calculated n and n_{max} (see table 3.2).

Constellation	T_s (ms)	Ch_r (kbps)	x_{dp}(km)	n	n_{min}
Iridium	8.28	50	1059	4	2
ICO	6.67	36	633	6	5

Table 3.2: Application to Iridium and ICO channels

Comparing the results from table 3.2 and figure 3.2, it is clear that by adopting the satellite systems frame structure and channel rates we can attain a significant reduction in the number of cells in the footprint. On the other hand, to adopt this approach, we have to introduce changes to the GSM protocol to match the characteristics of the satellite channel. However, in both cases, the PRACH burst can fit into one timeslot without having any differential path delay and synchronisation problems between the network and users in the same cell.

3.4 Earth-fixed cell pattern

As the earth-fixed footprint area is moving with respect to the satellite-fixed footprint area, it is 'seen' from the satellite from different angles during the interval between satellite handovers. Therefore, in specific areas, the TA demand has to be met for all the positions of the satellite with respect to the earth-fixed footprint area, during this specific interval. In figure 3.7 all the positions of the satellite during this interval are located between points α and β. In this figure, the circle marked with the solid line is the earth-fixed footprint area with centre κ and the circles marked with the thinner and the thicker lines are the concentric rings with centres α and β respectively. As already mentioned, these two points are the two 'edge' positions within the satellite-fixed footprint area, that the earth-fixed footprint area centre could occupy between handovers. Therefore, the areas in which the differential path delay requirements for every position of the satellite between handovers are satisfied, are formed by the intersection of concentric rings with centres at points α and β as

illustrated in figure 3.7. For example, all the shadowed subareas form the λ area in both halves of the satellite-fixed footprint.
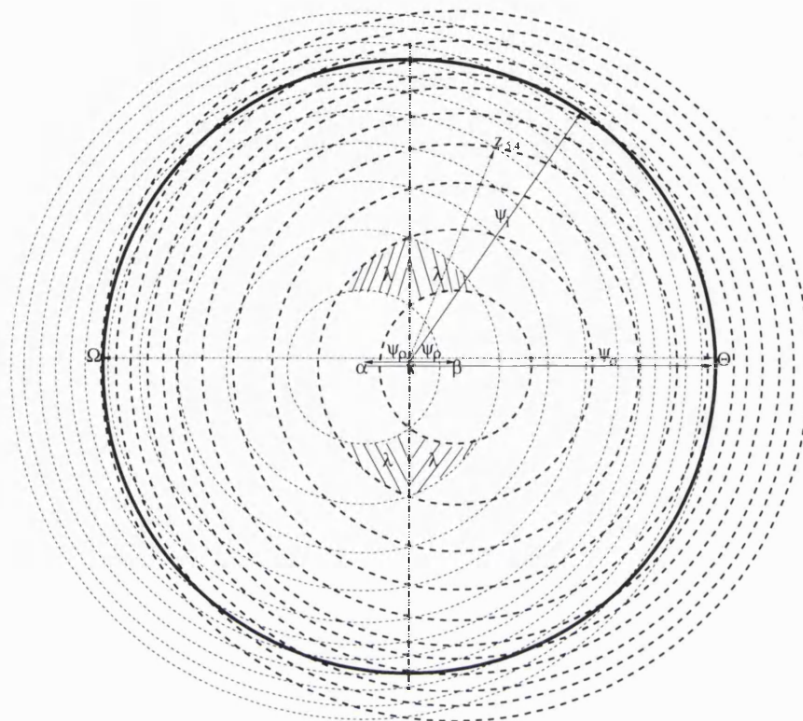
Figure 3.7: Earth-fixed footprint division

3.4.1 System analysis

In the analysis below we consider that $\frac{\Psi_\rho}{2} \leq \Psi_i$. For larger values of Ψ_ρ the number of satellites needed becomes huge, and the system becomes non-realistic [20].

Figure 3.8 illustrates part of figure 3.7, seen from a different angle. Specifically, Ψ_i, Ψ_o and Ψ_ρ are the earth-centred angles as presented in section 2.4.4. In this particular figure, the satellite is moving along the route marked with A, B and C. In the same figure, h and e_{min} are the height and the elevation angle of the satellite
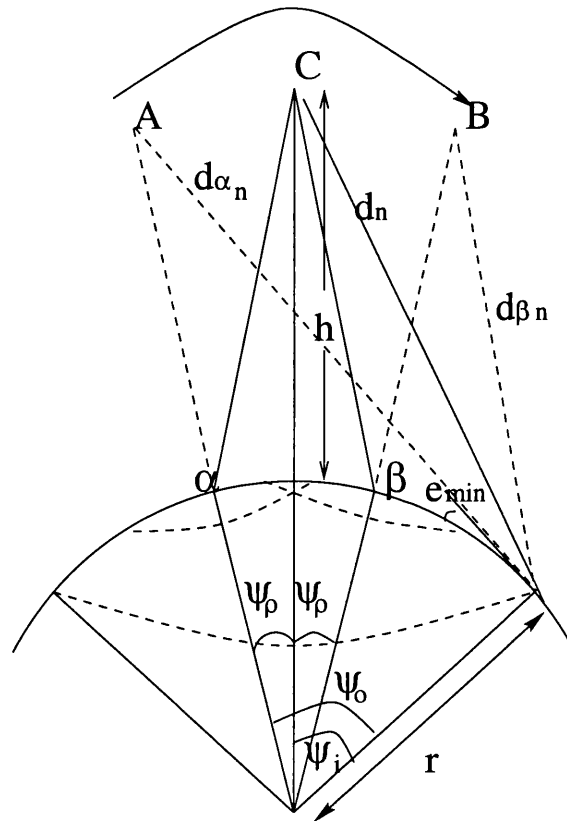
Figure 3.8: Geometric considerations for the 'earth-fixed' pattern

system respectively. Additionally, d_{α_n} is the longest distance of the satellite, when

it is located at point A, from the 'satellite-fixed' footprint circumference, while d_{β_n}

is the shortest distance of the satellite, when it is located at point B, from the

'satellite-fixed' footprint circumference. This distance is d_n when the satellite is

located at point C.

As the particular problem is more complicated than the one in the 'satellite-

fixed cell' case, we choose to calculate the number n_{min} of arcs for each footprint

half, based on the second approach from section 3.3.1. In the following analysis,

we choose to calculate the number of areas for the right half of the footprint. The

number of rings with centre α, whose part is included in this half is equal to:

$$n_\alpha = \left\lceil \frac{d_{\alpha_n} - h}{x_{dp}} - (n_{\alpha_{cr}} - 1) \right\rceil, \tag{3.8}$$

where $n_{\alpha_{cr}}$ is the number of the first ring for which $\psi_{\alpha_{n_{cr}}} > \Psi_\rho$ and $\psi_{\alpha_{n_{cr}}}$ is the earth angle between centre α and the outer circle of the $n_{cr_{th}}$ ring with centre α.

A general equation giving ψ_{α_i} for any i_{th} ring is given after combining equations 3.5 and 3.6:

$$\psi_{\alpha_i} = acos \left(\frac{(ix_{dp} + h)^2 + r^2 + (r + h)^2}{2r(r + h)} \right), \tag{3.9}$$

The number of rings with centre β, whose part is included in this half is equal to:

$$n_\beta = \left\lceil \frac{d_{\beta_n} - h}{x_{dp}} \right\rceil. \tag{3.10}$$

Let us also define as $z_{x,y}$ the point where the x_{th} ring with centre α and the y_{th} ring with centre β are met. Then, $\kappa z_{x,y}$ is the earth-centred angle between this point and the centre of the earth-fixed footprint κ (figure 3.7) and can be calculated if we apply the law of cosines to the spherical triangles $\alpha z_{x,y}\beta$ and $\alpha z_{x,y}\kappa$. Specifically, if we apply the law of cosines first to $\alpha z_{x,y}\beta$ we obtain the $\widehat{\alpha}$ angle:

$$cos(\widehat{\alpha}) = \frac{cos(z_{x,y}\beta) - cos(\alpha\beta)cos(\alpha z_{x,y})}{sin(\alpha z_{x,y})sin(\alpha\beta)}, \tag{3.11}$$

Additionally, if we apply the law of cosines to $\alpha z_{x,y}\kappa$ we get:

$$cos(z_{x,y}\kappa) = cos(\alpha z_{x,y})cos(\alpha\kappa) + sin(\alpha z_{x,y})sin(\alpha\kappa)cos(\widehat{\alpha}), \qquad (3.12)$$

If we combine equations 3.11 and 3.12 we obtain $\kappa z_{x,y}$ which is given by the following equation:

$$\kappa z_{x,y} = acos\left(cos(\alpha z_{x,y})cos(\alpha\kappa) + \frac{cos(z_{x,y}\beta) - cos(\alpha\beta)cos(\alpha z_{x,y})}{2cos(\alpha\kappa)}\right), \qquad (3.13)$$

where $\alpha\kappa = \frac{\alpha\beta}{2} = \Psi_\rho$, $\alpha z_{x,y} = \psi_{\alpha_x}$ and $\beta z_{x,y} = \psi_{\beta_y}$. The last two earth angles can be calculated from equation 3.9.

Taking all the above definitions into consideration, the total number of areas is given by:

$$\sum_{i=1}^{n_{\beta cr}-1} A_1(i) + \sum_{i=n_{\alpha cr}}^{n_\alpha} A_2(i), \qquad (3.14)$$

where $n_{\beta_{cr}} = n_{\alpha_{cr}}$.

The $A_1(i)$ factor is for all the zones that are created because of the rings with $\psi_{\beta_i} \leq \Psi_\rho$. An example of such a ring is the thick solid small ring in figure 3.9. If we assume that the α_k up to the α_l rings 'meet' the particular β_i ring then $\lambda - kappa + 1$ zones are created. Therefore:

$$A_1(i) = \lambda - \kappa + 1, \quad if$$

$$\psi_{\alpha_{\kappa-1}} - \Psi_\rho \leq \Psi_\rho - \psi_{\beta_i} < \psi_{\alpha_\kappa} - \Psi_\rho, \quad \psi_{\alpha_{\lambda-1}} - \Psi_\rho < \Psi_\rho + \psi_{\beta_i} \leq \psi_{\alpha_\lambda} - \Psi_\rho$$
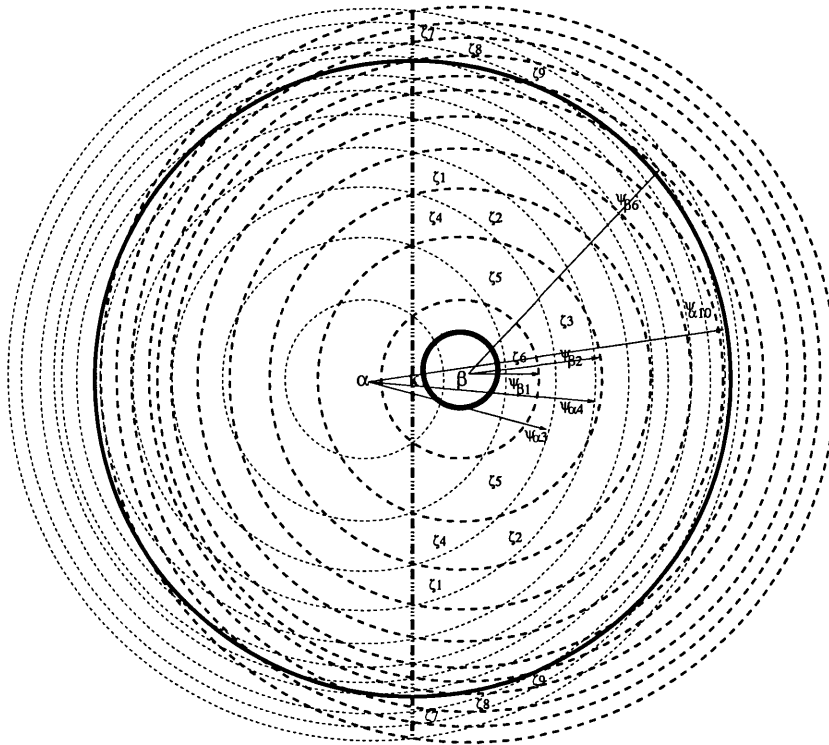
Figure 3.9: Earth-fixed footprint division

Additionally, the $A_2(i)$ factor is comprised from four subcases. Specifically:

$$A_2(i) =$$

$$\begin{cases} 1, & if\ i = n_{cr} \\[2mm] i-j+1, & if\ \psi_{\beta_j} + \Psi_\rho \geq \psi_{\alpha_i} - \Psi_\rho > \psi_{\beta_{j-1}} + \Psi_\rho\ and\ \psi_{\beta_{j-1}} + \Psi_\rho \leq \psi_{\alpha_{i-1}} - \Psi_\rho \\[2mm] i-j+2, & if\ \psi_{\beta_j} + \Psi_\rho \geq \psi_{\alpha_i} - \Psi_\rho > \psi_{\beta_{j-1}} + \Psi_\rho\ and\ \psi_{\beta_{j-1}} + \Psi_\rho > \psi_{\alpha_{i-1}} - \Psi_\rho \\[2mm] A_2(i) - (i - \mu) - x, & if\ \kappa z_{i,i} > \Psi_i,\ \kappa z_{i,\mu} \geq \Psi_i > \kappa z_{i,\mu-1}\ and\ \kappa z_{i-1,\mu+x} \geq \Psi_i > \kappa z_{i-1,\mu+x-1} \end{cases}$$

The first subcase is obvious. In the second subcase, where the $psi_{\alpha_i} - \Psi_\rho$ angle is smaller than the $\psi_{\beta_j} + \Psi_\rho$ angle, then if $\psi_{\beta_{j-1}} + \Psi_\rho \leq \psi_{\alpha_{i-1}} - \Psi_\rho$ the α_i ring 'meets' the β rings $i - j + 1$ times and therefore $i - j + 1$ zones are created. An example can be seen in figure 3.9 for $i = 4$ and $j = 2$ where the zones ζ_1, ζ_2 and ζ_3 are created. In this figure, for graphical simplicity, we assume that $\psi_{\beta_{ncr}} = \psi_{\beta_1}$ and $\psi_{\alpha_{ncr}} = \psi_{\alpha_1}$.

In the third subcase, where $\psi_{\beta_{j-1}} + \Psi_\rho > \psi_{\alpha_{i-1}} - \Psi_\rho$ the points where the α_i ring meets the β rings are $i - (j - 1) + 1$ and therefore the number of zones is $i - j + 2$. An example can be seen in figure 3.9 for $i = 3$ and $j = 2$ where the zones ζ_4, ζ_5 and ζ_6 are created.

In the fourth subcase, we can either have a second or a fourth subcase but we have to subtract the number of zones that are located outside the (κ, Ψ_i) circle. An example is illustrated in figure 3.9 for $i = 10$, $x = 1$ and $\mu = 6$ and the zones ζ_7, ζ_8 and ζ_9 are subtracted.

In the above equations, ψ_{α_x} and ψ_{β_y} are the earth centred angles that can be calculated from equation 3.4 and correspond to the x_{th} ring with centre α and the y_{th} ring with centre β respectively.

3.4.2 Results and discussion

In figures 3.10 and 3.12 the number of areas versus Ψ_ρ earth-centred angle is illustrated for a LEO and MEO system respectively, for three different values of elevation angle. The number of areas is obviously equal to the one calculated for a 'fixed-satellite' pattern when $\Psi_\rho = 0$, but it is increasing dramatically with Ψ_ρ. Figures 3.11 and 3.13 present the number of satellites for the same LEO and MEO systems, versus Ψ_ρ. The number of satellites is calculated from equation 3.7 for $\Psi = \Psi_o - \Psi_\rho = \Psi_i$ (figure 3.7).

In a division pattern such as this particular one, it is not simple to calculate the distance between the areas in every point of the footprint and to determine the size of the cells that would be inscribed between the borders of these areas.
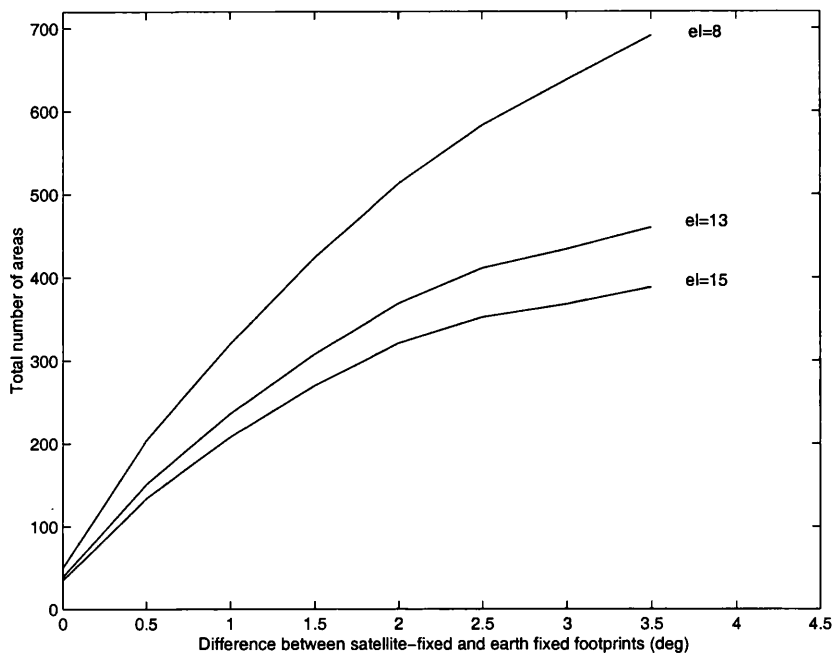
Figure 3.10: Number of areas for a LEO system (h=800 km) versus Ψ_ρ
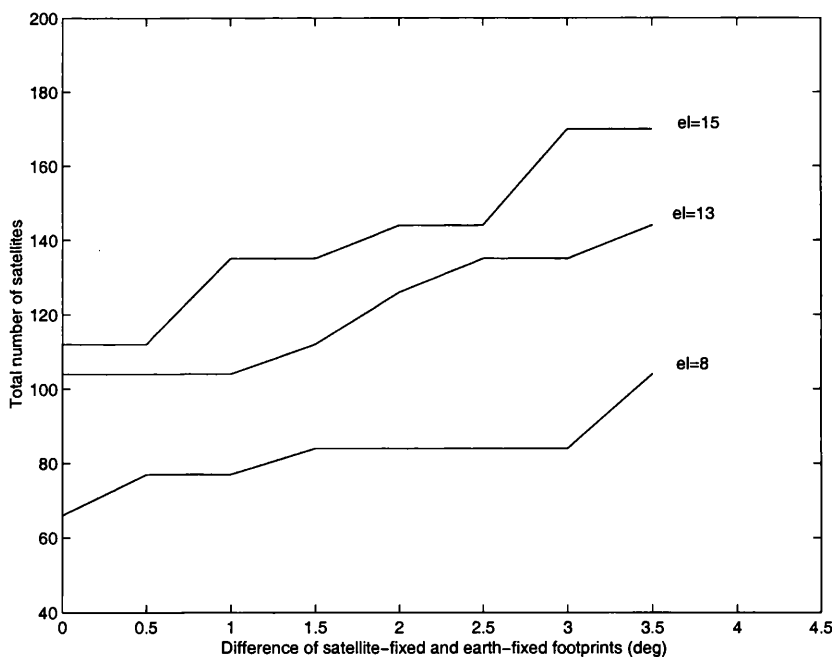
Figure 3.11: Number of satellites for a LEO system (h=800 km) versus Ψ_ρ

However, in order to provide some representative values, the maximum distance between consecutive areas that are located in the footprint $\Omega\Theta$ diameter (figure 3.7) are presented for a LEO and a MEO system respectively in figures 3.14 and 3.15.
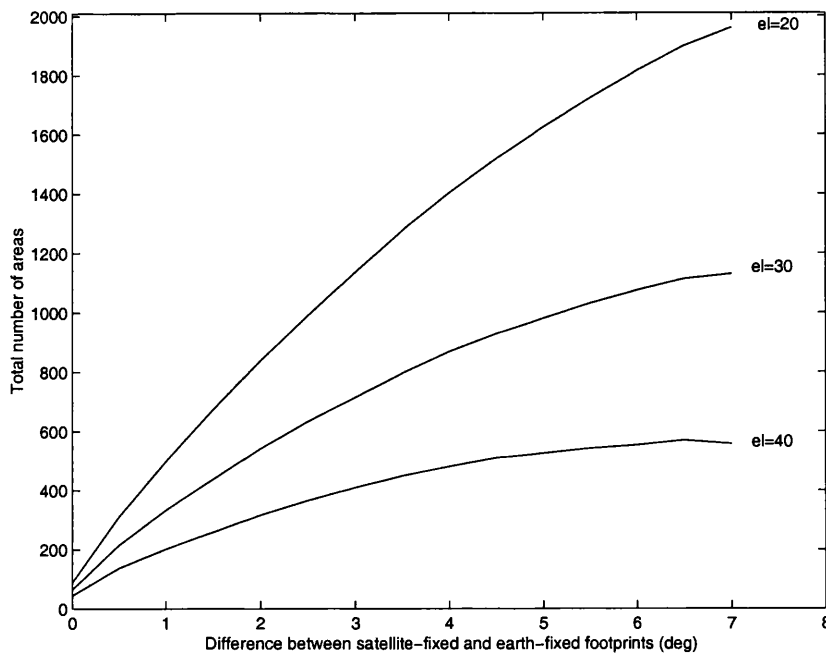
Figure 3.12: Number of areas for a MEO system (h=10000 km) versus Ψ_ρ
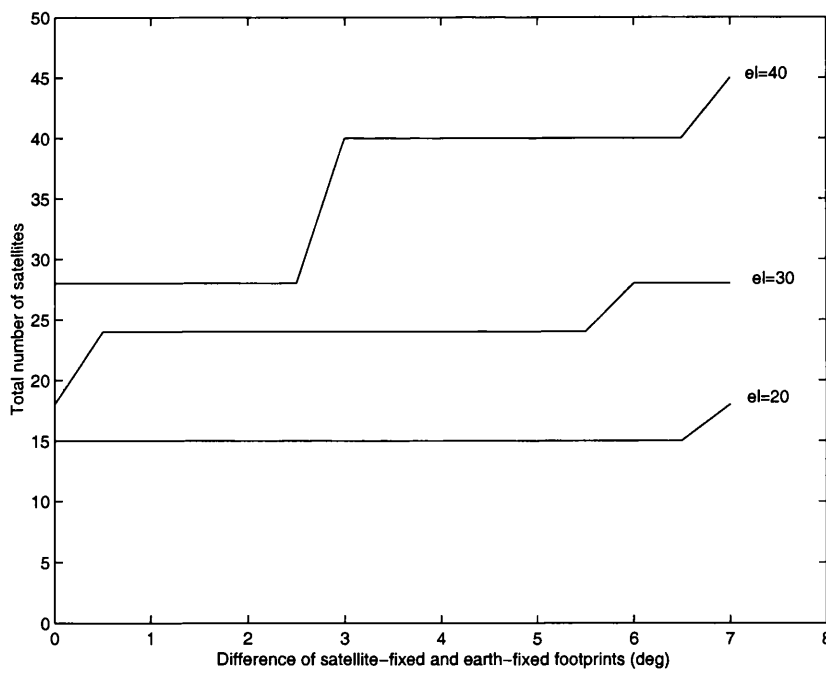
Figure 3.13: Number of satellites for a MEO system (h=10000 km) versus Ψ_ρ

In figure 3.14 the minimum distance (and candidate cell diameter) is $d_{min} = 200m$, while the average distance, for inter-ring distances below 50 km, is approximately

$d_{avg} = 21.9km$. Similarly, in figure 3.15 the minimum distance is $d_{min} = 2.7km$ and the average distance for inter-ring distances below 100km, is $d_{avg} = 31.07km$. The very small distances shown as deep valleys in these figures can be avoided if an algorithm were added to ensure that differences between consecutive ψ_{α_x} and ψ_{β_y} are kept above a minimum value. However, even if this was implemented, it would be very difficult to maximise the inter-ring distances near the intersection points. In the earth-fixed cell satellite system, if the satellite cells have a radius in the range of 1 km (as the diameter dmin was calculated to be 2.7 km) the system would be completely unpractical. Hence, it could be concluded that a better option in an 'earth-fixed cell system' scenario would be to adjust GSM in order to match the satellite channel. In that case, we are using the x_{dp} values presented in table 3.2 and we are applying these values for Iridium and ICO type of systems for different values of Ψ_{rho}. The results are presented in table 3.3 and it can be seen that the number of areas is significantly reduced[1].

3.5   Conclusions

In this chapter we have presented footprint division patterns for GSM-mobile satellite systems. These patterns were introduced in order to address the timing synchronisation problems that appear in the resource contention phase for GSM due to the large channel propagation delay introduced in a satellite environment. There

---

[1]The actual number of satellites used for Iridium is 66 and for ICO is 10 plus 2 spare. A possible reason why there is this discrepancy in table 3.3, with 77 Iridium and 15 ICO satellites in 0 $\Psi_p(degrees)$, could be that the formula defines an upper bound for the number of satellites. Note however that originally there were supposed to be seven planes and 77 satellites in total for Iridium. Hence the name Iridium, whose atomic number is 77 [36].

$\Psi_\rho(degrees)$	Iridium		ICO	
	areas	satellites	areas	satellites
0	2	77	5	15
1	2	84	9	15
2	2	84	9	15
3	2	104	8	15
4	2	112	10	15
5	3	126	11	15
6	3	135	10	15
7	3	170	11	15
8			11	18
9			12	24
10			10	24
11			10	24
12			12	24
13			12	24

Table 3.3: Application to Iridium and ICO type of constellations
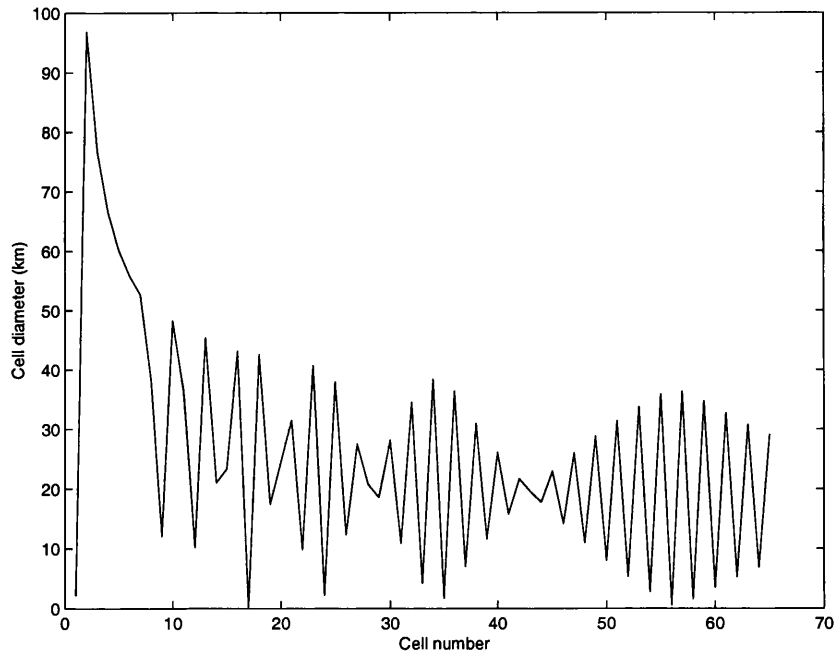
Figure 3.14: Distance between consecutive areas that are located in the footprint $\Omega\Theta$ diameter for a LEO system (h=800 km, $el = 13°$, $\Psi_\rho = 2°$) versus the number of area

exists a trade-off between the number and size of cells that the footprint is segmented into, and the hand-over frequency introduced. Alternative implementations that will increase the size and reduce the number of cells required have been suggested, while addressing the synchronisation problems during the contention phase in all cases.
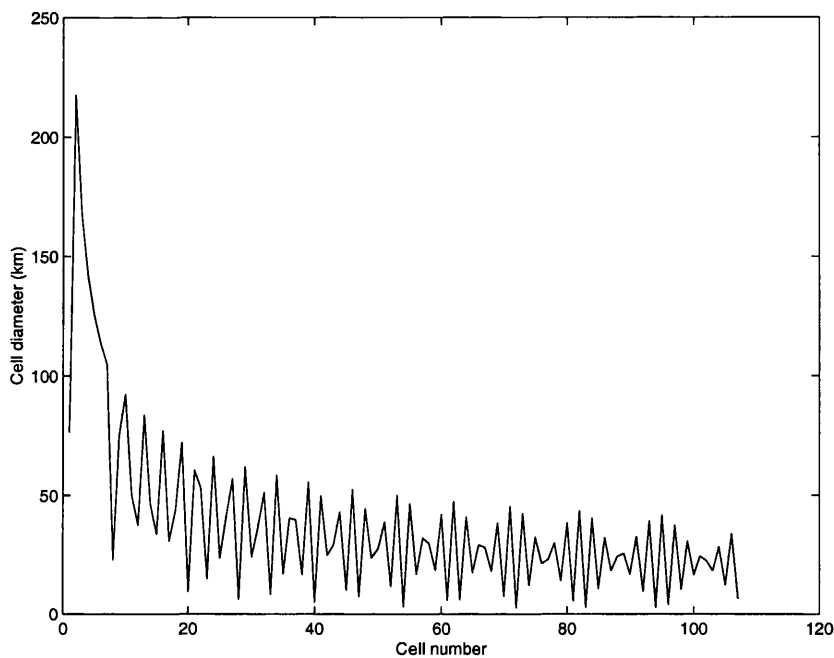
Figure 3.15: Distance between consecutive areas that are located in the footprint $\Omega\Theta$ diameter for a MEO system (h=10000 km, $el = 30°$, $\Psi_\rho = 4°$) versus the number of area

The suggested schemes have been examined for several combinations of altitude and elevation angle satellite systems. Additionally, both 'satellite-fixed cell' and 'earth-fixed cell' systems were considered. In an 'earth-fixed cell' system scenario, it was observed that the system complexity would increase immensely and alternative schemes would be most appropriate to mitigate the problem.

In the next chapter we will examine aspects of the GPRS RLC/MAC protocol over a MEO satellite channel. Different aspects of the protocol will be explored in order to test its feasibility and applicability in a satellite environment.

Chapter 4

Packet Switching Services over

Satellite Radio Systems

4.1 Introduction

In this chapter we explore the performance of the GPRS RLC/MAC protocol over a MEO satellite system. In the first section a short presentation of the tool used for the simulation, namely OPNET, is given. In the second section the simulation framework will be explained and in the third section some mathematical validation of it will be described. Finally, in the fourth section the results are presented and analysed.

4.2   Simulation tool and framework

In the next section a brief description of the simulation tool used for this research, is included. Additionally, the simulation framework is presented.

4.2.1   Optimised Network Engineering Tools

Optimised Network Engineering Tools (OPNET) is a system capable of simulating large communication networks, with detailed protocol modelling and performance analysis. OPNET features include:

- **Hierarchical, object based modelling (fig. 4.1).** At the lowest level, the behaviour of the algorithm or the protocol is encoded by a state-transition diagram with embedded code based on C language constructs. At the next level, discrete functions such as buffering, processing, transmitting and receiving data packets are performed by separate objects. These objects, called modules, are connected to form a higher level node model. At the highest level, node objects are connected by links to form a network model, which is the scope of the simulation.

- **A dynamic, event-scheduled Simulation Kernel.** OPNET simulations are based on a discrete-event modelling approach, where the progression of the model over simulation time is decomposed into individual points, called events, where changes can take place. Each event represents a need for the model to possibly effect some change in its state or to make some decision. As new events are executed, simulation time is increased monotonically. For
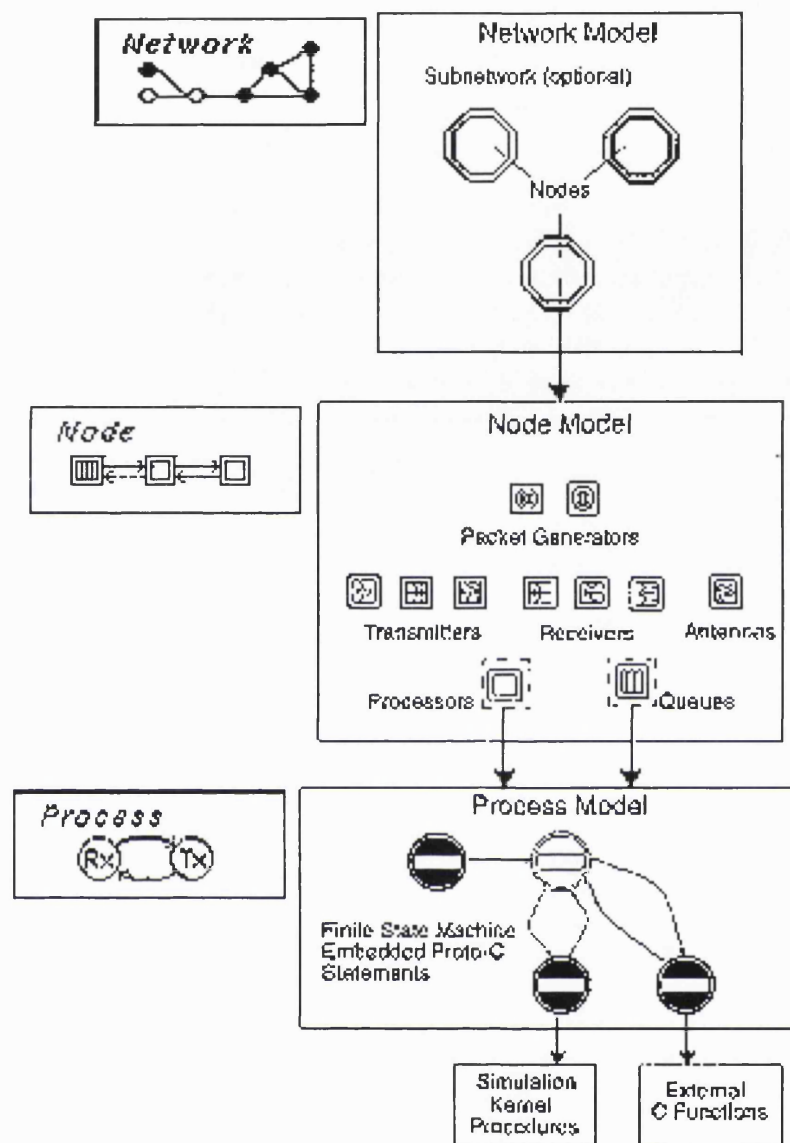
Figure 4.1: The OPNET modelling structure [37]

events that are considered to have duration equal to zero, multiple such events can occur simultaneously.

- **Graphical specifications of models**. Wherever possible, models are entered via graphical editors.

- **Integrated data analysis tool.** OPNET includes a tool for graphical presentation and processing of simulation output.

4.2.2   Simulation framework

In this section the modelling framework of the GPRS simulation is examined. In this simulation the uplink performance of the GPRS protocol over a satellite link is presented.

The following assumptions were made in order to make the modelling task tractable:

- The model does not include any handovers since we are focused on the RLC/MAC layer protocol. The mobiles are placed in the same cell (spotbeam). If system handover would be introduced, then we would expect a system performance degradation since the resource contention phase, associated with the MAC layer, might happen more than once during a packet transmission period.

- The differential path delay from any two points within the cell is smaller than or equal to the differential path delay allowed for the use of GPRS PRACH (see chapter 3), and the GSM timing advance feature.

- The satellite is 'transparent', i.e. the BTS is located on earth.

- The mobile stations are already GPRS attached (they are in the STANDBY state).

- The protocol is examined only for transmission of data generated by the mobile station.

- The multiframe timing structure is not included in the model.

- The Mobiles are all class-A, and they all have the same priority in their data queueing discipline and the resources allocation algorithm.

- There is no circuit-switched traffic in the system. All 8 timeslots are used for GPRS services.

The general configuration used in the simulation is presented in fig 4.2. The next sections present a brief overview of the system elements.

Packet Messages

The messages used for communication on the uplink are the PRACH and data packets. The PRACH packet is sent by the mobile following the slotted-aloha protocol, when there is a request for transmission resources for the generated packet traffic. All the packets, apart from the PRACH, are transmitted in RLC blocks. The PRACH is occupying only a timeslot in a TDMA frame. The PRACH can be transmitted from a mobile every time the USF downlink flag is set to FREE. In this simulation the flag is set to FREE every block period.

The messages used in the downlink are the (N)ACK, Control (Resource Assignment (RA)) and PNCH packets.
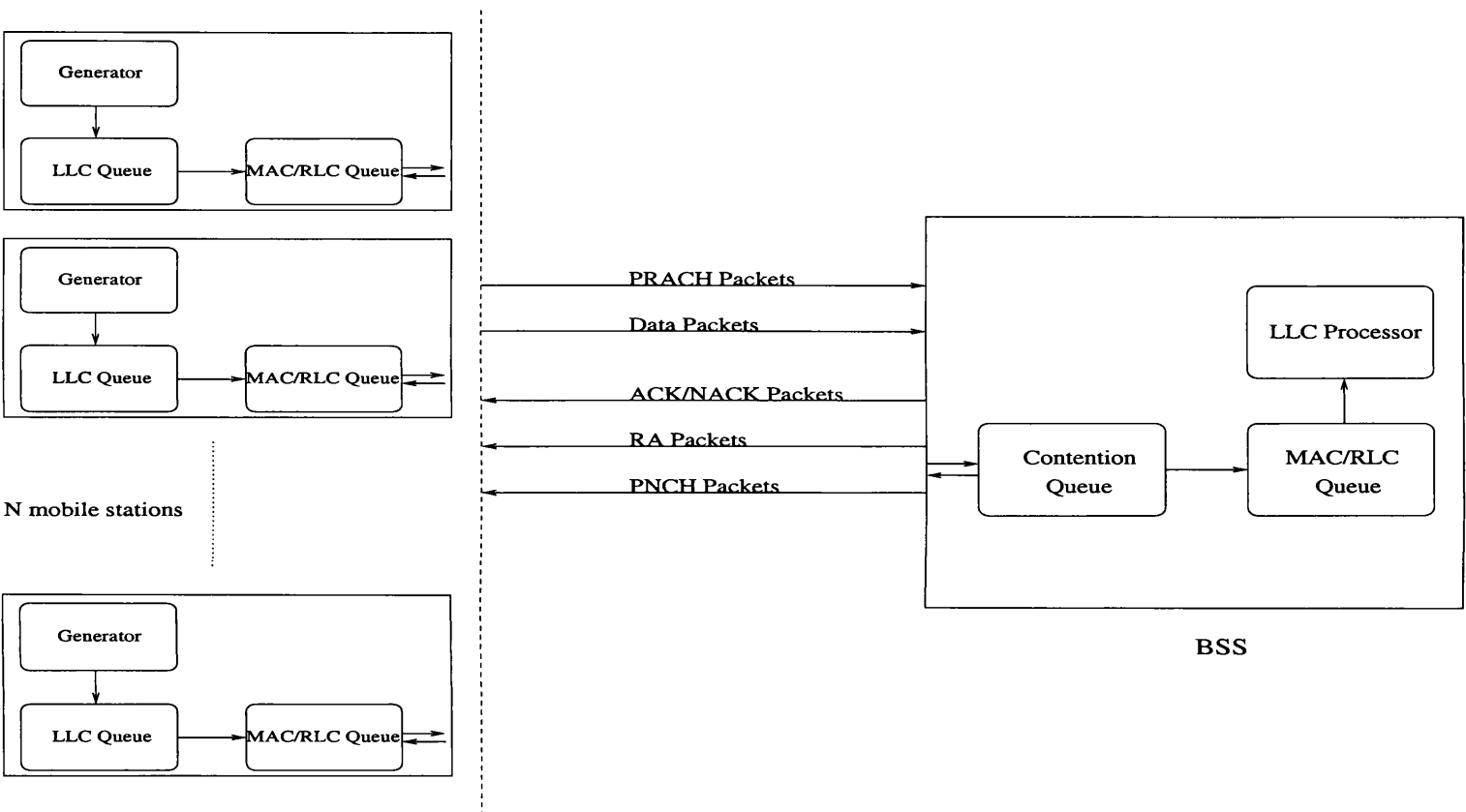
Figure 4.2: The general model

The Mobile Station Node

The mobile stations comprise of a packet generator, a Logical Link Control (LLC) and a Medium Access Control (MAC)/Radio Link Control (RLC) queue. The ideal generator produces packets with a Poisson arrival process and three different distributions for packet length [41].

- The first one is a Cauchy (0.8, 1) distribution (FUNET model), based on statistics from email usage, with maximum packet size of 10 kbytes:

$$f(x) = \frac{1}{\pi(1 + (x - 0.8)^2)} \tag{4.1}$$

- The second one is based on statistics on an assessment of railway application requirements (Railway model), has a maximum packet length of 1000 bytes and a probability density function of:

$$f(x) = \frac{1}{170}exp(-\frac{x}{170}) \tag{4.2}$$

- The third one is based on statistics from a fleet management application (Mobitex model):

$$Uplink : 30 \pm 15bytes, \quad Downlink : 115 \pm 57bytes \tag{4.3}$$

The packets that are generated are forwarded to the LLC queue, where they are fragmented into LLC frames. The LLC frames are forwarded to the MAC/RLC

queue where they are broken down to RLC blocks, in order to be sent over the GPRS channels. The functions of forming and sending the PRACH channel, as well as buffering the unacknowledged blocks, are also included in the RLC/MAC queue process.

Two different timers are used in order to sustain the required quality of service, and to avoid congestion. The T_o timer is set when a PRACH is sent and a PQN or an RA response packet is expected. If neither of these channels is received within the To time, the mobile station assumes that a collision has occurred and re-sends the PRACH after a calculated backoff time [40]. The T_{max} timer is the maximum time that a mobile can contend for radio resources. When it expires, the frame is considered to be dropped.

LLC frames are broken down into RLC blocks, after the coding of data. Each one of the RLC blocks is assigned an SNR which is a number computed by an external function, so that the radio environment of the satellite channel would be simulated.

BSS Node

In the Base Station Subsystem (BSS) there are two queues, the contention and the MAC/RLC queue. The contention queue is checked every timeslot that a PRACH is expected (USF==FREE) from the downlink. If there are more than one PRACH arrivals, a collision event is assumed and the PRACH packets are destroyed. If there is only one PRACH arrival then the packet is forwarded to the RLC/MAC queue.

The algorithm used for the distribution of the radio resources (USF values for

each GPRS timeslot) is the following one. There are two main matrices used for the application of the algorithm. In the first matrix, all USFs per timeslot are free (maximum number of USF per timeslot is 8). Every time a mobile is contending for an x number of timeslots, the x timeslots that have the largest available number of USFs are chosen. If there are not x available timeslots, then *(x-1)* or *(x-2)* (etc...) number of timeslots is assigned. The second matrix stores all the assigned USFs per timeslot. This matrix is checked every block period to send all 8 (maximum) USF values in the downlink. A round robin scheme is used to serve all USF values per timeslot.

In the RLC/MAC module, there is a central queue for the PRACH messages waiting to be served and a buffer for the RLC blocks arriving from the mobile. If there is an event of RLC data arrival, the (correct) block is buffered until the complete LLC frame is received. In the case of some of the blocks being in error, a NACK packet including the error bitmap of the blocks is sent at regular intervals [40]. When the complete correct frame is received, it is then forwarded to the LLC layer, where the statistics for the evaluation of the protocol are updated.

4.3   Analytical models

The analytical models which will be presented in this section were developed mainly for the validation of the simulation. Since GPRS is very complicated, for mathematical analysis purposes, the models are applied only to simplified cases of the protocol.
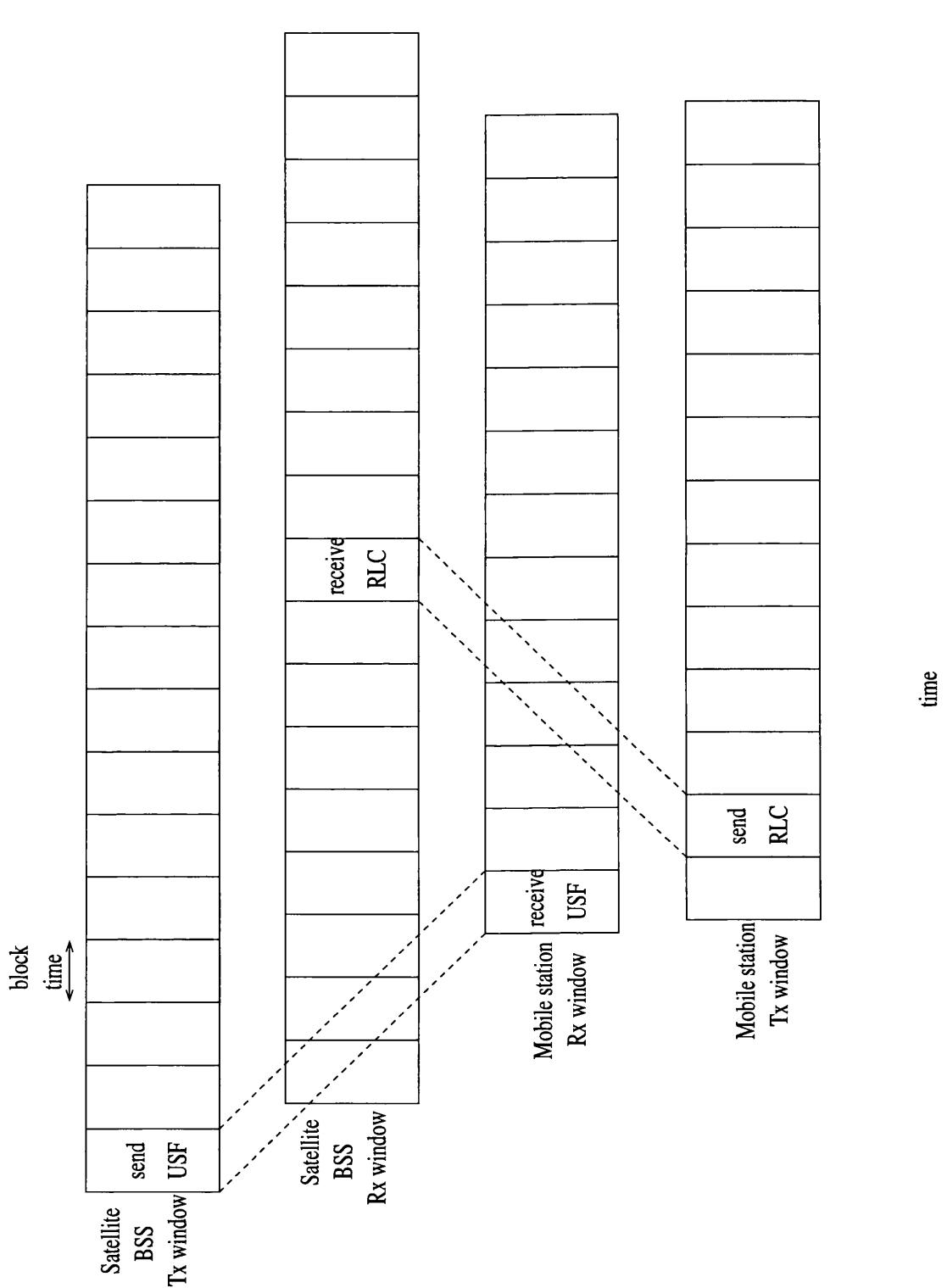
Figure 4.3: Transmission and reception windows structure in the MS and the BSS

Initially, we should first present a brief description of how the reception and the transmission windows are structured in the MS and the BSS. Since both the MS and the BSS are located on the ground, a single satellite hop time delay is required between the transmission and the reception of the same block. In figure 4.3 a simple example is illustrated. Specifically, the propagation delay for a single hop is a little bit more than 3 blocks time. Additionally, the reception and transmission windows in the MS are three timeslots apart (section 3.2). In this particular example, the BSS sends a message containing the appropriate USF flag to the MS and the MS replies sending an RLC block.

4.3.1 First model

The first model is the simplest one. We assume the following:

- There is only one mobile source of data.

- The data generated has a Poisson arrival rate and deterministic packet length.

- The packet length can fit into one block of transmitted data.

- The interarrival rate is much bigger than a block period.

- There is no noise in the system.

- There is only one timeslot (number 0) per frame, which is used for the transmission of both the control and data blocks.

- The satellite is transparent and a MEO system is assumed.

- There is no difference in the propagation delay from the satellite to the mobile on different ends of the satellite's beam coverage and from the satellite to the BSS.

The main concept is, therefore, that we have a source generating data with Poisson arrival, a fixed service time and one server (timeslot). Thus, the system could be analysed using M/G/1 queue equations. We use the queueing approach because even though the load is very low (to ensure that there are no two packet arrivals within the same block time), the service time, due to the propagation delay, will be much greater than the interarrival times.

The basic diagram used to explain the timing structure of the system is shown in figure 4.4. The USF flag on the downlink channel is USF==FREE, until a PRACH is sent and then the timeslot is only used for transmission in the uplink until the service of the block is over. This figure follows the general principles of figure 4.3 and it includes all the BBS and MS Rx and Tx windows. Actually, the MS Tx and Rx windows are combined in one set of windows as the time difference between them is minimal (3 timeslots). The height of the satellite is at maximum 9500 km. This height gives a maximum round trip delay of 3.12 blocks (a block period is considered to be 20 ms if we take into account compensation for the idle frames). This height was chosen in order to make the handling of the propagation delay more flexible. The handling of a control procedure (RA or ACK preparation) by the BSS or a data transmission preparation by the MS adds a delay of half a block. The transmission of the control channels equals one block period.
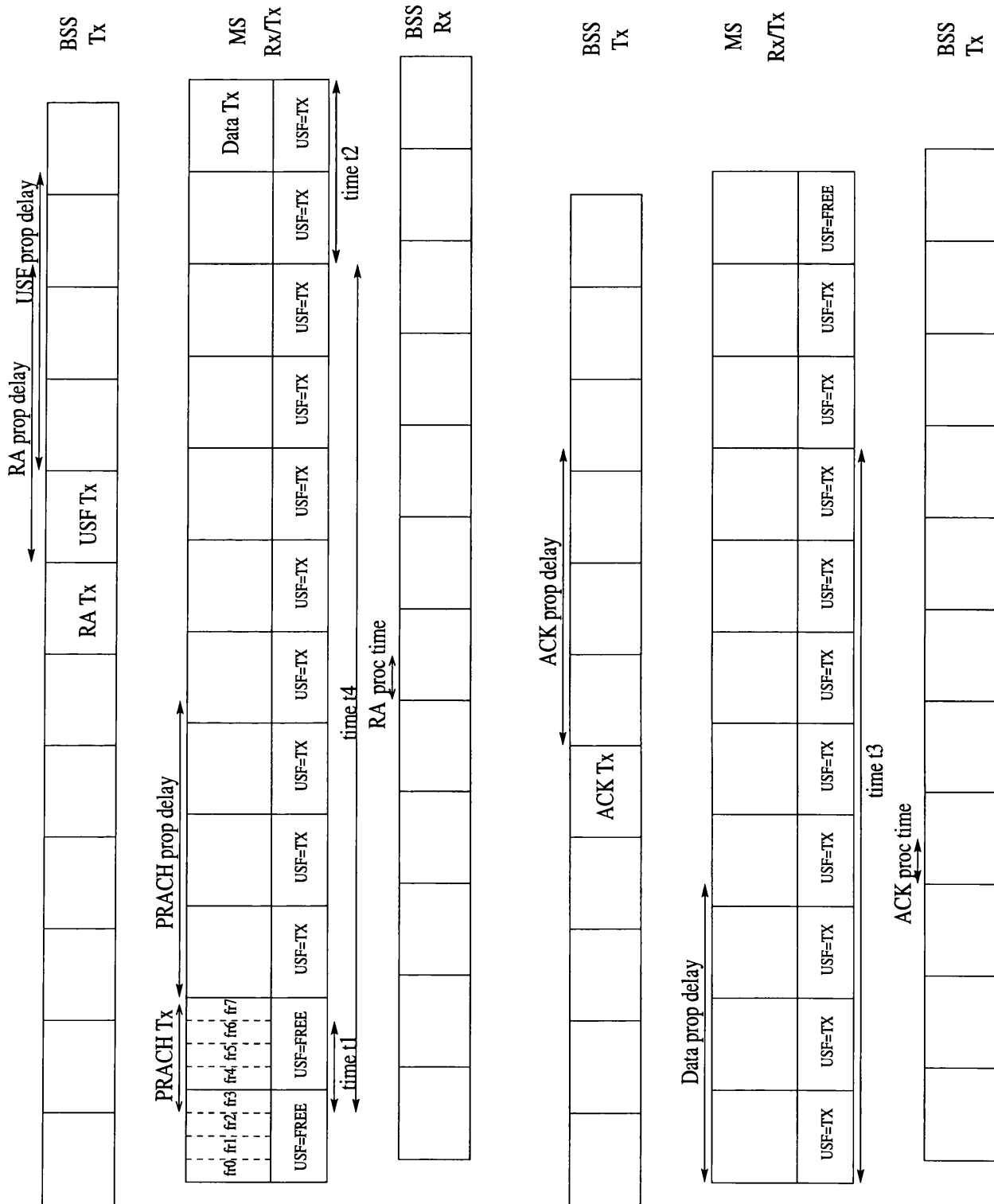
Figure 4.4: The timing structure of the first model of analysis

The Pollaczek-Khintchine formula gives the average number L of customers (packets) in the system for a stable M/G/1 queue:

$$L = \frac{\lambda^2 E[S^2]}{2(1 - \lambda E[S])} + \lambda E[S], \tag{4.4}$$

where λ is the packet arrival rate , $E[S]$ the average service time and $E[S^2] = Var(S) + (E[S])^2$.

By combining equation 6.1 with Little's formula:

$$L = \lambda W, \tag{4.5}$$

we can obtain the average time W that a customer spends in the system:

$$W = \frac{\lambda E[S^2]}{2(1 - \lambda E[S])} + E[S] \tag{4.6}$$

There are two cases that have to be examined in this model. The first one is the worst one where the packet is already queued and it starts being served at the beginning of the 2nd block of the MS window in figure 4.4. In this case the service time is fixed,

$Var(S)$ =0 and

$$E[S]_{max} = 380ms.$$

The second one is the case where the queueing time is zero because the system is empty and the packet could arrive anywhere during the t_1 time period. Therefore, the time that the packet is waiting to be served could be 4,3,2 or 1 frames. Thus, the mean of the waiting time is 2.5 frames, the maximum 4 frames and the minimum 1 frame[1]. Therefore, the average service time is

$$E[S]_{min} \quad = \quad 372.5ms$$

and the variance is calculated as

$$VAR[S]_{min} \quad = \quad 7.50ms.$$

The table 4.1 shows the maximum and the minimum theoretical W achieved using the formula 4.6, as well as the simulation results for different interarrival rates λ. As can be seen in the table there is a big difference between the first and the second value of W. This happens because the first point is very close to the point where the system begins to be unstable ($arrival \quad rate \geq service \quad rate$)

$1/\lambda$ (sec)	0.4	0.5	0.6	0.8	1.0	1.2
W analytical min (sec)	3.03	0.946	0.694	0.544	0.489	0.462
W analytical max (sec)	3.99	0.982	0.708	0.552	0.496	0.468
W simulation (sec)	3.70	0.9590	0.693	0.544	0.486	0.467

Table 4.1: Analytical and simulation W without noise in the system

---

[1]The time between the arrival of the packet until the beginning of the next block is considered to be negligible compared to the other packet delays and therefore, is not included in the calculations

4.3.2 Second model

In this case the basic model is the same as the first one. The only difference is that we assume a block error rate q only for the data block. In this case the periods t_2 (we also include the time period during which the MS has to wait for the appropriate USF value to arrive) and t_3 in figure 4.4 will be repeated $\overline{X}$ times ($\overline{X}-1$ number of NACKs and 1 ACK), where $\overline{X}$ is the mean number of retransmissions that will be needed for each block. If X is the number of trials required until the first success, then X is a geometric random variable with parameter p=1-q and mean and variance:

$$\overline{X} = \frac{1}{p}, \quad VAR[X] = \frac{q}{p^2}.$$

The new

$$E[S] = \overline{t}_4 + \overline{X}t_5$$

and

$$VAR[S] = VAR[t_4] + VAR[X](t_2 + t_3)^2,$$

because t_4 and X ($t_2 + t_3$) are independent variables.

Using a $q = 0.0002$ we have:

$$E[S]_{max} \quad = \quad 380.04ms$$

$$VAR[S]_{max} \quad = \quad 8ms.$$

$$E[S]_{min} \quad = \quad 372.54ms$$

$$VAR[S]_{min} \quad = \quad 15.50ms.$$

The table 4.2 shows the maximum and the minimum theoretical W achieved using the formula 4.6, as well as the simulation results for different interarrival rates λ.

$1/\lambda$ (sec)	0.4	0.5	0.6	0.8	1.0	1.2
W analytical min (sec)	3.19	0.978	0.717	0.553	0.496	0.458
W analytical max (sec)	4.19	1.015	0.726	0.564	0.503	0.473
W simulation (sec)	4.02	0.989	0.715	0.551	0.493	0.470

Table 4.2: Analytical and simulation W with noise in the system

4.3.3 Third model

In this model we assume that the LLC frame consists of n RLC blocks. Then the actual transmission time is equal to

$$T_{tx} = (n\overline{X})blocks \times 20\frac{ms}{blocks}$$

and the total sum of the propagation times and the transmission of the NACKs and ACK is repeated $E[N(n)]+1$ times, where $E[N(n)]$ is the average number of the NACKs sent for an LLC frame initially consisting of n RLC blocks. Therefore if a block time is t_b:

$$E[S] = \bar{t}_4 + T_{tx} + (E[N(n)] + 1)(t_2 + t_3 - t_b)$$

and

$$VAR[S] = VAR[t_4] + VAR[T_{tx}] + VAR[N(n) + 1](t_2 + t_3 - t_b)^2,$$

E[N(n)] is equal to [10]:

$$E[N(n)] = \sum_{m=1}^{\infty} mP\Big(N(n) = m\Big) \tag{4.7}$$

and

$$P\Big(N(n) = m\Big) = \sum_{k=1}^{n} \binom{n}{k} q^k p^{n-k} P\Big(N(k) = (m-1)\Big) \tag{4.8}$$

Using a $q = 0.0002$ and $n=10$ we have:

$$E[S]_{max} \quad = \quad 543.362ms$$

$$VAR[S]_{max} \quad = \quad 16.8ms.$$

$$E[S]_{min} \quad = \quad 535.862ms$$

$$VAR[S]_{min} \quad = \quad 21.48ms.$$

Table 4.3 shows the maximum and the minimum theoretical W achieved using

the formula 4.6, as well as the simulation results for different interarrival rates λ.

$1/\lambda$ (sec)	0.6	0.7	0.8	1.0	1.2	1.5
W analytical min (sec)	2.546	1.392	1.097	0.876	0.757	0.693
W analytical max (sec)	2.839	1.437	1.113	0.898	0.778	0.712
W simulation (sec)	2.723	1.398	1.102	0.882	0.766	0.703

Table 4.3: Analytical and simulation W with noise in the system for a packet consisting of 10 blocks

4.3.4   Fourth model

In the case where the length in bytes of an LLC frame is a continuous random variable x with probability density function *f(x)*, then the probability density function of the number of RLC blocks n, if we assume that each block contains 20 information bytes, is:

$$P(n) = \int_{20(n-1)}^{20n} f(x)dx \qquad (4.9)$$

We also know that the probability density function of y is the geometric distribution, where y is the the number of transmissions needed for one RLC block:

$$P(y) = pq^{y-1} \qquad (4.10)$$

The total number of RLC block transmissions is a compound random variable k and if $\overline{k}$ is its average value, then the total data transmission time is equal to

$$T_{tx} = \overline{k} blocks \times 20\frac{ms}{blocks}$$

By applying the theory of compound randomness, one can conclude that:

$$\overline{k} = \overline{n} \times \overline{y}$$

The total sum of the propagation times and the transmission of the NACKs and ACK is repeated $E[N(n)]$+1 times, where $E[N(n)]$ is the average number of the NACKs sent for an LLC frame initially consisting of n RLC blocks. $E[N(n)]$ is equal to

$$E[N(n)] = \sum_{n=1}^{\infty} N(n) \int_{160(n-1)}^{160n} f(x)dx \qquad (4.11)$$

By using 4.2 as the distribution for x and taking into account that the length has a maximum of 1000 bytes, probability density function 4.9 becomes:

$$P(n) = \frac{1}{(1 - exp(-\frac{1000}{170}))} \int_{20(n-1)}^{20n} \frac{1}{170}exp(-\frac{x}{170})dx \qquad (4.12)$$

which after the integration becomes:

$$P(n) = \frac{exp(20/170) - 1}{(1 - exp(-1000/170))} \times exp(20n/170) \qquad (4.13)$$

and equation 4.11 becomes:

$$E[N(n)] = \frac{exp(20/170) - 1}{(1 - exp(-1000/170))} \sum_{n=1}^{\infty} N(n)exp(20n/170) \qquad (4.14)$$

By substituting all the above to the following equation:

$$E[S] = \bar{t}_4 + T_{tx} + (E[N(n)] + 1)(t_2 + t_3 - t_b),$$

we get the following values for $E[S]_{max}$ and $E[S]_{min}$:

$$E[S]_{max} = 516.65ms$$

$$E[S]_{min} = 508.62ms$$

The variance is calculated as follows:

$$VAR[S] = VAR[t_4] + VAR[k] \times t_b + VAR[N(n) + 1](t_2 + t_3 - t_b)^2,$$

The $VAR[N(n)]$ can easily be calculated since we know the distribution given in 4.12. However, for calculating the $VAR[k]$ a different process must be followed:

Let's assume that $H_k(z)$ is the probability generating function of k, $F_n(z)$ is the probability generating function of n and $G_y(z)$ is the probability generating function of y. Then:

$$H(z) = F(G(z)) \Rightarrow H'(z) = F'(G(z))G'(z) \Rightarrow H''(z) = F''(G(z))G'(z)^2 + F'(G(z))G''(z)$$

$$\Rightarrow H''(1) = F''(1)G'(1)^2 + F'(1)G''(1) \Rightarrow E[k^2] - E[k] = F''(1)E[y]^2 + E[n]G''(1)$$

$$\Rightarrow E[k^2] = E[n]E[y] + F''(1)E[y]^2 + E[n]G''(1)$$

However:

$$VAR[k] = E[k^2] - E[k]^2 = E[n]E[y] + F''(1)E[y]^2 + E[n]G''(1) - (E[n]E[y])^2 \quad (4.15)$$

and

$$F_n(z) = \sum_{n=0}^{\infty} P(n)z^n = \frac{exp(20/170) - 1}{(1 - exp(-1000/170))} \times \sum_{n=0}^{\infty} z^n exp(20n/170) \quad (4.16)$$

which becomes:

$$F_n(z) = \frac{exp(20/170) - 1}{(1 - exp(-1000/170))} \times \frac{1}{1 - (z exp(20/170))} \quad (4.17)$$

for $z/exp(20/170) < 1$

Similarly,

$$G_y(z) = \sum_{y=0}^{\infty} P(y)z^y = p/q \sum_{y=0}^{\infty} q^y z^y, \qquad (4.18)$$

which becomes:

$$G_y(z) = (p/q) \times (1/(1 - qz)) \qquad (4.19)$$

for $qz < 1$

By differentiating equations 4.18 and 4.16 and substituting in 4.15, and after the calculation of the total variance we have the following results:

$$VAR[S]_{max} = 37.3ms$$

$$VAR[S]_{min} = 39.1ms$$

$1/\lambda$ (sec)	0.6	0.7	0.8	1.0	1.2	1.5
W analytical min (sec)	2.033	1.251	0.997	0.792	0.728	0.646
W analytical max (sec)	2.225	1.292	1.117	0.821	0.722	0.708
W simulation (sec)	2.1199	1.268	1.013	0.810	0.731	0.654

Table 4.4: Analytical and simulation W with noise in the system for a packet of size following the exponential distribution

The table 4.4 shows the maximum and the minimum theoretical W achieved using the formula 4.6, as well as the simulation results for different interarrival rates λ.

4.4 Results

We have tested the general performance (section 4.4.1) of GPRS over the satellite link for two of the three packet distributions presented in section 4.2.2. The statistics used for the evaluation were end-to-end delay, blocking probability and throughput. The LLC end-to-end delay is the time elapsed between the instant at which the LLC packet is forwarded to the RLC/MAC layer, and the instant at which the final ACK message for this specific packet is received. The blocking in the specific model is only due to the random access attempts that failed because the T_{max} expired and hence, the blocking probability is the fraction of these failed attempts over the total attempts. The throughput is defined as the rate of error-free data that reaches the destination. Additionally, in section 4.4.2 the GPRS ARQ scheme is tested and alternative transmit window sizes, in combination with different intervals between consecutive (N)ACK transmissions, are used for the 4.1 distribution presented in section 4.2.2. Furthermore, the GPRS performance with multislot operation is compared to the one with single-slot operation in section 4.4.3. Finally, in section 4.4.4 an enhancement in system performance is analysed after the introduction of new schemes.

4.4.1 General performance

We have tested the Mobitex and the Railway models for the CS-1 and CS-3 coding schemes and different Block Error Rates (BLERs). The model was working on a single-slot operation, and the values for the timers were $T_o = 200ms$, $T_{max} = 500ms$.
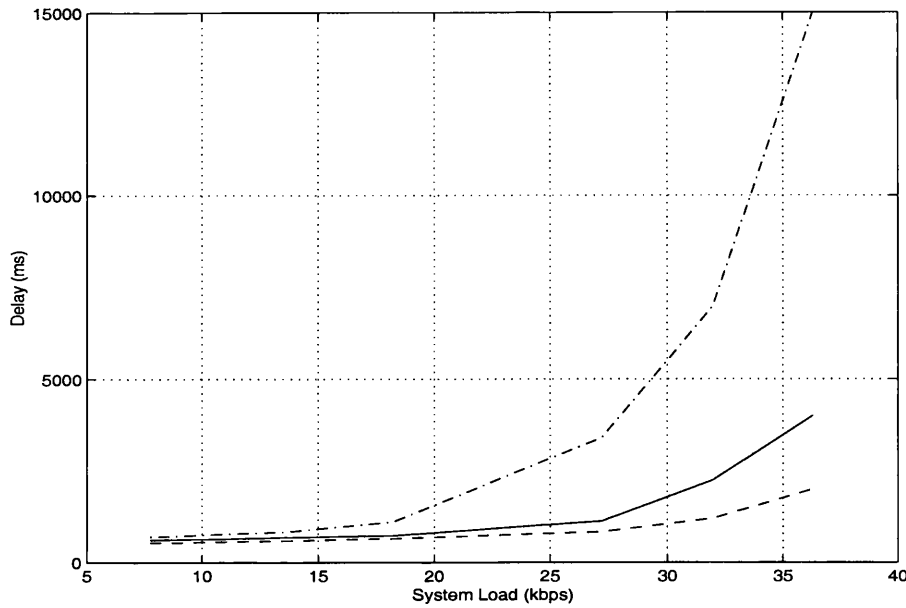
Figure 4.5: LLC end-to-end frame delay for the Railway model for: *i.*CS-1,BLER=0.0002 (solid line), *ii.*CS-1,BLER=0.08 (-.), *iii.*CS-3,BLER=0.02 (–)
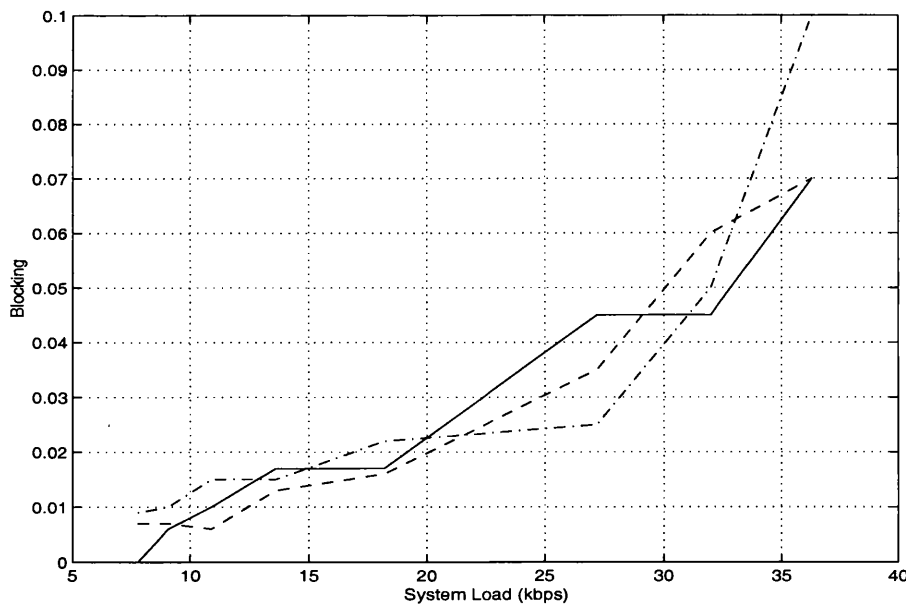
Figure 4.6: LLC frame blocking for the Railway model for: *i.*CS-1,BLER=0.0002 (solid line), *ii.*CS-1,BLER=0.08 (-.), *iii.*CS-3,BLER=0.02 (–)

The ARQ window size was set to 64. The statistics used were delay, blocking probability and throughput. Specifically, figures 4.5, 4.6 and 4.7 illustrate the above statistics for the Railway model, while figures 4.8, 4.9 and 4.10 show the corre-
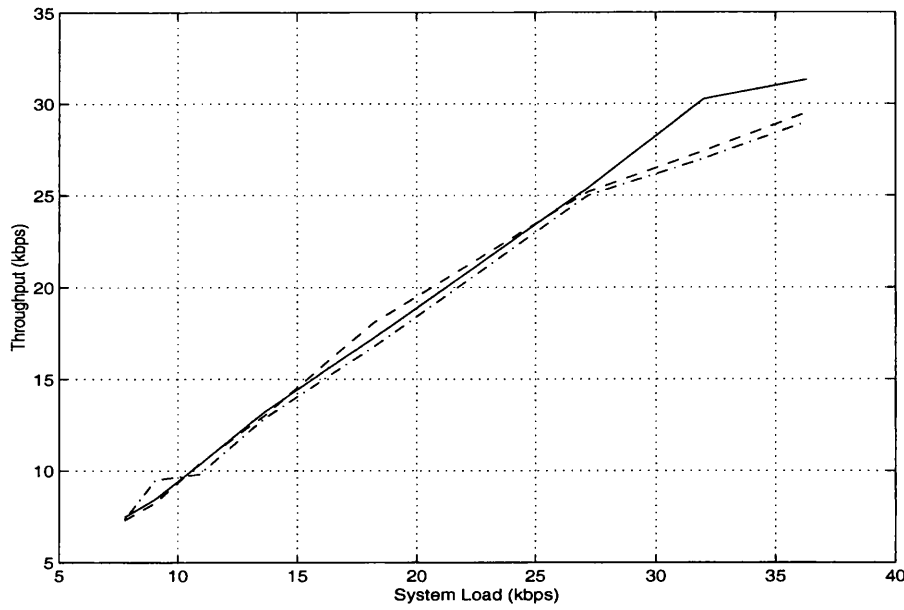
Figure 4.7: LLC frame throughput for the Railway model for: *i.*CS-1,BLER=0.0002 (solid line), *ii.*CS-1,BLER=0.08 (-.), *iii.*CS-3,BLER=0.02 (–)
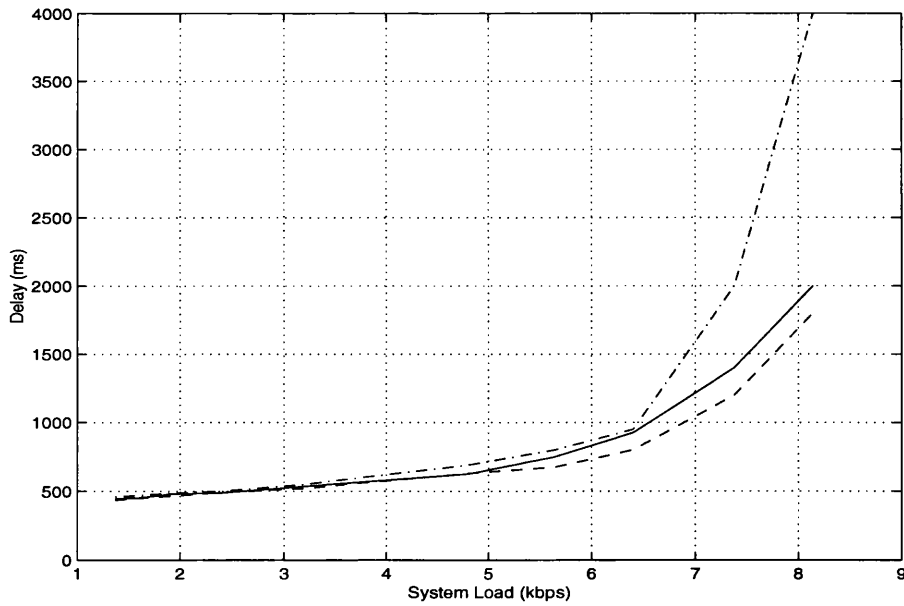
Figure 4.8: LLC frame delay for the Mobitex model for: *i.*CS-1,BLER=0.0002 (solid line), *ii.*CS-1,BLER=0.08 (-.), *iii.*CS-3,BLER=0.02 (–)

sponding results for the Mobitex model. Additionally, in section 4.4.2 the protocol is tested for the 4.1 distribution presented in section 4.2.2. The first observation from the results is that the traffic model greatly affects the efficiency of the protocol
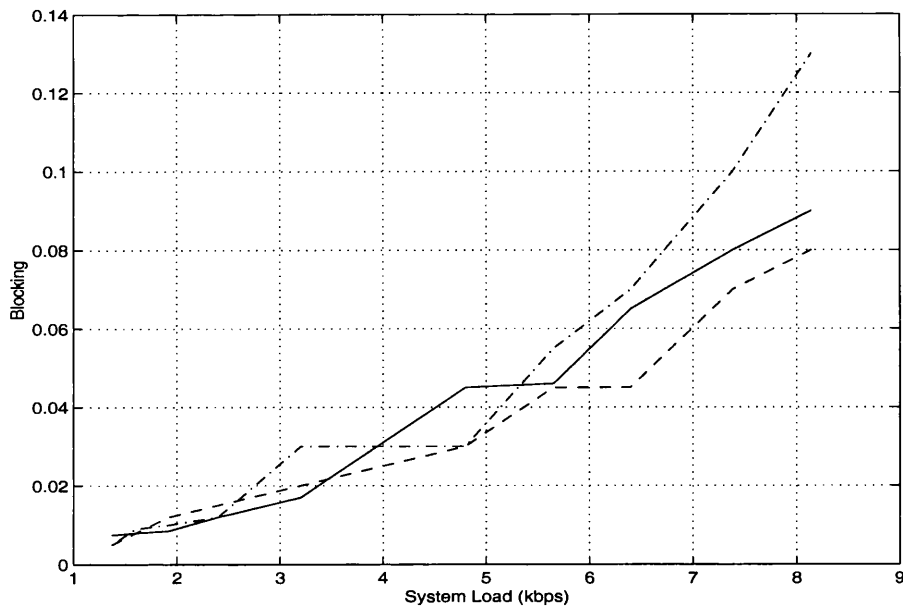
Figure 4.9: LLC frame blocking for the Mobitex model for: *i.*CS-1,BLER=0.0002 (solid line), *ii.*CS-1,BLER=0.08 (-.), *iii.*CS-3,BLER=0.02 (-)
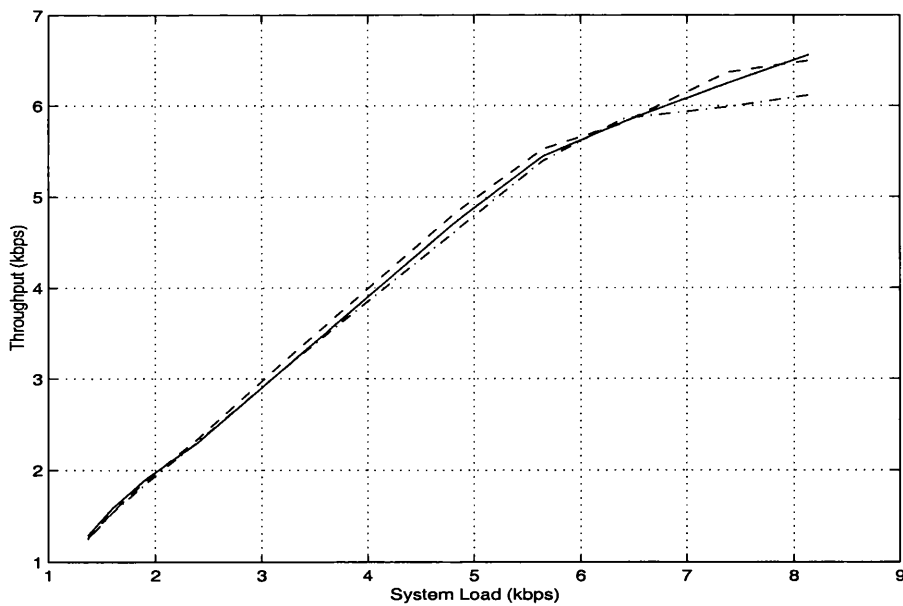
Figure 4.10: LLC frame throughput for the Mobitex model for: *i.*CS-1,BLER=0.0002 (solid line), *ii.*CS-1,BLER=0.08 (.-), *iii.*CS-3,BLER=0.02 (-)

just as in GPRS terrestrial systems ([41], [42]). Specifically, the fact that the three different packet models are of different LLC frame sizes, results in the saturation state being reached at a lower system load for the smaller packets, while for the

bigger packets saturation is reached at a much higher system loading (figures 4.7, 4.10 and 4.12). Therefore, the Mobitex model starts to reach saturation at approximately 8 kbps system load, the Railway model at 35 kbps and the FUNET model at 50 kbps. These critical points of system load are obviously lower than those of a terrestrial system ([42]). However, the general behaviour seems to be the same.

We also observe that a higher noise environment has more impact on the Railway model (figure 4.5), while in the Mobitex model the noise really has an impact on the LLC frame transfer delay as the saturation point is approached (figure 4.8). Similarly, the second coding scheme introduces an improvement in system delay in the Railway model and does not have a big effect in the Mobitex model apart from system loads close to saturation. This can be explained by the fact that in the Mobitex model the packets are very small and the real effect in the delay comes from the contention stage. The blocking probability remains at tolerable levels when delay also remains at tolerable levels.

4.4.2   Impact of the ARQ scheme

For the FUNET model we do not expect the contention phase to be the major limiting factor, mainly due to the large packet sizes. Therefore, we focus our attention on the transmission time of the packet. Specifically, we study how the protocol behaves under different combinations of BLERs and ARQ window sizes. As already mentioned in the 'General Background' chapter, the ARQ window size is 64 at the moment, because of the 64 bitmap used in the ACK/NACK GPRS packets. We will see here that different window sizes can improve the performance of GPRS under
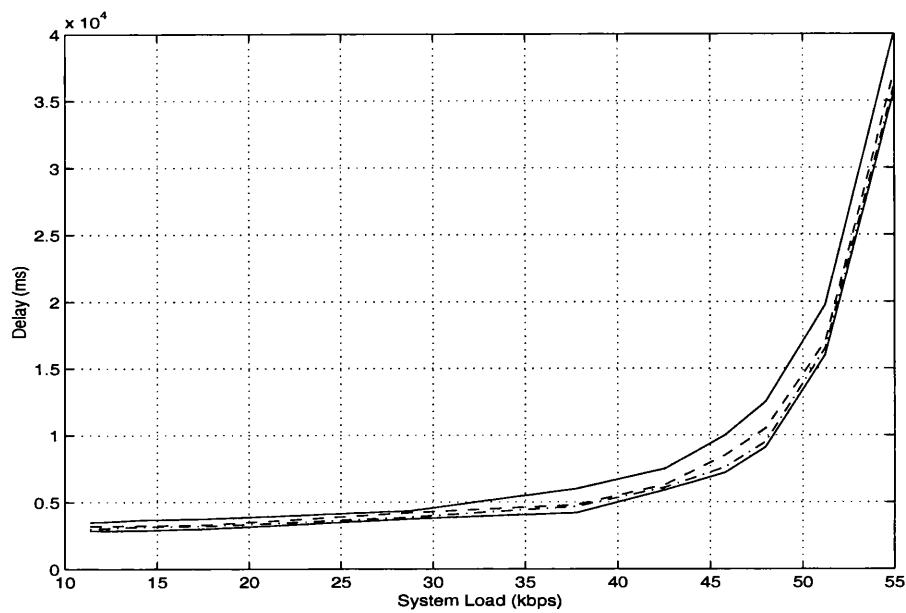
Figure 4.11: Delay for the FUNET model for BLER=0.0002: *i.* K=30, WIN=64 (solid line), *ii.* K=64, WIN=64 (- line), *iii.* K=60, WIN=80 (-. line), *iv.* K=64, WIN=110 (solid-dotted line)
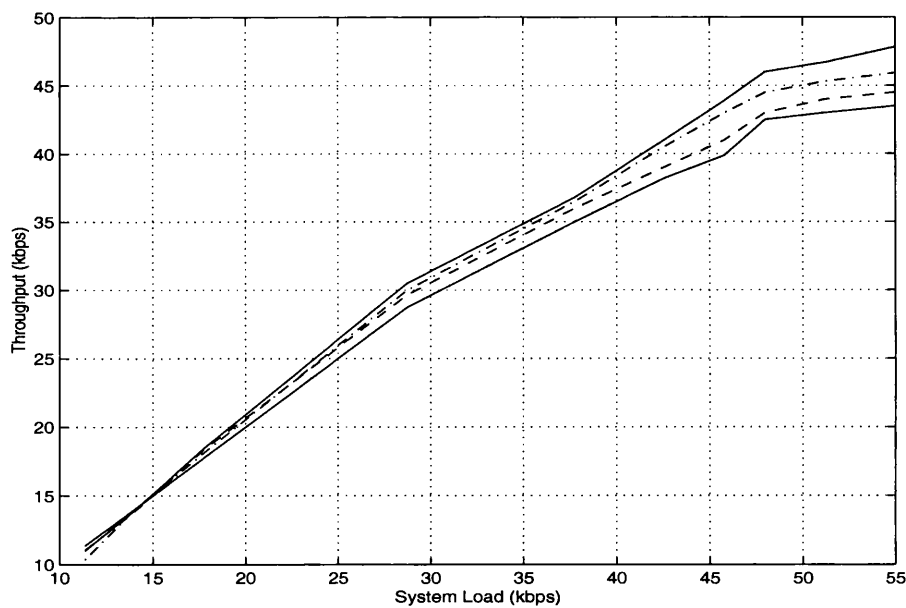
Figure 4.12: Throughput for the FUNET model for BLER=0.0002: *i.* K=30, WIN=64 (solid line), *ii.* K=64, WIN=64 (- line), *iii.* K=60, WIN=80 (-. line), *iv.* K=64, WIN=110 (solid-dotted line)

specific conditions.

Figures 4.11 and 4.12 show the LLC frame delay and system throughput for
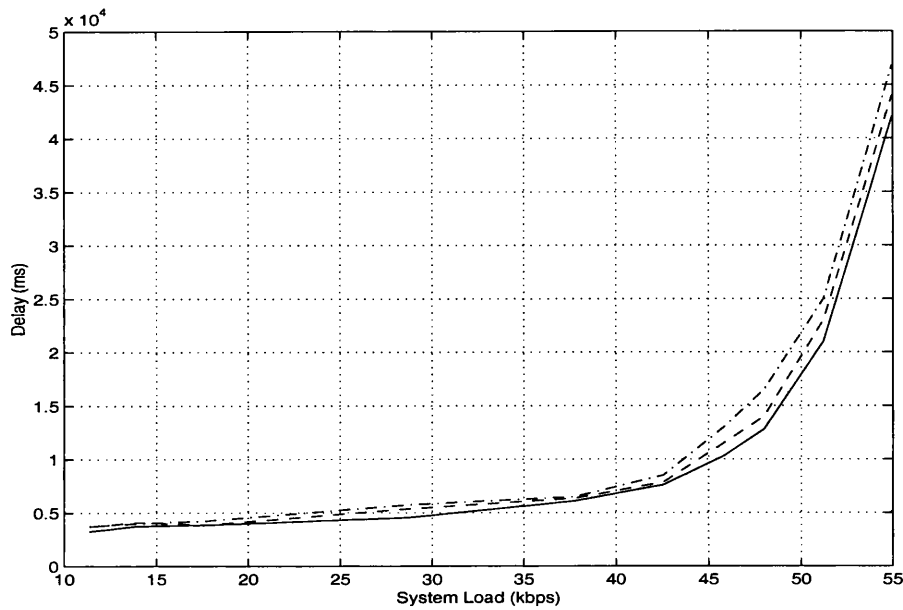
Figure 4.13: Delay for the FUNET model for BLER=0.02: *i.* K=30, WIN=64 (solid line), *ii.* K=64, WIN=64 (– line), *iii.* K=110, WIN=133 (-. line)
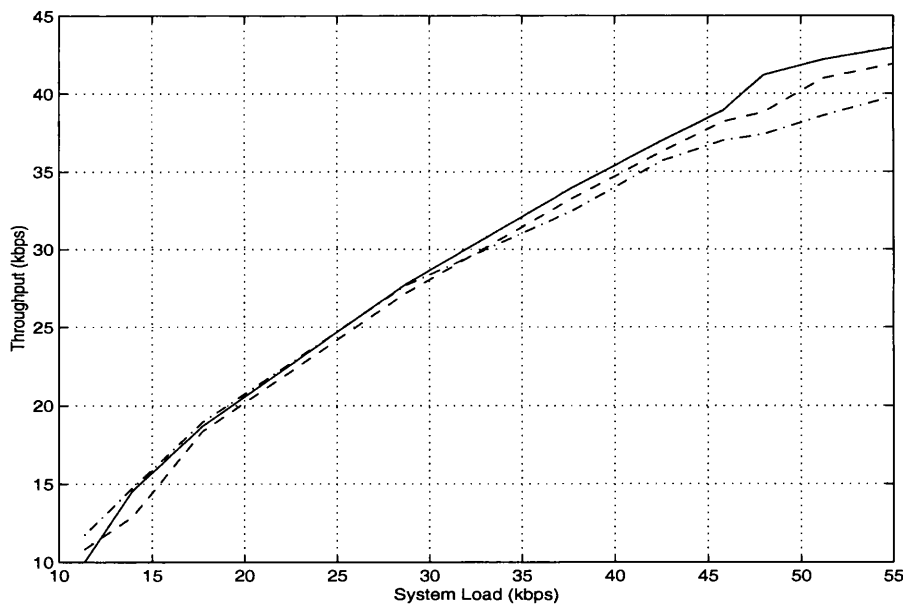
Figure 4.14: Throughput for the FUNET model for BLER=0.02: *i.* K=30, WIN=64 (solid line), *ii.* K=64, WIN=64 (– line), *iii.* K=110, WIN=133 (-. line)

BLER=0.0002, while figures 4.13 and 4.14 illustrate the LLC frame delay and system throughput for BLER=0.02. The reason that blocking is not illustrated when examining the FUNET model is that the blocking is caused by frame dropping due

to the expiry of the timer in the contention period. However, here the saturation state is reached because of the high frame transmission times and not because of the contention attempts. Therefore, the blocking usually stays at low levels, and ranges from 0 to 0.01. In the captions of the figures mentioned above, WIN is the ARQ window size and K is a parameter which defines how often (i.e. per number of RLC blocks) the mobile is polling the BSS for an ACK/NACK packet. In the lower noise environment we observe that the system behaves more efficiently when larger window sizes are used, while in the higher noise environment more frequent ACK/NACK packets have to be sent. The reason is that the mobile has to wait for an ACK for a minimum of 200ms after the window size is filled and cannot send any more frames. This is not desirable when the noise is low and not many retransmissions are required.

If the improvements that will take place by introducing a variable ARQ window size, justify the implementation of such a scheme, then further standardisation work is required. This would have an impact on the software of both the terminal and the BSS as well as on the structure of the ARQ (N)ACK channel. What would be required is a software that adjusts the window size and frequency of polling according to detected channel conditions, signal levels, etc. Additionally, in the case that polling is triggered by the mobile a new signalling message will have to be defined. Moreover, a further consideration would be the requirement for larger buffer sizes in the BSS and in the mobile terminal, because of the larger ARQ window size.

The GPRS protocol over the satellite link has been examined and presented in [38] and [31]. These publications compared the performance of the models using

different statistics and parameters and by applying the protocols under different conditions.

We will discuss some of the above results in the context of the ones presented in [31] since, as in our case, the fact that a MEO satellite system is used, allows for some level of comparison. Specifically, the FUNET and Railway models have been used in this publication for an ICO type of satellite network. As a first comment, we can say that as discussed in section 4.4.1, the author draws similar conclusions that the results are 'generally comparable' to those obtained from a terrestrial GPRS system. Comparing our results (figures 4.11-4.14) for the FUNET model with the ones presented in [31] we observe that the throughput and packet delay are in the same range of values for low system loads. Similarly, comparing the Railway results there is agreement especially for the low noise cases, as the model in [31] assumes an ideal propagation channel. Additionally, in both our case and the work presented in this paper, saturation is reached at lower load levels when the Railway model is used instead of the FUNET one, due to the higher collision rate. In both studies, for both the FUNET and Railway models, saturation point is reached at the load value at which the delay starts to increase rapidly.

The MAC layer, and especially the S-Aloha scheme, is extensively discussed in this publication. The high collision rate for smaller packets in a high propagation delay environment is identified as a problem, and solutions are proposed. Similarly, solutions to mitigate the effects of this specific problem, are presented in section 4.4.4. Additionally, the work presented in the next sections investigates performance under multislot operation as well as different ARQ window sizes for different noise

environments. On the other hand, in [31] handover and timing advance issues are also discussed.

4.4.3   Multislot operation
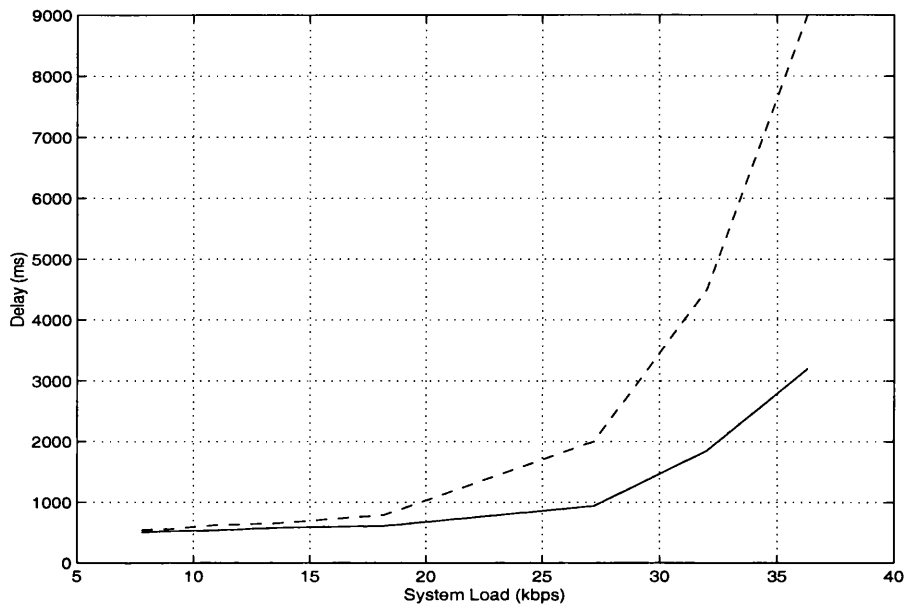
Figure 4.15: LLC frame delay for the Railway model for 3-slot operation: *i.* BLER=0.08 (– line) , *ii.* BLER=0.0002 (solid line)

In this section we examine all three models in a 3-slot environment. This means that the users, when contending for resources, ask for 3 slots to be allocated for transmission and will receive 3 slots maximum depending on resources available.

We observe that after comparing the results in all of the figures in this section (4.15-4.22) with figures produced in the previous two sections, where single-slot operation was used, there is clearly a system improvement in the case where 3-slot operation is used. Again, the improvement is more significant for the FUNET model and less significant for the Mobitex since, as already mentioned, the contention phase is what is placing the Mobitex model in saturation. The comparison is more obvious
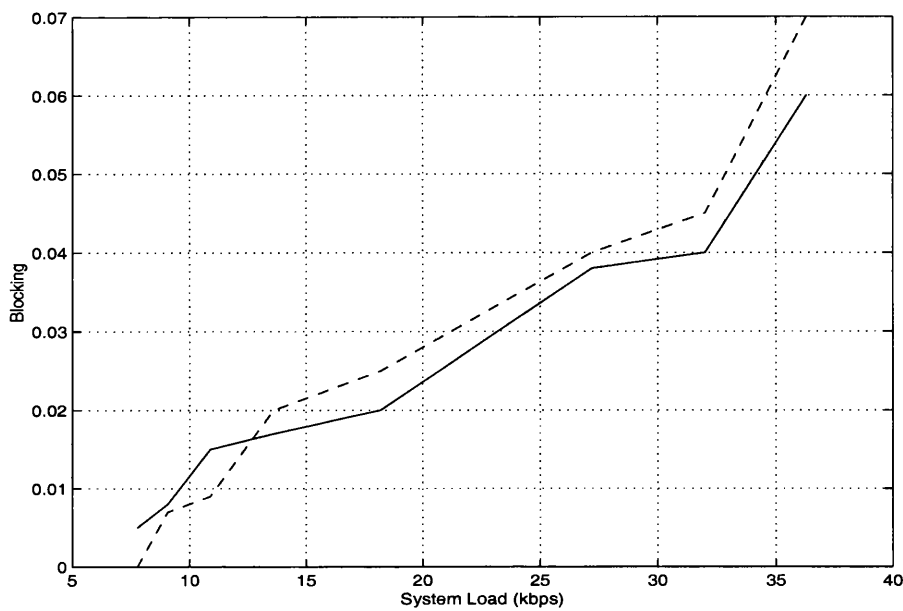
Figure 4.16: Blocking probability for the Railway model for 3-slot operation: *i.* BLER=0.08 (- line) , *ii.* BLER=0.0002 (solid line)
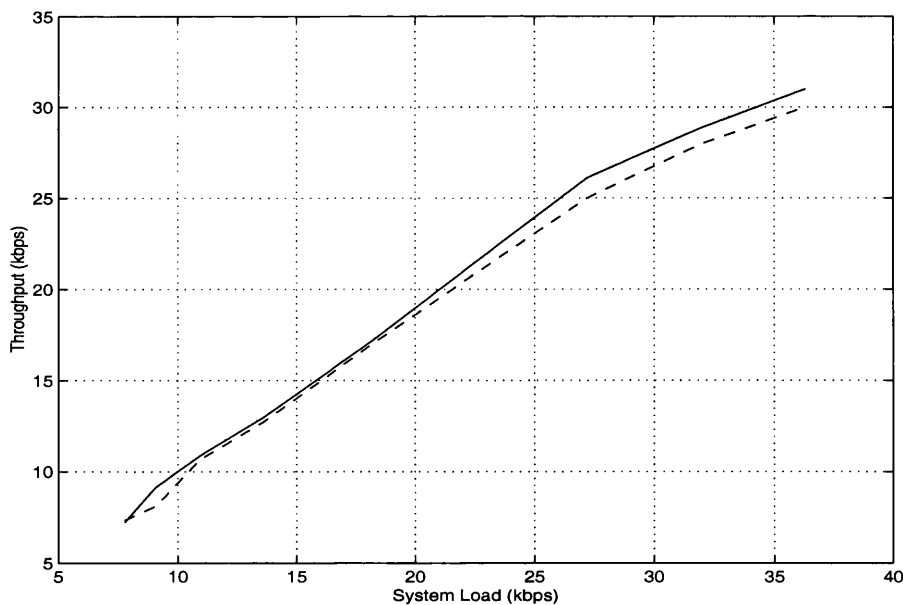
Figure 4.17: Throughput for the Railway model for 3-slot operation: *i.* BLER=0.08 (- line) , *ii.* BLER=0.0002 (solid line)

in tables 4.5-4.7 where the percentage of the delay improvement is shown for the 3-slot versus the single-slot operation. Specifically, the ratio $d_{3s/1s}$, presented in these tables, is equal to:
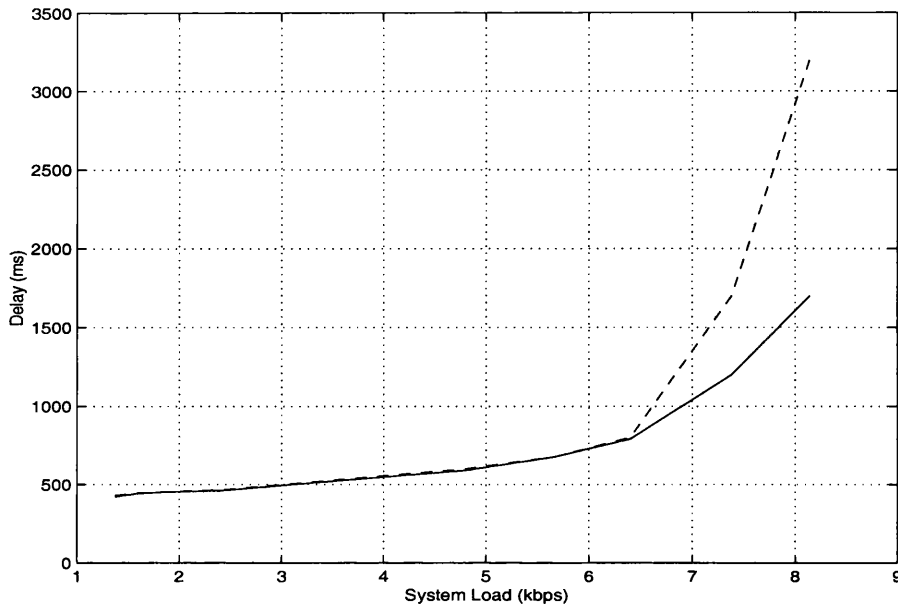
Figure 4.18: LLC frame delay for the Mobitex model for 3-slot operation: *i.* BLER=0.08 (– line) , *ii.* BLER=0.0002 (solid line)
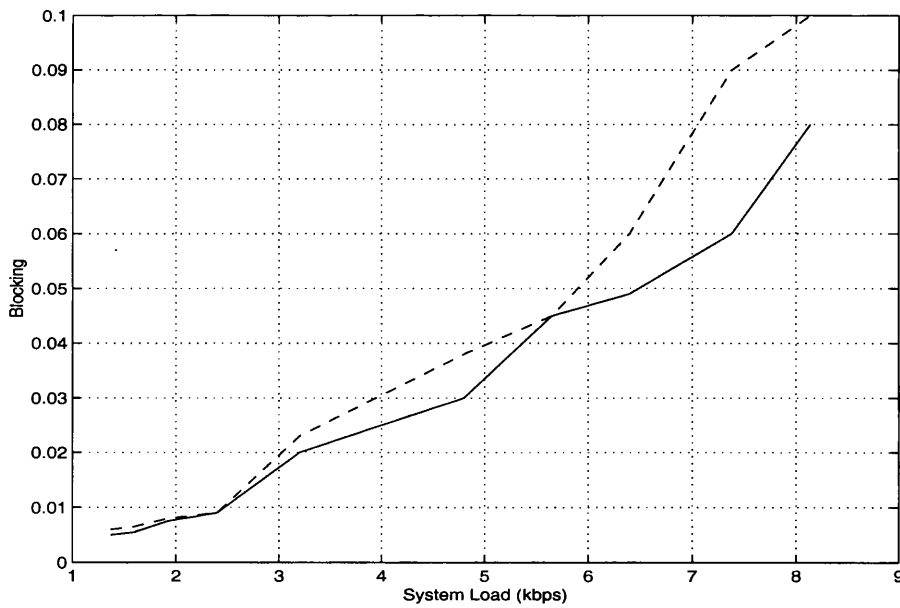
Figure 4.19: Blocking probability for the Mobitex model for 3-slot operation: *i.* BLER=0.08 (– line) , *ii.* BLER=0.0002 (solid line)

$$d_{3s/1s} = \frac{d_{1s} - d_{3s}}{d_{1s}}, \tag{4.20}$$

where d_{1s} is the LLC frame delay in the single-slot system operational mode, while
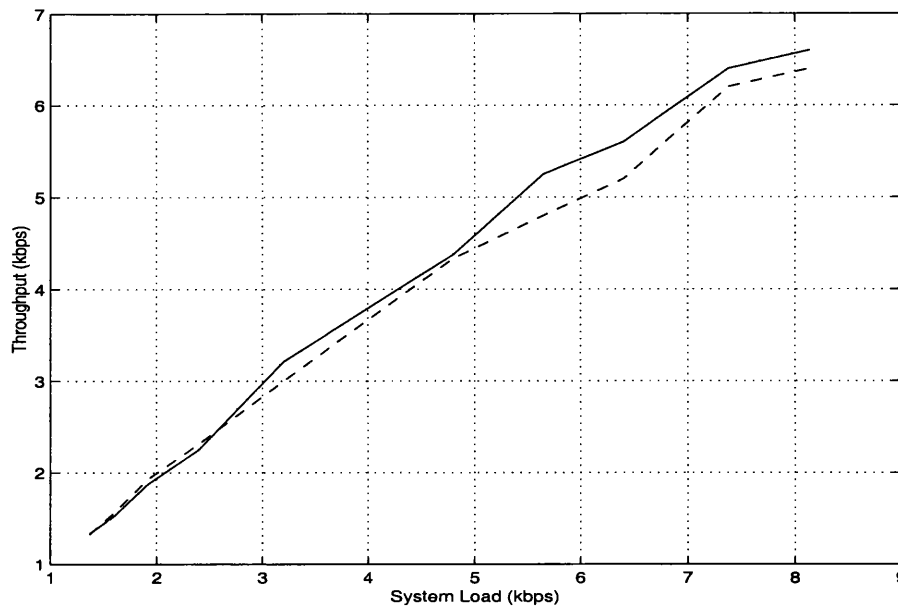
Figure 4.20: Throughput for the Mobitex model for 3-slot operation: *i.* BLER=0.08 (−line) , *ii.* BLER=0.0002 (solid line)

d_{3s} is the LLC frame delay in the 3-slot system operational mode. Comparing the relative improvement in the LLC frame delay, we see the actual impact that multislot operation has when the system load consists of large (FUNET model) packets, instead of small (Mobitex model) packets.

data rate (kbps)	$d_{3s/1s}$ (BLER=0.0002)	data rate (kbps)	$d_{3s/1s}$ (BLER=0.08)
1.37	0.04	1.37	0.065
1.6	0.03	1.6	0.051
1.92	0.056	1.92	0.068
2.4	0.056	2.4	0.07
3.2	0.056	3.2	0.073
4.8	0.06	4.8	0.13
5.65	0.1	5.65	0.16
6.4	0.14	6.4	0.16
7.38	0.14		

Table 4.5: Delay improvement when 3-slot scheme is used instead of single-slot scheme, for the Mobitex model for two different BLERs

data rate (kbps)	$d_{3s/1s}$ (BLER=0.0002)	data rate (kbps)	$d_{3s/1s}$ (BLER=0.08)
7.77	0.16	7.77	0.20
9.07	0.16	9.07	0.21
10.9	0.16	10.9	0.21
13.6	0.17	13.6	0.22
18.2	0.17	18.2	0.28
27.2	0.18	27.2	0.41
32	0.19		

Table 4.6: Delay improvement when 3-slot scheme is used instead of single-slot scheme, for the Railway model for two different BLERs

data rate (kbps)	$d_{3s/1s}$ (BLER=0.0002)	data rate (kbps)	$d_{3s/1s}$ (BLER=0.02)
11.37	0.43	11.37	0.44
13.93	0.44	13.93	0.44
17.71	0.44	17.71	0.45
28.74	0.44	28.74	0.45
37.78	0.45	37.78	0.46

Table 4.7: Delay improvement when 3-slot scheme is used instead of single-slot scheme, for the FUNET model for two different BLERs

4.4.4 Improved performance

Due to the requirement to achieve a proper correspondence and timing relationship between PRACH and both the RA and PNCH messages and since a MEO satellite system has a large propagation delay, a 'safety' margin of 200 ms should be imposed on the T_o timer. The rest of the contention process follows a normal terrestrial model. In the results that will be presented we will refer to the former scheme, as the 'first scheme'. Alternatively, we propose here a scheme, the 'second scheme', where every time the network receives a PRACH message, it stores in a database the time T_s the time when it is serving it either by sending a PNCH or a RA message. Any PRACH received from the same source for a time window of 200 ms after the time T_s is destroyed. The T_s entry is cleared after the 200ms in-
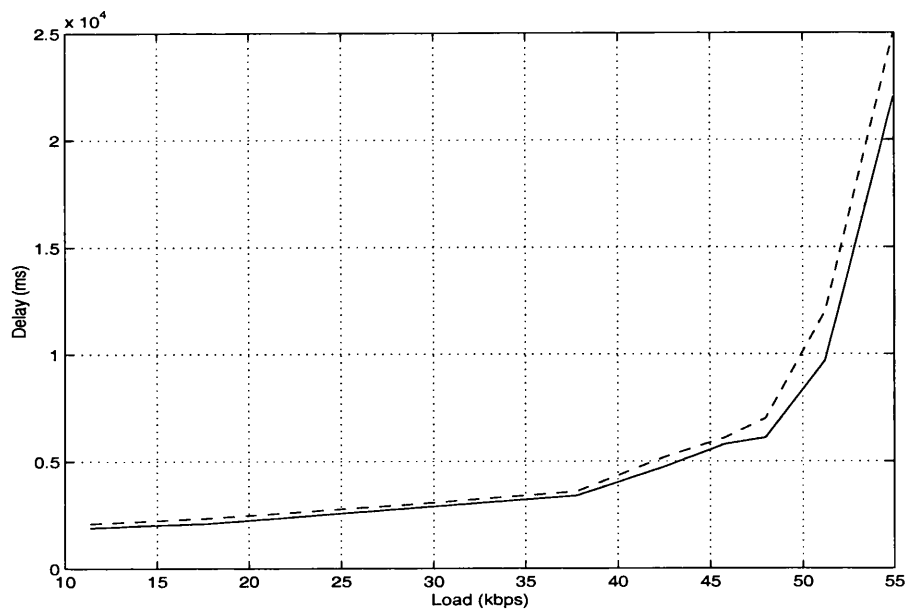
Figure 4.21: LLC frame delay for the FUNET model for 3-slot operation, BLER=0.0002 (solid line), BLER-0.02 (–line), K=30, WIN=64
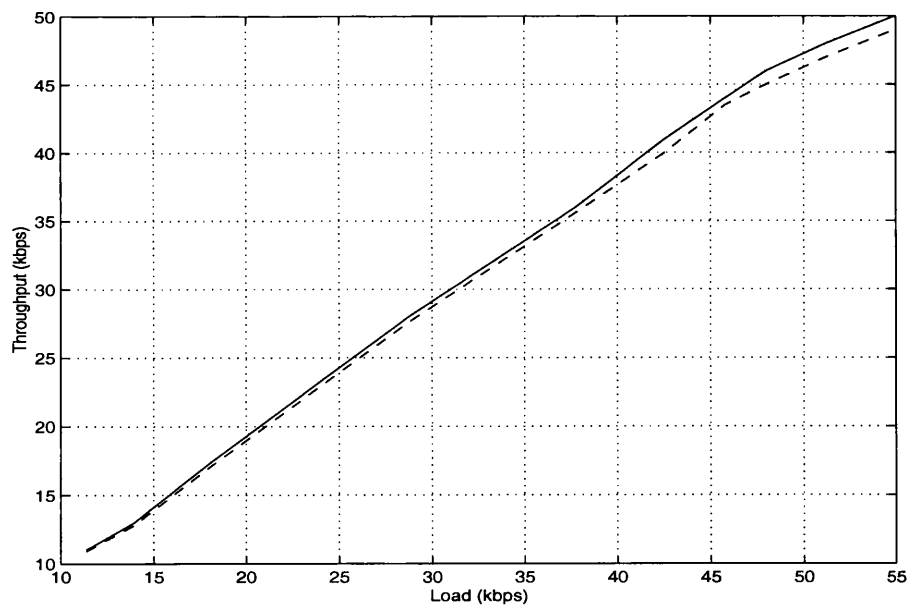
Figure 4.22: Throughput for the FUNET model for 3-slot operation, BLER=0.0002 (solid line), BLER-0.02 (–line), K=30, WIN=64

terval termination. By incorporating the second scheme in our GPRS model, we allowed the MS to contend for resources at more frequent intervals and the system still worked under a properly timed communication. Implementation of this scheme

would involve the maintenance of a list associating a T_s value with a source. Since, as already observed in the previous sections, the contention phase is not of so much importance for the FUNET model, we test only the Mobitex and Railway models for the above two schemes.

In figures 4.23-4.28 the Mobitex and Railway models are tested under the two schemes, for medium and high load conditions, for a single-slot GPRS operation, for CS-1 coding scheme, and a BLER=0.0002. It can be observed that the use of the second scheme (T_o values lower than 200ms) provides a system performance improvement as far as both the LLC frame delay and the system blocking probability are concerned. For the Mobitex model the big improvement is shown in the blocking probability (figures 4.24,4.25), while there is some improvement in the LLC frame delay for the higher loads (figure 4.23). Similar conclusions can be drawn from figures 4.26-4.28 for the Railway model.

In figures 4.23 and 4.26 we would expect the curve to exhibit an initial valley at low values of T_o. The reason for this is that as a mobile increases its frequency of contention there is a trade-off between improving the probability of being assigned a channel sooner and increasing the rate of collisions, which translates into an increased total delay. However, what is illustrated instead is an 'oscillating' curve for low values of T_o. The delay illustrated in the above figures is the average delay of the frames that were actually served. However, if this figure is combined with the ones illustrating the blocking probability, we could see that, since the blocking is improved for lower T_o values, the average number of frames that is served is larger. This means that the system is trying to find a balance between the number of con-

tention messages that are served and the average serving delay. This could explain the 'oscillation' shown in the delay diagrams for low values of T_o, as the system seems to be reaching stable conditions at higher values of this parameter, as the smaller number of served frames allows the system to reach a state of equilibrium. It would be expected to observe the same behaviour for the lower system load curves in both 4.23 and 4.26 diagrams. The explanation for not seeing that kind of pattern in these diagrams, could be that either the amplitude of oscillation is too small for it to appear using the selected scale, or that the 'oscillations' actually occur for lower values of T_o.
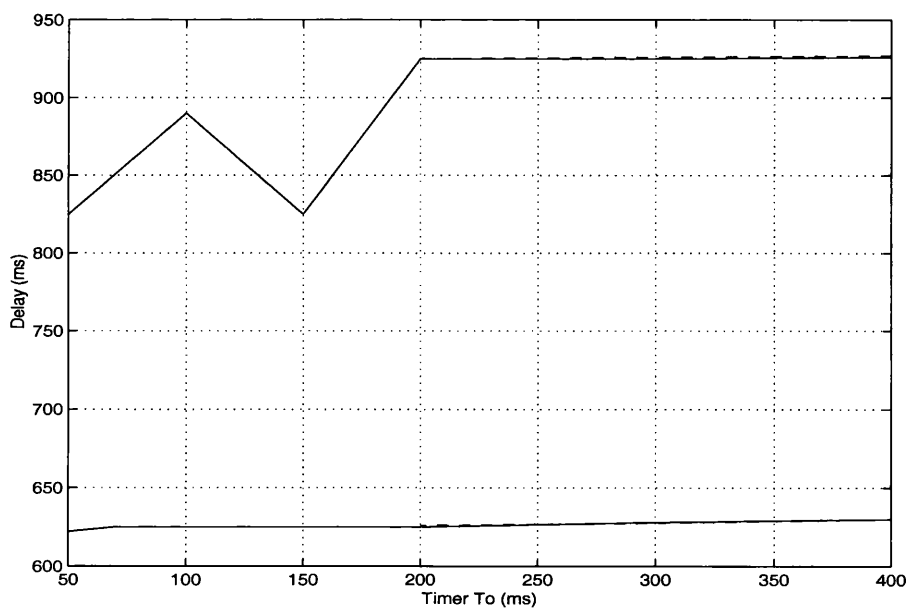
Figure 4.23: LLC frame delay for the Mobitex model for: i. system load = 6.4 kbps, second scheme (upper solid line) , ii. system load = 6.4 kbps, first scheme (upper dotted line) , iii. system load = 4.8 kbps, second scheme (bottom solid line), iv. system load = 4.8 kbps, first scheme (bottom dotted line)
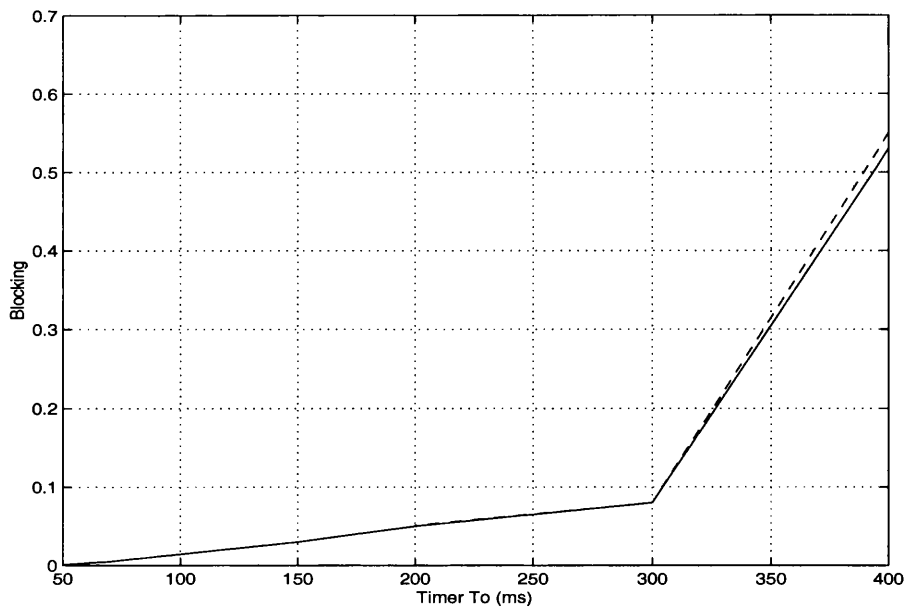
Figure 4.24: LLC frame blocking for the Mobitex model for: i. system load = 4.8 kbps, second scheme (solid line) , ii. system load = 4.8 kbps, first scheme (dotted line)
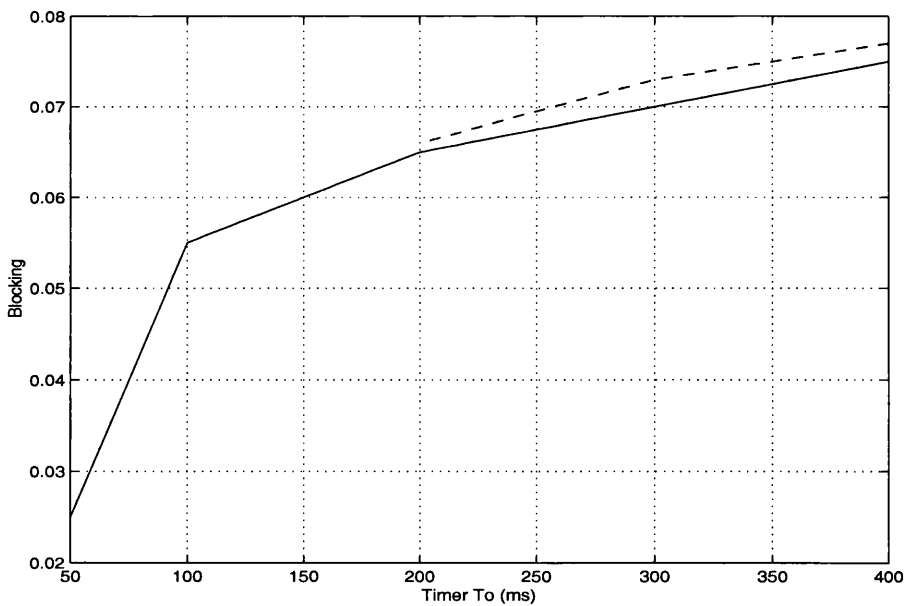
Figure 4.25: LLC frame blocking for the Mobitex model for: i. system load = 6.4 kbps, second scheme (solid line) , ii. system load = 6.4 kbps, first scheme (dotted line)

4.5   Conclusions

In this chapter we examined the aspects of the GPRS RLC/MAC protocol over a

MEO satellite channel. The simulation tool used for developing the specific system
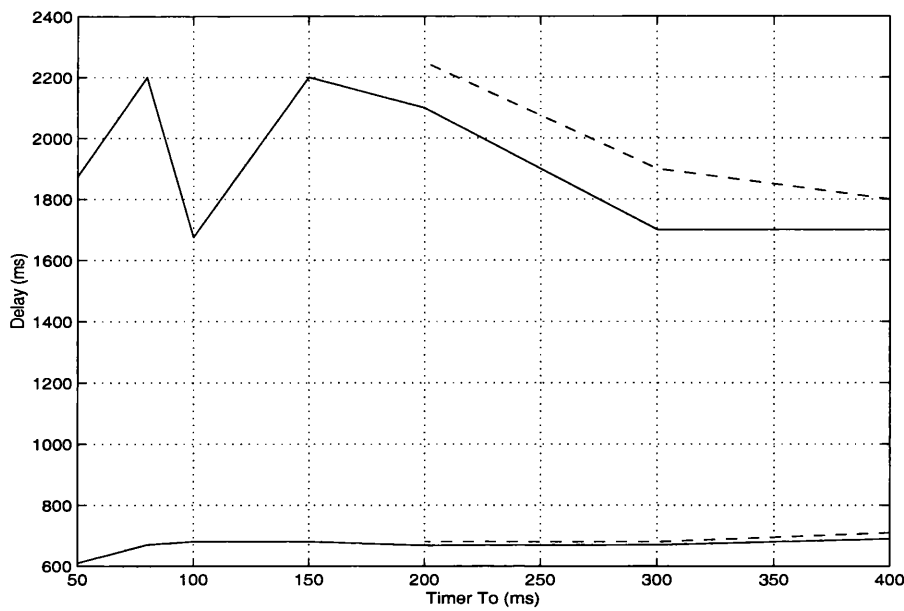
Figure 4.26: LLC frame delay for the Railway model for: i. system load = 32 kbps, first scheme (upper solid line), ii. system load = 32 kbps, second scheme (upper dotted line) , iii. system load = 13.6 kbps, second scheme (bottom solid line), iv. system load = 13.6 kbps, first scheme (bottom dotted line)
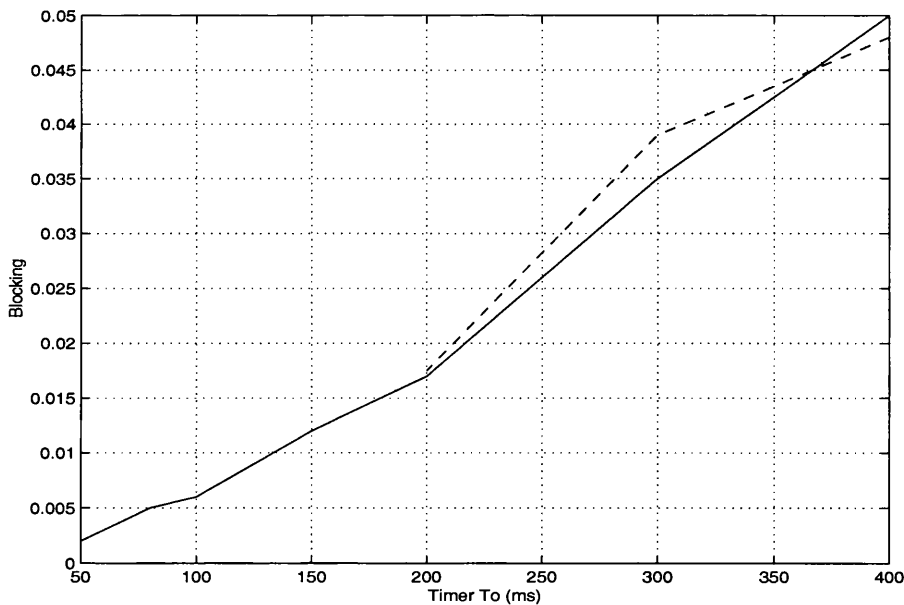
Figure 4.27: LLC frame delay for the Railway model for: i. system load = 13.6 kbps, second scheme (solid line) , ii. system load = 13.6 kbps, first scheme (dotted line)

model was OPNET. A detailed outline of the simulation model and its important

parameters was included in this chapter. Different aspects of the protocol were
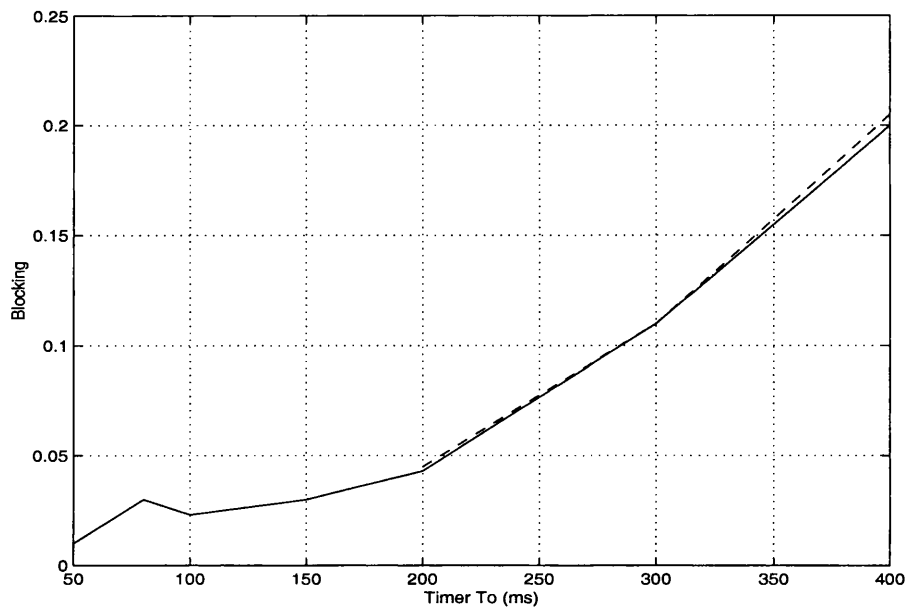
Figure 4.28: LLC frame delay for the Railway model for: i. system load = 32 kbps, second scheme (solid line) , ii.system load = 13.6 kbps, first scheme (dotted line)

examined in order to test its performance in a satellite environment.

Three different traffic models were used with different packet size distributions, in order to enable us to examine the impact of the satellite channel on the protocol's performance under different traffic conditions. The behaviour of the protocol was similar to the one observed when it is applied over a terrestrial channel, however the loads under which saturation was reached were obviously smaller. The system was examined both for single-slot and multislot operation and for different BLERs.

Additionally, different ARQ window sizes were proposed and examined in order to mitigate the degradation, mainly due to the long round-trip delay, of the system performance in different noise environments. Finally, a novel scheme was used to address the delay caused by timing synchronisation problems in the contention phase of the protocol.

In the next chapter we introduce novel algorithms for improving the distribu-

tion of the available resources between mobile satellite users in systems using DCA

techniques.

Chapter 5

Heuristic techniques for time invariant satellite DCA problems

5.1 Introduction

In this chapter we examine the application of heuristic techniques for dynamic channel allocation (DCA) in satellite systems. Specifically, we present the concepts of hill climbing, steepest descent and simulated annealing. Their application to the solution of the problem at hand is studied and the results are discussed. The objective of this chapter is to investigate the applicability of different heuristics, in conjunction with optimisation techniques, to the solution of resource allocation problems in mobile satellite systems. Hence, the resource allocation studies use a time invariant model, before applying them in the next chapter to a more complex time-variant model.

5.2 Combinatorial problems

Let us present the following problem:

$$Optimize \quad g(x), \quad | \quad \omega(x) \leq \alpha_i, \quad i = 1, ..., n. \tag{5.1}$$

In the above setting we are trying to optimise a function of a vector of decision variables x, taking some constraints under consideration. Combinatorial problems are the problems where the decision variables are discrete and hence the solution is formed by discrete objects.

Combinatorial problems that need for their solution a complexity effort that grows as a low-order polynomial, are said to be of class P. NP (Non-deterministic polynomial) problems are the set of problems for which a polynomial order optimisation algorithm has not been proven to exist yet. A problem Q is considered to be NP-hard if every NP problem can be transformed to Q. If Q is an NP problem as well, then Q is considered to be NP-complete. In other words, NP-complete problems are a set of problems, that if a method that guarantees a solution is found for one of them, it can be applied to all the members belonging to that group.

5.3 Heuristic Techniques

Combinatorial optimisation is finding a 'global' or 'local' optimal solution to a combinatorial problem. In order to give a more complete comprehensive description of a 'local' and a 'global' optimum, it is first necessary to describe the concept of

the neighbourhood.

Let $N(s)$ be the set of configurations that can be reached from a specific solution s after a simple change is performed. Then the set $N(s)$ is the neighbourhood for the specific solution s. We also specify a cost function $C(s)$, that gives the cost for each configuration s. We should add here that the optimisation could mean minimisation or maximisation. We will describe below a minimisation process. If we start from a random solution s, by choosing each time a solution s' so that $C(s') \leq C(s)$, we move towards a local optimum of the problem with respect to this cost function. We say that we have reached the local optimum at s_{lo} when all the solutions belonging to $N(s_{lo})$ produce a rise in the system cost. In contrast, the global optimum is the solution s_{gl} such that:

$$\forall \quad s \quad C(s_{gl}) \leq C(s).$$

The global optimum is also a local one, however there may be many local minima with higher values of this function.

In order to treat combinatorial problems, two types of methods can be used: i) Methods that guarantee to produce a globally optimal solution and ii) Heuristics, which are methods that do not guarantee a globally optimal solution, but may reach a locally optimal solution. Specifically, the following definition can be used:

'A heuristic is a technique which seeks good (i.e. near-optimal) solutions at a reasonable computational cost without being able to guarantee either feasibility or optimality, or even in many cases to state how close to optimality a particular feasible solution is' [43].

Heuristics that produce small computational complexity are being used to treat NP-hard combinatorial problems.

A main characteristic of a heuristic is that by applying the same algorithm to the same problem, with different initial conditions, different local optima can be reached. Hence, by increasing the number of the trials, the probability of identifying the global optimum is increased as well. In the following sections, heuristic techniques that were used in our research are presented in detail. These are the hill climbing, steepest descent, and simulated annealing.

5.3.1 Hill Climbing

The term hill climbing is mainly used for maximisation algorithms. We are using the concept of hill climbing for minimisation purposes, which are the ones that apply to our problem. According to this concept, whenever we are located in a specific state we only except 'downhill' moves, i.e. moves that increase the system cost function. Uphill moves, that could produce a temporary increase in the system cost but could eventually lead to lower local optimum, are not accepted. This solution is chosen randomly amongst the ones that also produce a decrease in the cost. Hill climbing can only guarantee a local optimum (minimum in our case) will be reached.

5.3.2 Steepest Descent

Steepest descent is another heuristic technique that can only guarantee convergence towards a local minimum. Specifically, out of all the possible transitions from a specific system state, we except the one that is producing the largest decrease in

the system cost. Therefore, we can say that steepest descent is a special case of hill climbing. It is obvious that by following such a technique the local optimum that will be reached strongly depends on the initial conditions of the problem.

5.3.3 Simulated Annealing

Simulated annealing is a global optimisation technique. The method was initially presented by Metropolis et al [48] for statistical mechanics purposes in order to simulate the behaviour of cooling fluids. According to them, the system is initially put in a very high temperature that gradually decreases until the system is considered to be frozen. During the cooling process, the system goes through different configurations that produce a change ΔE in the system energy. If $\Delta E \leq 0$ the change is accepted, while if it is $\Delta E > 0$ the probability P_α of accepting the change is

$$P_\alpha = e^{-\Delta E / kT},$$ (5.2)

where k is the Boltzman's constant and T the system absolute temperature. The fact that T is progressively reduced during the cooling process, results in a reduced probability of accepting a new configuration that increases the energy levels as the process progresses. Additionally, during the process smaller increases are always more likely to be accepted than larger ones.

Kirkpatrick et al [44] and Cerny [45], later suggested that this method could be applied for the solution of NP-complete combinatorial problems. Specifically, they used simulated annealing in an attempt to minimise a cost function C by making

random changes in the initial configuration. The changes are accepted if, in analogy

with equation 5.2, the difference $\Delta = C(n + 1) - C(n)$ of the new cost function

$C(n + 1)$ and old cost function $C(n)$ is negative, or with probability:

$$p = e^{-\frac{\Delta}{T}}, \quad \Delta > 0$$

where T is called the system 'temperature'. The process stops when no further

improvement in cost is expected. The main advantage of the simulated annealing

method is that, by allowing 'uphill' moves (i.e. positive cost function changes), we

increase the probability of finding the global, rather than the local minimum of a

problem.

5.4 Fixed and dynamic channel allocation techniques

In today's rapidly expanding mobile communication networks, there is an in-

creasing demand for the limited radio spectrum allocated for the use of both ter-

restrial and satellite mobile phone systems. In order to improve the efficient use

of the available channels, the concept of frequency reuse is adopted. According to

this concept, the same set of frequencies can be used in cells that are at a distance

where the co-channel interference is acceptably low. This distance is called the min-

imum reuse distance. Generally, there are two main categories for the techniques

of sharing the available number of channels amongst the mobile users. The first

category is the currently used Fixed Channel Allocation (FCA), where each cell is assigned a fixed set of frequencies which are reassigned in distant cells that satisfy the minimum reuse distance constraint. This is achieved by using a suitable channel reuse pattern. The second category consists of DCA techniques, where the available channels are collected in a common pool. The channels that are extracted from this pool can then be used in any cells in the system as long as these cells satisfy the minimum reuse distance criteria.

The frequency assignment problem is a challenging one. Some characteristic examples in the bibliography investigate several different approaches to its solution. DCA schemes with channel borrowing techniques were used for both land [50] and satellite [51] networks, while in [52] solutions that combine both FCA and DCA schemes for mobile satellite systems are presented.

5.5   Resource Allocation Heuristics

The frequency assignment problem consists of assigning frequencies to mobile stations in a way that for a given level of accepted interference, the number of users that can be served is maximised. This problem has been studied from a network planning perspective [47, 49], where the achievement of an optimal solution requiring the minimum number of channels is NP–hard [46]. In this section, several algorithms with assignment heuristics have been tested for their efficiency and practicality. Another approach which is found in the literature is to formulate the problem as a graph colouring problem and use appropriate algorithms to obtain upper and lower

bounds on the minimum number of required channels [47].

In this section we will study the capacity of a satellite mobile system with DCA

for a set of heuristic techniques. The problem under consideration is as follows:

we have a set of stationary mobile satellite users and a fixed number of channels,

all belonging to a common pool. We assume that we know the location of the

mobiles. The main difference from traditional frequency assignment schemes is that

in our model the channels are distributed to mobiles rather than to cells. The same

channel can be assigned only to users that are located in a distance greater than the

minimum interference distance. This distance for a seven cell FCA frequency reuse

pattern, is equal to (see figure 5.1):
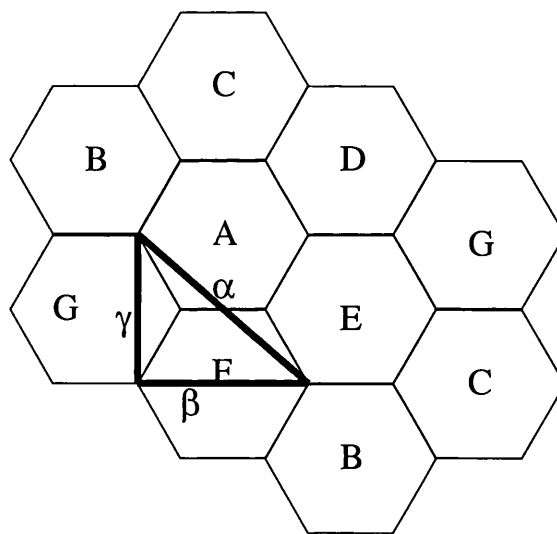
Figure 5.1: A seven cell frequency reuse pattern

$$d_{min} = \sqrt{7/3} \ 2H, \tag{5.3}$$

where $\alpha = d_{min}$, $\beta = 2R$ and $\gamma = 2H$, where R and H are the radii of the circum-

scribed and inscribed circles of the hexagonal cell respectively.

Therefore, the criterion for distributing the channels amongst the users is the distance d_{sa} between mobiles assigned the same frequency, as the co-channel interference is strongly dependent on this distance. Hence, a better 'packing' of the channels is achieved when the values of the minimum distance $d_{sa_{min}}$ between mobiles using the same frequency and the average d_{sa} increase. Several algorithms that attempt to find a good solution were investigated.

The model described in this chapter, models the *instantaneous* conditions of the network and does not include any particular handover schemes. Therefore, the analysis and results apply to all types of cell-based satellite systems. In the next chapter, we will model time variant systems by optimising the frequency allocation at particular time instances, each time using the principles described here.

Optimisation framework

We considered a 10000 km^2 satellite footprint area[1] as an area of study. The number of channels available is 56. For each algorithm under consideration, several problems were studied. For each problem, different sets of initial conditions were used to increase the probability of finding a good solution. In the first case, all mobiles have the same frequency assigned to them, whilst in the second case each mobile is assigned a random frequency. The first case is the worst possible initial state, providing a big scope for improvement.

---

[1]This footprint area leads to small cell sizes when compared to the ones of traditional satellite systems. The choice of a small cell size has the advantage of limiting the simulation run times, as for the same mobile density per cell, a model with a smaller cell size would require a fewer number of events to reach stable conditions. Results for larger cell sizes (like the ones currently used in S-PCN) should be similar to the results described in this chapter, as for the same mobile density per cell, the same performance could be achieved by using a larger pool of frequencies.

According to the heuristic applied each time, each user can move to a different channel, until the maximum $d_{sa_{min}}$ is achieved. By the end of the process the frequencies are distributed to the mobiles evenly. The statistics gathered are the values of $d_{sa_{min}}$ and average d_{sa}.
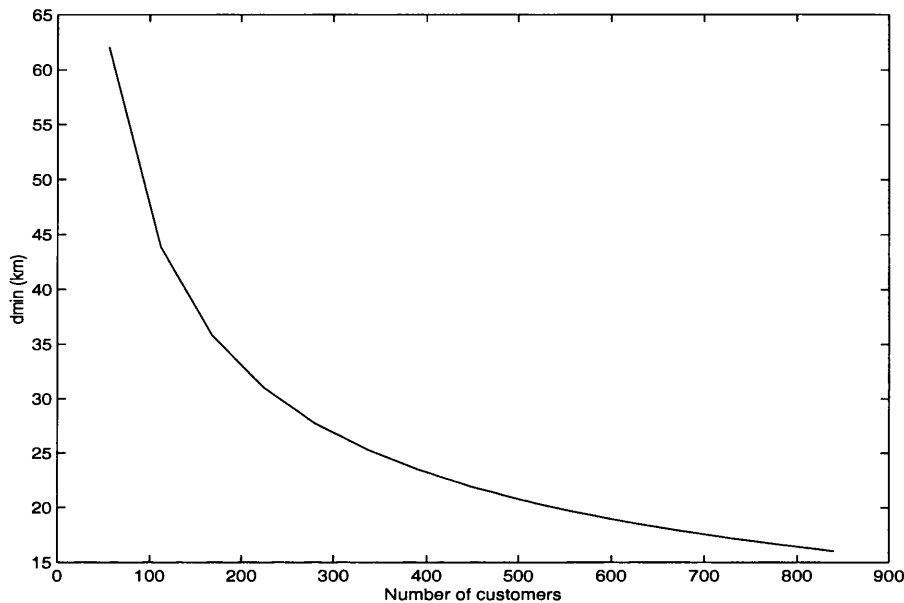
Figure 5.2: The interference radius d_{min} versus the maximum number of customers that can be served in the area A

Number of mobiles	d_{min} (km)
100	46.42
200	32.82
300	26.81
400	23.21
500	20.76
600	18.95
700	17.55
800	16.41

Table 5.1: The maximum number of customers that can be accommodated in a seven cell frequency re-use scenario versus the minimum interference distance d_{min}

We have an area A of 10000 km^2. If we assume a 7 cell frequency reuse pattern,

this gives a number N of cells equal to:

$$N = \frac{A}{R^2 \frac{3\sqrt{3}}{2}},$$
(5.4)

where R is the cell radius. If we also consider equation 5.1, then the minimum

interference distance d_{min} is equal to:

$$d_{min} = \sqrt{7}\sqrt{\frac{A}{N}\frac{2}{3\sqrt{3}}}.$$
(5.5)

If we have k seven cell clusters that 'fit' into the area A then we have:

$$d_{min} = \sqrt{7}\sqrt{\frac{A}{7k}\frac{2}{3\sqrt{3}}}.$$
(5.6)

Each seven cell cluster is assigned 56 frequencies. In the best case where the cus-

tomers are evenly distributed in the cells we can accommodate a maximum of 56

customers per cluster. In figure 5.2 we can see how the number of maximum cus-

tomers that can be served in the area A varies with the minimum interference

distance d_{min}.

In the table 5.1 we present several values from equation 5.6 that will allow

comparison between the system performance for the above scenario and the system

performance after the use of heuristics.

Algorithms and cost functions

Several distinct algorithms were investigated and the more efficient ones are presented here:

i Let $d_{ab_{f_x}}$ be the distance between any pair (a, b) of mobiles using the same frequency f_x. Both mobiles that belong to the pair (i,j) for which $d_{ij_{f_x}}$ becomes minimum, are tested with all other frequencies. We assign to one of the two mobiles (i or j) the frequency f_s, for which the $\max\{d_{ik_{f_s}}, d_{jk_{f_s}}\}$ takes the maximum value over all pairs (i,k) and (j,k) with $f_s \neq f_x$ and all k having f_s assigned to them.

ii The optimisation techniques described in section 5.3 used with the cost function:

$$C(n) = \sum_i \sum_j \frac{1}{d_{ij}{}^n} \quad | \quad n > 0, \tag{5.7}$$

where d_{ij} is the distance between any (i,j) couple of active mobiles.

iii The optimisation techniques described in section 5.3 used with the cost function:

$$C(n) = \sum_i \sum_j u(i, j),$$

$$u(i, j) = \begin{cases} \kappa, & \text{if } d_{ij} < d' \\ \frac{1}{d_{ij}{}^n} \quad | \quad n > 0, & \text{if } d_{ij} \geq d' \end{cases} \tag{5.8}$$

where κ is a constant with $\kappa \gg d_{max}$, d_{ij} is the distance between any (i,j) couple of active mobiles, d_{max} is their maximum distance possible and d' is an optimisation parameter.

The first algorithm is a heuristic that cannot be easily classified under the three different categories that were presented in section 5.3. However, it can be considered as a form of hill climbing since each change always leads to an improved system state. This is an algorithm that aims at a very local improvement of the situation, by changing the frequency in the pair of mobiles, whose interference contributes the most in the system interference.

The second and third algorithms were used for different values of n in combination with the three heuristics presented in section 5.3. Specifically, the second cost function is improving the situation by leading to a better (smaller) system cost function, which at the same time leads to the increase of d_{ij} and eventually of $d_{sa_{min}}$.

According to the third cost function we penalise all the active mobiles that are closer than the minimum interference distance by having them contributing a disproportionately large 'weight' to the cost function, while at the same time favouring the reuse of each channel at mobile stations slightly further apart than d'. An appropriate value for the distance d' is evaluated experimentally for every different parameter case in our simulation.

5.5.1   Results and Discussion

Initially, we have applied the second cost function in combination with all three

heuristics presented in section 5.3. The best results were achieved when n in equation

5.8 took the values 1 and 2. Specifically, for n=2 we obtained good results for all the

three heuristic techniques, while for n=1 only the hill climbing technique produced

good results.

Figure 5.3: Results for the $d_{sa_{min}}$ for the random initial configuration after the use of the three heuristics presented in section 5.3 and the second cost function (n=1,2) presented in section 5.5 (simulated annealing and n=2 marked with solid line, hill climbing and n=2 marked with '-.', steepest descent and n=2 marked with '..', hill climbing and n=1 marked with '–')

In figures 5.3 and 5.7 the $d_{sa_{min}}$ and the average d_{sa} are presented for the random

initial configuration while in figures 5.5 and 5.9 the $d_{sa_{min}}$ and the average d_{sa} are

presented for the worst initial conditions. In particular, in all graphs the results

marked with the solid lines, the lines marked with (-.) and the dotted lines are

produced after using simulated annealing, hill climbing and steepest descent respec-

Figure 5.4: Results for the $d_{sa_{min}}$ for the random initial configuration after the use of the first algorithm presented in section 5.5 (marked with '...') and the combination of simulated annealing with both the second (n=2) (solid line) and third cost functions (marked with '–') presented in section 5.5

tively, for the value of n=2. The lines marked with (- -) are the results we obtained

using the hill climbing technique for n=1. The conclusion that can be reached

from all four figures is that the best system performance is obtained by simulated

annealing and the worst by the use of n=1 in combination with hill climbing.

no of mobiles	Random initial conditions		Worst initial conditions	
	$d_{sa_{min}}$(km)	d_{sa} avg(km)	$d_{sa_{min}}$(km)	d_{sa} avg (km)
100	54.23	70.10	54.50	70.65
200	37.64	43.32	37.35	43.05
300	29.75	32.92	29.59	33.37
400	26.15	28.49	25.88	28.20
500	21.89	25.51	22.63	24.81
600	20.56	22.35	20.51	22.33
700	18.93	20.28	18.83	20.43
800	17.60	18.80	17.20	18.63

Table 5.2: Results for d_{sa} and $d_{sa_{min}}$ for the first algorithm presented in section 5.5

The use of the first algorithm presented in section 5.5 improved the system

Figure 5.5: Results for the $d_{sa_{min}}$ for the worst initial configuration after the use of the three heuristics presented in section 5.3 and the second cost function (n=1,2) presented in section 5.5 (simulated annealing and n=2 marked with solid line, hill climbing and n=2 marked with '-.', steepest descent and n=2 marked with '..', hill climbing and n=1 marked with '–')

no of mobiles	Random initial conditions		Worst initial conditions	
	$d_{sa_{min}}$(km)	d_{sa} avg(km)	$d_{sa_{min}}$(km)	d_{sa} avg (km)
100	59.58	72.73	59.42	72.73
200	39.84	46.64	39.37	46.59
300	27.78	33.05	27.99	32.77
400	24.11	28.35	23.41	28.23
500	21.32	25.08	21.53	25.01
600	19.86	22.89	18.79	22.83
700	16.91	20.88	17.18	20.90
800	15.96	19.32	15.24	19.10

Table 5.3: Results for d_{sa} and $d_{sa_{min}}$ for a combination of simulated annealing and the second cost function (n=2) presented in section 5.5

performance, although it was taking into consideration only local changes in order to improve the situation. However, the use of simulated annealing in combination with the third algorithm, has proved to be the most efficient one and produced the best results achieved. In figures 5.4 and 5.8 the $d_{sa_{min}}$ and the average d_{sa} are presented for the random initial configuration while in figures 5.6 and 5.10 the $d_{sa_{min}}$ and the

Figure 5.6: Results for the $d_{sa_{min}}$ for the worst initial configuration after the use of the first algorithm presented in section 5.5 (marked with '...') and the combination of simulated annealing with both the second ($n=2$) (solid line) and third cost functions (marked with '–') presented in section 5.5

no of mobiles	Random initial conditions		Worst initial conditions	
	$d_{sa_{min}}$(km)	d_{sa} avg(km)	$d_{sa_{min}}$(km)	d_{sa} avg (km)
100	60.42	72.44	60.43	72.35
200	45.74	47.55	46.05	47.50
300	34.01	35.36	34.05	35.41
400	31.01	31.66	30.26	31.25
500	26.35	27.56	26.35	27.07
600	24.01	24.56	23.67	24.58
700	22.01	22.57	22.02	22.52
800	21.01	21.36	21.00	21.37

Table 5.4: Results for d_{sa} and $d_{sa_{min}}$ for a combination of simulated annealing and the third cost function presented in section 5.5

average d_{sa} are presented for the worst initial conditions. Specifically, in all figures

the solid lines are illustrating the results produced after using simulated annealing

with $n=2$ for the second cost function presented in section 5.5. These results are

included in these figures for comparison purposes with the other algorithms. The

dotted lines are the results obtained after the use of the first algorithm presented
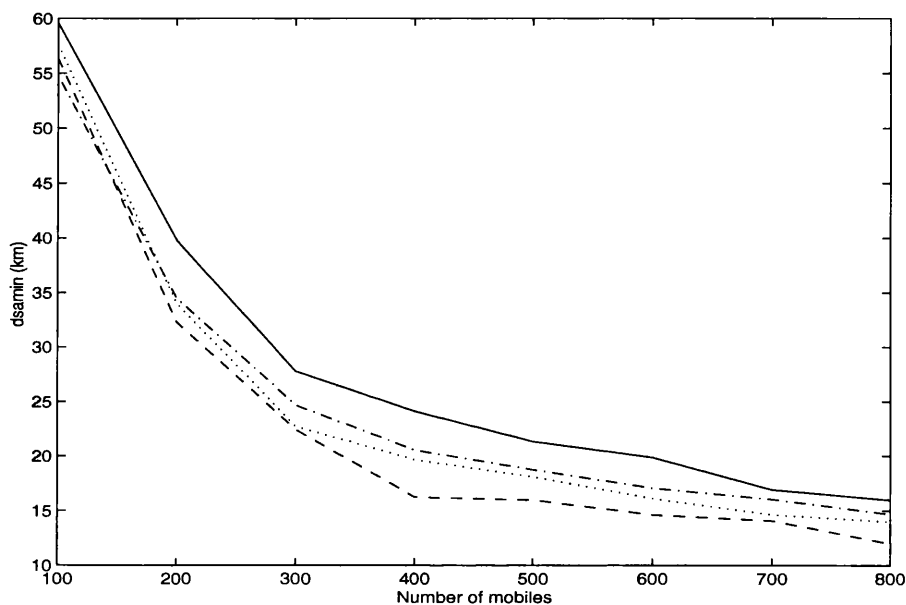
Figure 5.7: Results for the d_{sa} average for the random initial configuration after the use of the three heuristics presented in section 5.3 and the second cost function (n=1,2) presented in section 5.5 (simulated annealing and n=2 marked with solid line, hill climbing and n=2 marked with '-.', steepest descent and n=2 marked with '..', hill climbing and n=1 marked with '–')

in section 5.5 and the lines marked with (- -) are the results produced after the use

of simulated annealing and the third cost function from section 5.5. The results

illustrated in the four above figures can also be found in tables 5.2-5.4. From these

tables we can observe that simulated annealing with the use of the third cost function

produced higher values for both $d_{sa_{min}}$ and d_{sa} average. Additionally, we observe

that the values of $d_{sa_{min}}$ seem to approach the values of d_{sa} average, more than

they do with the use of either of the other algorithms. This is becoming more clear

for higher values of the number of mobiles. Two conclusions can be reached from

this. The first one is that with the use of this specific algorithm we have a very

even distribution of the frequencies amongst the users since the average value of the

distance d_{sa} is very close to the minimum value of d_{sa}. The second one is that in
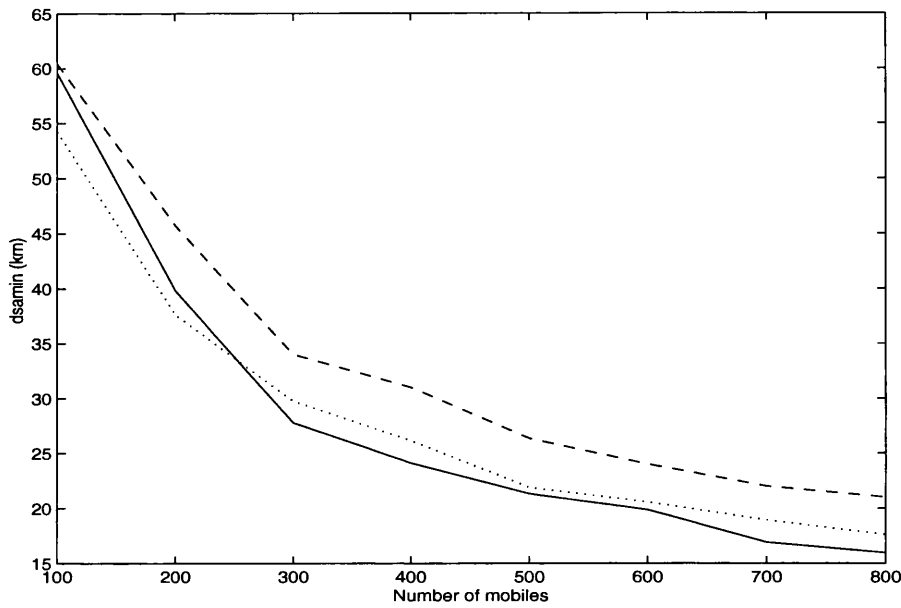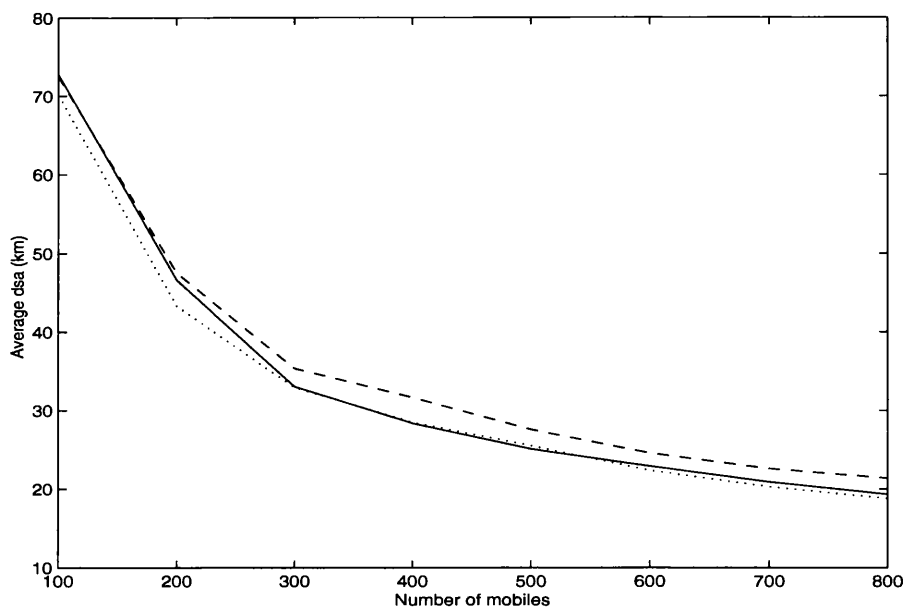
Figure 5.8: Results for the d_{sa} average for the random initial configuration after the use of the first algorithm presented in section 5.5 (marked with '...') and the combination of simulated annealing with both the second ($n{=}2$) (solid line) and third cost functions (marked with '–') presented in section 5.5

these specific cases we seem to approach the upper limit of these values and therefore we approach the optimal solution of the problem.

We also observe in all the above figures and tables 5.2-5.4 that by imposing a larger degree of difficulty in the initial conditions, the system performance becomes slightly worse but is not seriously affected, especially in the case of simulated annealing.

If we compare the results from figure 5.2 with those in figures 5.3 and 5.5 we can see that the steepest descent and the hill climbing techniques after the use of the specific cost function do not give improved results over the distribution for the fixed cell scenario. The results are improved only with the use of simulated annealing which can also be seen if we compare tables 5.1 and 5.3, but as the number of customers increases, simulated annealing with the use of this specific cost function
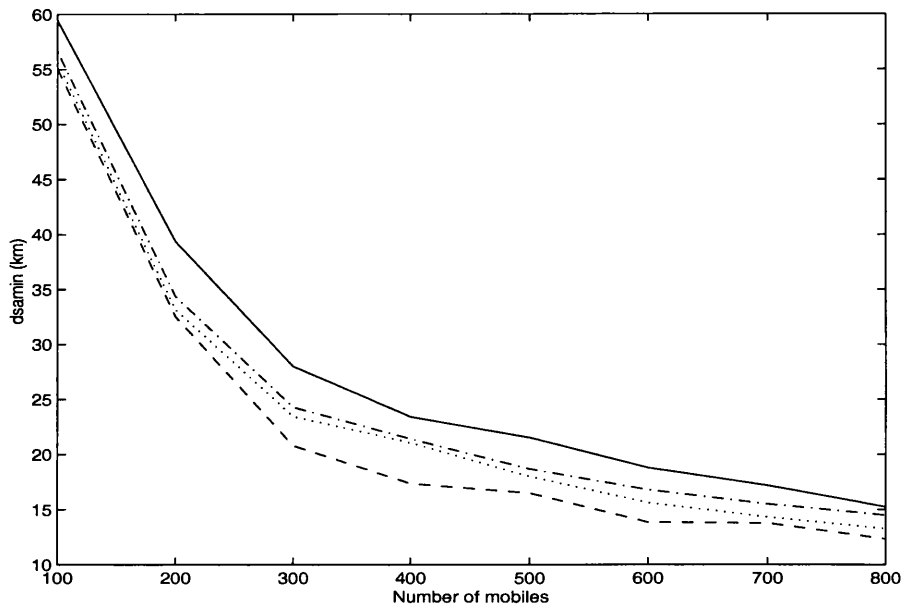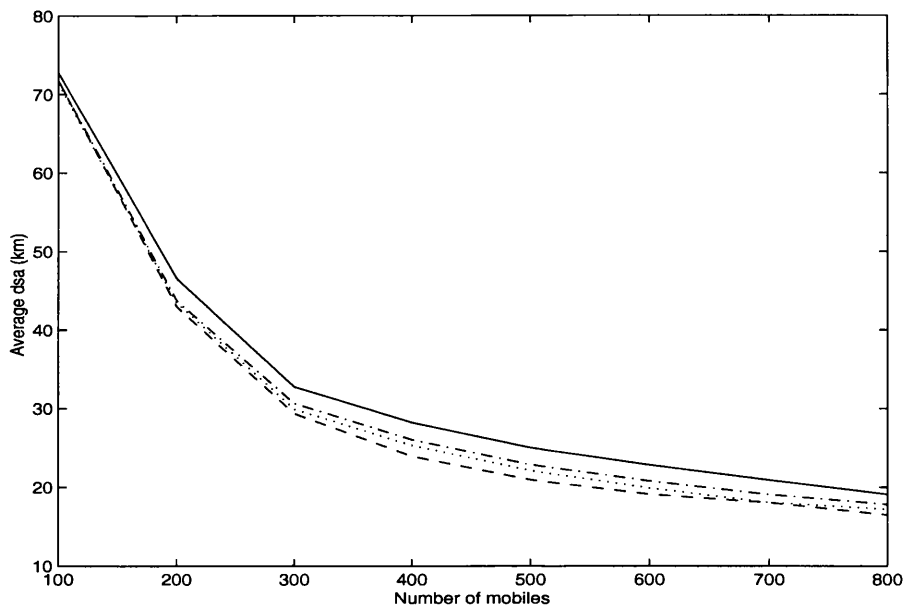
Figure 5.9: Results for the d_{sa} average for the worst initial configuration after the use of the three heuristics presented in section 5.3 and the second cost function (n=1,2) presented in section 5.5 (simulated annealing and n=2 marked with solid line, hill climbing and n=2 marked with '-.', steepest descent and n=2 marked with '..', hill climbing and n=1 marked with '–')

gives poor results. However, if we compare table 5.1 with tables 5.2 and 5.4 we observe that these two heuristics give a far better distribution of frequencies, as the number of customers that can be served for a specific d_{min} increases. It should also be noted once more that the numbers that appear in table 5.1 is the upper limit of the customers that can be served in a fixed cell scenario.

5.6 Conclusions

In this chapter, we have used heuristics for improving the distribution of the available channels between mobile satellite users in systems using DCA techniques. The systems examined here were time invariant ones, which means that the research was examining techniques that would allow us to explore the maximum capacity of a
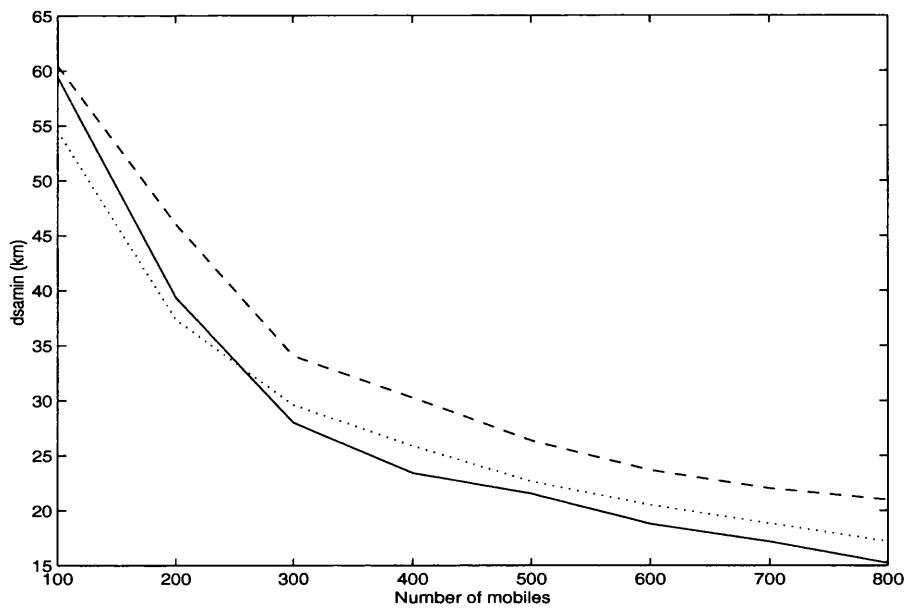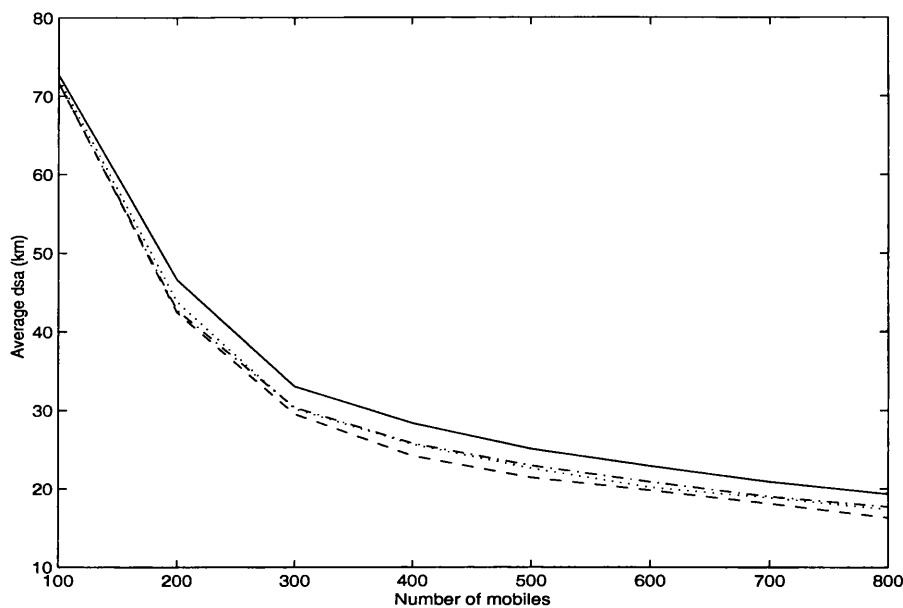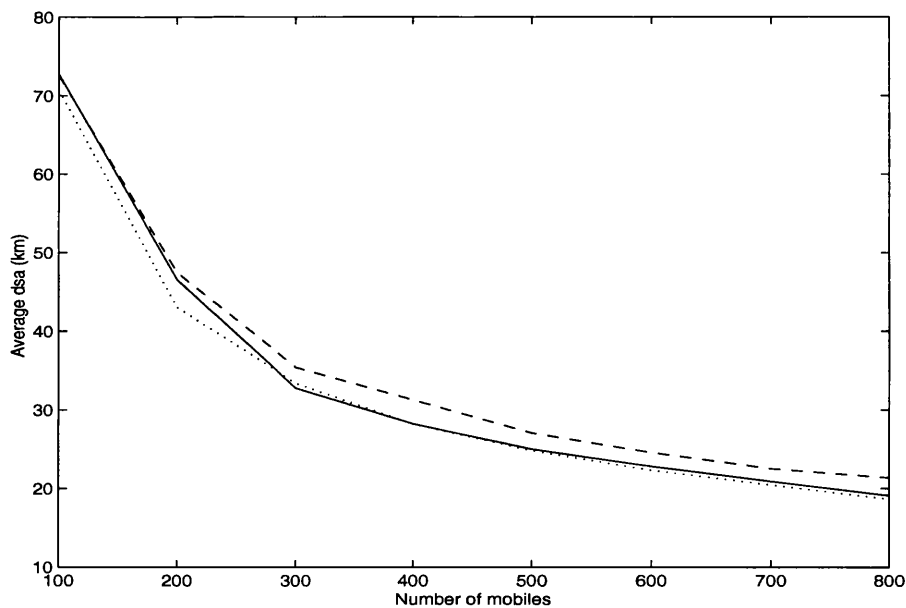
Figure 5.10: Results for the d_{sa} average for the worst initial configuration after the use of the first algorithm presented in section 5.5 (marked with '...') and the combination of simulated annealing with both the second (n=2) (solid line) and third cost functions (marked with '–') presented in section 5.5

system, in terms of number of customers supported, at a particular instant in time, taking into account only the spatial distribution of frequencies.

Three heuristic techniques, namely hill-climbing, steepest descent and simulated annealing, were used in combination with different cost functions. The results were compared to the ones achieved through a traditional seven-cell frequency re-use system with FCA technique. The use of the heuristics introduced a significant improvement in the system performance and the most efficient results were produced with the use of simulated annealing. This was expected due to the nature of this very effective optimisation technique. Another factor that allowed us to experience system performance improvement was the fact that the techniques we proposed allowed the dynamic allocation of available resources to users rather than to cells.

The objective of this chapter has been the applicability of these heuristics, in

conjunction with the optimisation techniques, to the solution of instantaneous re-source allocation problems in mobile satellite systems. In the next chapter we will present DCA channel assignment strategies, similar to the ones presented in this chapter, for time variant mobile satellite communication systems.

Chapter 6

Improved dynamic channel allocation for satellite systems: novel algorithms and constant temperature annealing

6.1 Introduction

In the previous chapter the ground work has been set in order to assess different optimisation techniques, as well as cost functions, that would be applicable to the research work presented in this chapter. It has been shown that simulated annealing produced very efficient results when compared to other heuristics. In particular, the combination of simulated annealing with a specific cost function produced a set of results that approached the optimal solution of the specific problem at hand

and performed equally satisfactorily under both good and bad initial conditions. Additionally, the idea of allocating frequencies to mobiles rather than to cells in a DCA environment, was tested and shown to improve the results compared with the ones produced in a traditional FCA environment. In this chapter, the problem is being explored in a more realistic case scenario with an additional degree of freedom. Specifically, the added complexity lies in the fact that the problem is now time variant; mobility patterns are introduced for the users and therefore different statistics will have to be collected in order to assess the efficiency of a technique that has produced satisfactory results in a time invariant mode.

In this chapter different channel allocation techniques are evaluated to improve the performance of mobile satellite communication systems. Specifically, novel DCA schemes are introduced and compared with existing FCA and DCA schemes as far as their effect on the system blocking probability is concerned. Additionally, a constant temperature simulated annealing algorithm is applied to our dynamic model to achieve better results.

6.2   General model background and description

6.2.1   Model Background

The initial idea, as introduced by Nokia Mobile Systems, was a system named 'Octopus' according to which mobile satellite users would be served by steerable spotbeams that would track down and 'follow' the user, on an individual basis, during the communication period. However, before elaborating on the development

of the idea presented in this chapter, some additional feedback on array satellite systems' literature is needed and is provided below.

- Traditional array satellite systems: As also mentioned in chapter 2 the satellite cell is defined as an area covered by a satellite spotbeam. Different frequencies are assigned to different spotbeams and same set of frequencies can be re-used, as in a land-mobile system, as long as the spatial isolation between spotbeams allows for acceptable signal-to-noise ratio levels. In a traditional multibeam array satellite system the spotbeams generated are of fixed size and shape and move along with the satellite without any steering capabilities.

- Beam steering and dynamic shaping: Future satellite array systems can allow for beam steering (systems presented also in chapter 2), by controlling the phases of the signals prior to the feed elements that are used for the antenna beam forming. Additionally, systems with independent beam steering capabilities can be enhanced with dynamic beamforming techniques [53]. The result is a satellite system that can cover 'hot-spots' in the satellite coverage footprint in a dynamic way, by producing spotbeams that are steered and shaped dynamically in order to match the coverage needs of an arbitrarily changing traffic pattern.

- Mobiles are allocated an individual beam: The above concept of multiple independently steered, arbitrarily shaped beams, through the use of digital beamforming, is extended further now in [54]. In the system proposed here, each mobile is tracked by a separate individual beam that carries a single

channel. In order to avoid interference with other mobiles that are served by different beams, but using an identical carrier frequency, the main lobe for the served mobile has to be aligned, at each moment in time, to the specific location of the mobile and at the same time produce nulls for the other mobiles served by that frequency.

Octopus was planned to adopt similar principles as the third system mentioned above. The model used for the research, as described below, is built on these principles.

6.2.2 Model Description

As already mentioned, in this chapter our model is 'dynamic' in time rather than the 'static' one used in chapter 5. However, the first initial important assumption being made is that the location of the mobiles, as in chapter 5, is known at any moment in time. This is a departure from existing practice, but we expect it to be feasible in future systems, where either a dedicated satellite based navigation system or a location system based on the communication satellites could be used.

Additionally, we assume a DCA system in which all the channels belong to a common pool. Every time there is a request for a new connection, the assignment of a new channel is decided based on a cost function that is described in the next section. Ideally, such a cost function would be directly related to the blocking probability of the system, allowing the selection of the channel that would lead to the lowest value of the blocking probability. However, this is not computationally

feasible, necessitating the use of system cost functions.

In our model, every user in a call is tracked down by a single steerable individual channel/frequency 'pencil-shaped' beam, as described in the previous section.

In the specific model we assume a fixed-earth cell satellite system (the basic principles of which have been presented in Chapter 2). Therefore, in our model a cell is defined by a geographical region, using a frequency carrying broadcast and synchronisation type of information as well as resource and mobility management messages. Steerable spotbeams can follow the user within a cell or across cells. It is important to note here that, in our proposed model, the concept of a cell is limited to functions like the ones mentioned above. However, the cell based FCA and DCA models that will be used for comparison will be applied in this (fixed-earth) cell layout, as also described in section 'Simulation Framework' of this chapter (the principles of this are also presented in [59]).

Since the location of the mobiles is known we can estimate the co–channel interference between any pair of mobiles using the same frequency, based on their location within the satellite footprint. Hence, whenever we need to allocate or reallocate a channel to a user, we choose a channel that is being re-used at a distance that satisfies the interference level criterion. Therefore, we can state that in this particular model our channels are distributed to mobiles rather than to cells as in chapter 5. The channel used by a mobile remains the same, even if the mobile changes 'cell', as long as the same channel is not assigned to a mobile that is closer than the minimum co–channel interference distance. The redistribution of the channels amongst the users occurs when the distance between users that are assigned

the same frequency is such that the co–channel interference is above the acceptable levels. For the scope of this work, this distance is assumed to be equal to the one already calculated in section 5.5.

The principle of optimally allocating channels to users and not to cells has been analysed in [47] for a terrestrial circuit-switched mobile network using graph colouring techniques.

Additionally, the advantage of dynamically sharing resources amongst users on a per packet basis in a packet-switched broadband satellite network, while satisfying interference constraint criteria, is addressed in [26]. In this paper the topology of the satellite network is constantly changing since the system deploys steerable beams that point independently and hop in between fixed earth cells on a per packet basis. Simple heuristic models along with Monte Carlo simulations are used to evaluate the system performance. The model presented in this paper uses some of the basic principles that Octopus was planned to adopt, but the feature of hopping on a per packet-basis, is introduced in order to tackle the burstiness of a packet-switched broadband system.

6.3 DCA technique

In this section we present the cost functions that may be used when a new channel is allocated to a mobile or a re–allocation takes place due to the increasing level of interference between mobiles. We have examined numerous cost functions. In this section we present the three that produced the best results. The first two

can be considered to be two different versions of a cost function used by *Del Re et al* [55]. The third one is a quite simple round robin method.

In order to present a more coherent view of the first two algorithms, some definitions have to stated. We define as Λ_j the frequency assigned to the jth mobile. Let $N_j(r)$ be the set of mobiles that are inside a circle of radius r and centre the location of the jth mobile. Therefore, $N_j(d_{min})$ is the set of users in the jth mobile's interference area.

We also define as $\Omega(j)$ the set of frequencies that are available for the jth mobile which are the channels that are not assigned to any mobile belonging to $N_j(d_{min})$.

The algorithms can be described as follows:

i $\quad \forall\, i \in \Omega(j)$, we have the following cost function:

$$C_j(i) = \sum_k \delta_k(i) \quad | \quad k \in N_j(2d_{min}),$$

$$\delta_k(i) = \begin{cases} 1, & \text{if } i = \Lambda_k \\ 0, & \text{otherwise} \end{cases}$$

(6.1)

Then we choose the channel that results in the maximum value of the cost function. For each channel that is available for the jth mobile, we therefore choose the one that is used most often by mobiles that are located in the annular ring with inner radius d_{min} and outer radius $2d_{min}$. This ring includes the part of the interference area of any user belonging to $N_j(d_{min})$ that is not in the jth mobile's interference area.

ii $\forall\, i \in \Omega(j)$, we have the following cost function:

$$C_j(i) = \sum_k \delta_k(i) \quad | \quad k \in N_j(d_{min}),$$

(6.2)

$$\delta_k(i) = \begin{cases} 1, & \text{if } i \in \Omega(k) \\ 0, & \text{otherwise} \end{cases}$$

Then we choose the channel that produces the minimum cost function. For each channel that is available for the jth mobile, we therefore choose the one that is available to the least number of mobiles that belong in $N_j(d_{min})$.

iii In the third algorithm all the available frequencies are placed in a queue in ascending order. The dimension of the queue is equal to the number of different available frequencies N. Every time there is a request for a new channel, we assign the frequency that is stored in queue position $i + 1$, where i is the position of the frequency which was the last to be assigned. If the frequency in position $i + 1$ does not satisfy the interference criterion, then we move along the queue into a round robin manner until we find a suitable frequency. If there is no such frequency available, the call is blocked.

In the first two algorithms, if more than one channel produces the same cost function, we arbitrarily choose the one with the smallest index. Finally, if no channel satisfies the necessary criteria, i.e. $\Omega(j) = \emptyset$, then the call is blocked. The general idea behind the first two algorithms is to assign a channel that will produce a 'smoothed' general distribution of channels. Therefore, everytime there is a request for a new channel assignment from the ones that are available for the spe-

cific mobile, i.e. are not used by any other mobile within the interference distance area, we assign the one that is being used the most times by the largest number of mobiles that are located just outside this interference area. The third algorithm does not guarantee a sophisticated frequency arrangement but is of low complexity and produces satisfactory results, which in combination with the annealing process (section 6.6.2) become very powerful.

6.4   Simulation framework

In this section we present the parameters used in our simulation framework and analyse the assumptions made. The parameters are:

- Our main simulation area consists of 49 cells. The size[1] of the cells is given by H which is taken here to be 10Km. The cell pattern is shown in figure 6.1. This 49 cell pattern infinitely repeats itself on both axes. The network is parallelogram shaped and the angle between the axes is 60°. Each cell is identified by its coordinates (i,j), where i,j belong to the set $\{1, 2, ...7\}$. An example of the coordinate repetition is shown in figure 6.1 for the cells that are drawn with dashed lines. Therefore, if one user for example exits the 49 cell pattern from cell (7,3) it re–enters it either from cell (1,2) or cell (1,3).

---

[1]This choice of H gives a cell diameter of approximately 23 Km. This leads to a cell size which is comparable in size to the cell size computed in Chapter 3 based on timing synchronisation considerations. This choice of cell size has the advantage of limiting the simulation run times, as for the same mobile density per cell, a model with a smaller cell size would require a fewer number of events to reach stable conditions. Results for larger cell sizes (like the ones currently using for S-PCN) should be similar to the results described in this chapter, as for the same mobile density per cell the same blocking performance could be achieved by using a larger pool of frequencies.

- The total number of channels used is 35.

- For the comparison purposes with FCA, as well as cell-based DCA algorithms, we use a seven cell frequency reuse pattern applied to a fixed-earth cell satellite system. Specifically, for the FCA model the frequencies are allocated to specific earth cells. Additionally, for the cell-based DCA model when the user is entering a new earth cell, he/she contends for a new frequency to be allocated to the new beam covering the specific earth cell, using a DCA algorithm introduced in [55]. On the other hand, in the case of beam or satellite handovers in earth-fixed cell systems, there is no contention since a user terminal keeps the same channel during the call duration whichever is the serving beam or satellite (the principles of this are also discussed in [59]).

- The average number of users per cell is 20, which results in a total of 980 mobiles.

- Each mobile is moving in a straight line in a random, uniformly distributed direction with a speed that is also uniformly distributed and ranges from stationary to a maximum 100 km/h.

- The duration of the call is a random variable having an exponential distribution with 50 seconds average call holding time.

- The idle time between calls is a random variable having an exponential distribution with the average selected to give the required traffic volume.

Our main assumptions can be summarised as follows:

- The call set up time and the propagation path delays are negligible compared to the call duration time.

- The footprint area is considered to be two dimensional rather than three dimensional. For this reason our criteria of interference is based on the distances of the mobiles on this plane.
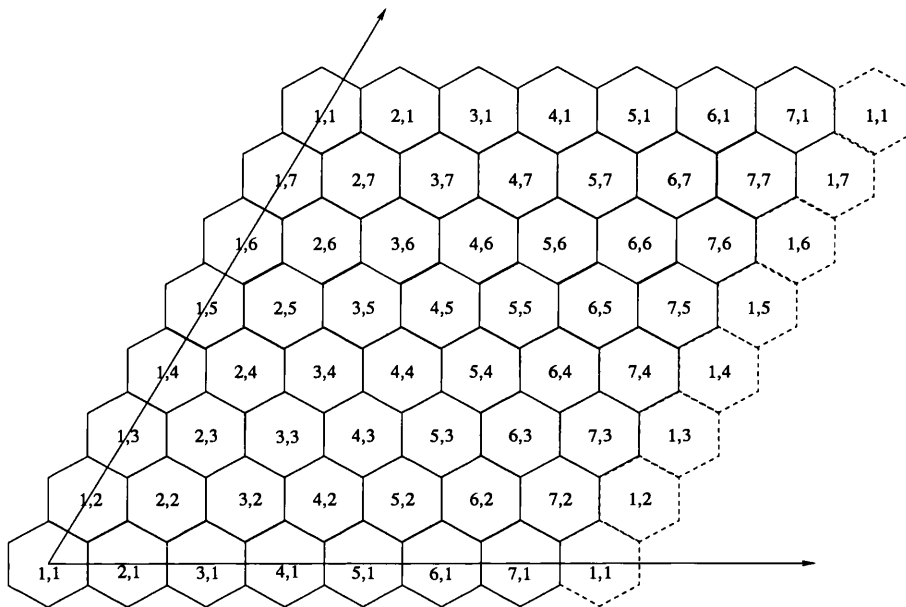
Figure 6.1: The cellular network used in the simulation

6.5 Results and discussion

In order to assess the efficiency of our proposed algorithms, we measure the blocking probability for each. The blocking probability includes all the calls that are dropped, both when there is a request for a new call assignment and also when there is no channel available to satisfy a request when a channel re–allocation is necessary.

In figure 6.2 the results for the system blocking probability for all three algorithms are presented. Specifically, the points marked with '*', 'x' and 'o' illustrate the results for the first, the second and the third algorithms respectively. We observe that the first and third algorithms produce almost similar results although the third algorithm is much simpler to implement and execute. This point will be emphasised in section 6.6.2. However, the best results are produced by the second algorithm.

We further compare the results given by our two best performing cost functions first with an FCA algorithm for a seven frequency reuse pattern, applied in a fixed-earth cell satellite system. We also compare our results with the original algorithm produced by *Del Re et al* in [55]. This algorithm is similar to the one used in equation 6.2. The main difference is that the channels are assigned to cells rather than to mobiles. Additionally, this algorithm is now applied to our fixed-earth cell satellite model.
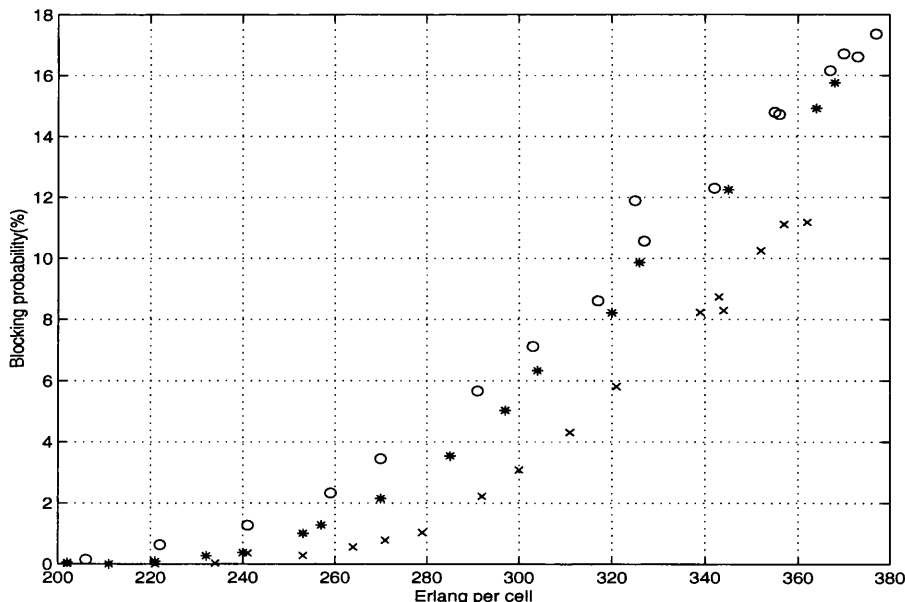
Figure 6.2: Results of all three proposed algorithms (third, first and second algorithm marked with 'o', '*' and 'x' respectively)
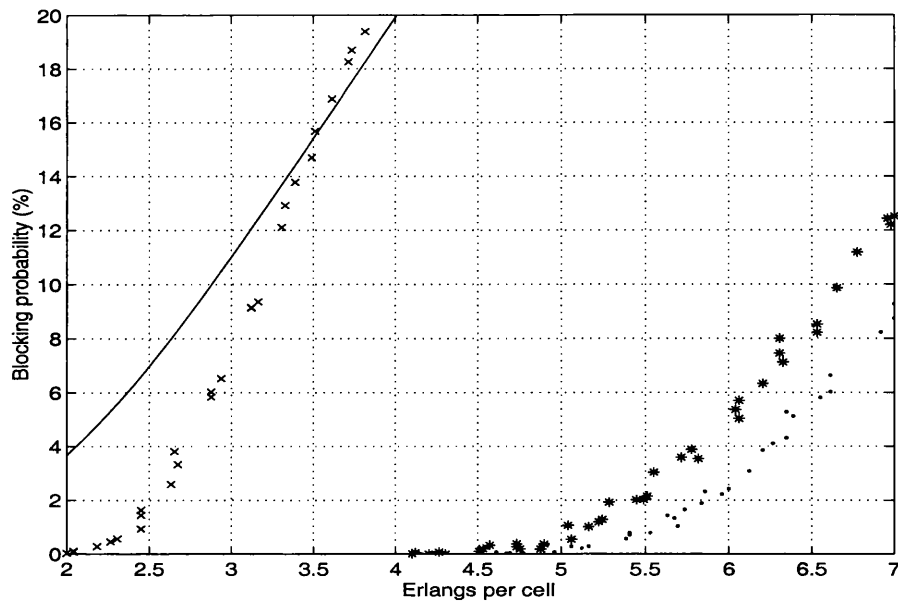
Figure 6.3: Comparison of the two best performing algorithms (first and second algorithm marked with '*' and '.' respectively) with results using a cell-based DCA algorithm from [55] (marked with 'x') and FCA (solid line)

In figure 6.3 we present the results. The solid line shows the results for the FCA algorithm. The results marked with 'x' are generated after using the algorithm found in [55]. Finally the results marked with '*' and '.' are generated using equations 6.1 and 6.2 respectively. We observe a very significant improvement in the system performance by using the mobile distance as a criterion for sharing the available frequencies amongst the users. The traffic that can be accommodated per cell is seen to more than double by the adoption of our algorithms. Also, the slope of the asymptote is lower, implying that this improvement will increase with increasing traffic.

For comparison purposes with a cell-based DCA algorithm, we used Del Re's algorithm ([55]) and applied it to our own model's fixed-earth cell pattern. The particular part of Del Re's algorithm that we extracted from the 'fixed-satellite cell'

model that it was used in, was used in the specific publication only for optimal allocation of channels when there was a request for one and can be applied into any type of cellular network. Therefore, this algorithm was depending on instantaneous snapshots of the frequency spatial distribution across the cells, and not on handover schemes that could be associated with this particular satellite system. We felt that this particular algorithm when applied to our model, having a fixed-earth cell pattern, could be used as a reference point to calibrate our own initial results against a cell-based resource allocation technique. In both cases, there is a one-to-one relationship between channel and user, a DCA algorithm is used and the cost function of allocating channels uses similar principles. One of the main differences was that the channels in our case were allocated to mobiles rather than to cells.

The particular model used by Del Re et al in [55] has evolved since the date of the particular publication and variations of it have been extensively used in the literature. We selectively mention some of these DCA pieces of work below. Specifically, in [57] the LEO fixed-satellite cell model is enhanced with intelligent handover prioritisation schemes. Additionally, it is compared with a GEO type of model and it outperforms it as far as the traffic density capabilities are concerned. In [61] a comparison of the resource management efficiency of a stand-alone terrestrial network with an integrated terrestrial-(LEO)satellite one is performed. The integrated solution outperforms the stand-alone one for all the allocation strategies (FCA-type and DCA-type) used in the specific publication. In [62] the algorithm initially used in [55] is now applied to a S-UMTS LEO fixed-satellite cell system, in order to assess channel allocation strategies in a 3rd generation environment.

All the publications mentioned above make use of variations of the specific algorithm used in [55] as they focus on voice centric DCA satellite networks. In [60] and in [26] packet centric satellite networks with dynamic resource allocation mechanisms are examined. Specifically, in [26] the advantage of dynamically sharing resources amongst users on a per packet basis in a packet-switched broadband satellite network, while satisfying interference constraint criteria, is addressed. The model presented in this paper uses some of the basic principles that Octopus was planned to adopt, but the feature of hopping on a per packet-basis, is introduced in order to tackle the burstiness of a packet-switched broadband system.

In a recent publication ([58]) DCA algorithms are enhanced with a channel reshuffling scheme. The latter is activated when an incoming call or a handover request cannot be accommodated. The authors conclude that such an enhancement offers a significant performance improvement in interference limited satellite networks. This concept of enhancing DCA algorithms with some form of external heuristic is similar to the principles behind the model described in this chapter. Specifically, in the next section we will introduce a novel variation of the simulated annealing technique to reshuffle the channels at fixed time intervals, in order to even out the distribution of the channels across an interference limited network. An important point to be made regarding the heuristic technique that will be introduced in the next section is that it is flexible enough to be used in cell-based as well as in non cell-based systems (as is the case with our model). When applied to a cell-based system, the distance parameters refer to distances between cells.

Therefore, in the sections below we introduce techniques to further improve

our system blocking probability and we compare the new results with our initial

algorithm's ones.

6.6 Simulated Annealing for Dynamic Problems

6.6.1 Concept description

In this section we will present the application of simulated annealing to a 'dynamic' problem in order to improve the efficiency of the channel allocation process and to thereby further reduce the blocking probability.

The main difficulty encountered is the selection of a suitable cooling schedule to be used for the optimisation of the evolving system. One approach would be to fully anneal for every different 'micro–state' of our system. However, this would be extremely time consuming. Furthermore, since there is no guarantee of consistency between successively optimised states, the number of simultaneous frequency transitions required between them can become prohibitive in terms of the required signalling. Therefore, this approach was rejected and another solution is required to continuously optimise the evolving system over time.

One extreme cooling schedule that has been used for static systems [56] is to use a constant temperature that is low enough so that local optima are visited and high enough to climb out of them. This approach is suitable for our system, since it can guarantee an upper limit to the number of channel transitions that will occur between successive states.

The cost function that was used for the annealing process was the following:

$$C(n) = \sum_i \sum_j u(i,j),$$

$$u(i,j) = \begin{cases} \kappa, & \text{if } d_{ij} < d_{min} \\ d_{ij}, & \text{if } d_{ij} \geq d_{min} \end{cases} \tag{6.3}$$

where κ is a constant with $\kappa \gg d_{max}$, N is the number of active mobiles, d_{ij} is the distance between any (i,j) pair of active mobiles, d_{max} is their maximum distance possible and d_{min} is the minimum interference distance calculated from equation 5.1.

The above cost function is similar to the one used in section 5.5. The novel idea of penalising all the active mobiles that are closer than the minimum interference distance, by having them contributing a disproportionately large 'weight' to the cost function, led to a very good system performance and produced the best results amongst all the other cost functions for a time invariant system. After a number of trials with different cost functions, this particular one produced the best results for the time variant system as well. The optimum value of n in equation 5.8 for this specific problem conditions was chosen to be -1.

6.6.2　Application and discussion

We have applied the process of simulated annealing with constant temperature to the system presented in section 6.3, in order to further improve the blocking probability exhibited by our system. The general simulation state diagram is pre-

sented in figure 6.4. Specifically, the channel assignment for new call requests or the reassignment of resources in mobiles that present an increased level of interference is decided based on one of the three DCA algorithms presented in section 6.3. Additionally, at every fixed system timestep T_p we apply simulated annealing (cost function 6.3) to our system and perform an upper limit of M channel transitions in order to 'smoothen' the assigned channels distribution. The timestep T_p is equal to:

$$T_p = K \times T_b,$$

where K is a multiple of system timing blocks (e.g. GSM blocks) and T_b is the block duration.
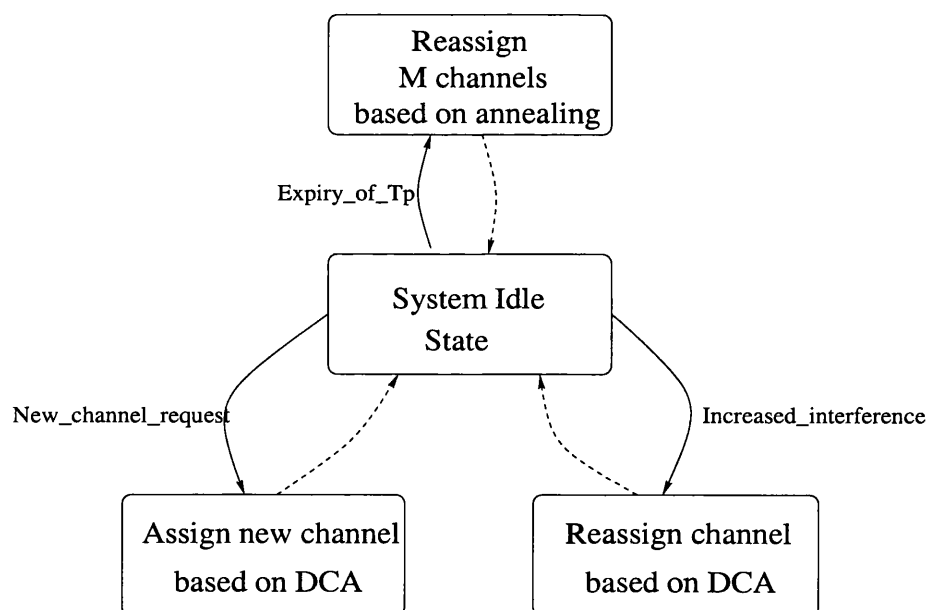
Figure 6.4: System general state diagram

i: $M = 1, K = 1$

We first started our simulations by allowing a maximum of one channel transition per timing block. Some preliminary studies indicated that higher values of M when $K = 1$ may give better results, but we wanted to limit the signalling to acceptable levels.

After examining the results, we observed that the improvement was significant. Specifically, in figure 6.5 the results marked with '*' illustrate the outcome of the first DCA algorithm from section 6.3, while the results marked with 'o' show the outcome of the same algorithm after the annealing process.
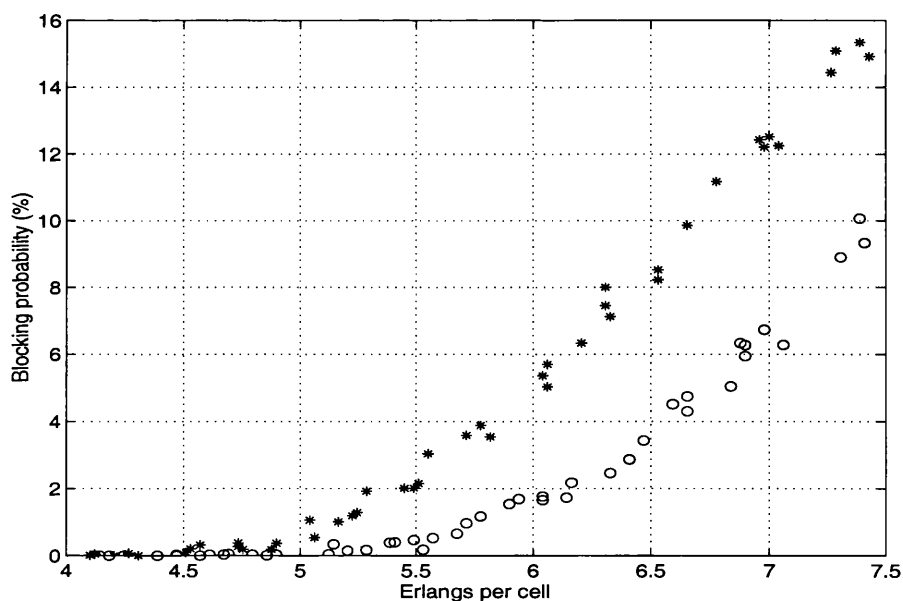
Figure 6.5: Results from the first algorithm with (marked with 'o') and without (marked with '*') the effects of annealing

Similarly, in figure 6.6 we present the results following the adoption of the second algorithm from section 6.3 with the '.' marks, while the '+' marks show the effect of annealing on this particular algorithm. Finally, in figure 6.7 we present the results
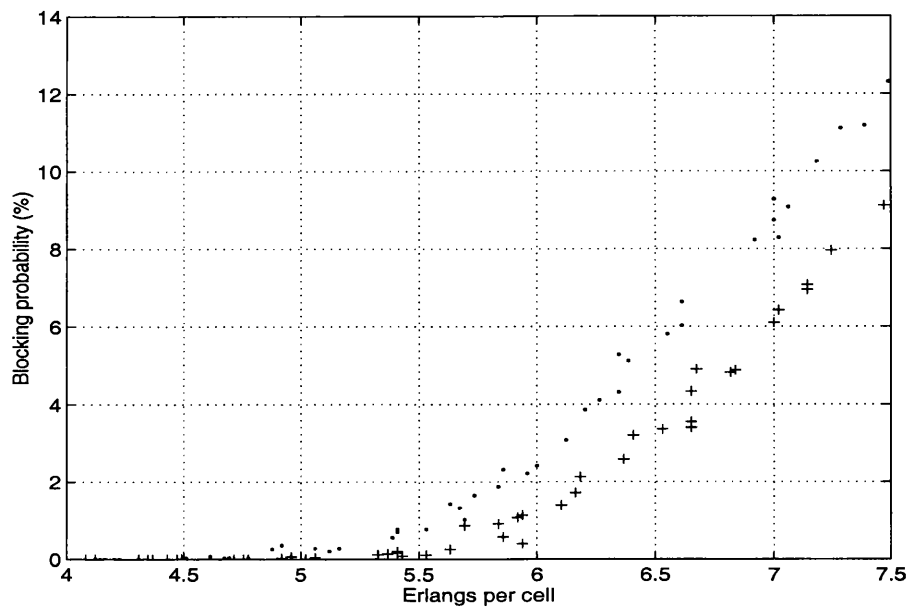
Figure 6.6: Results from the second algorithm with (marked with '+') and without (marked with '.') the effects of annealing
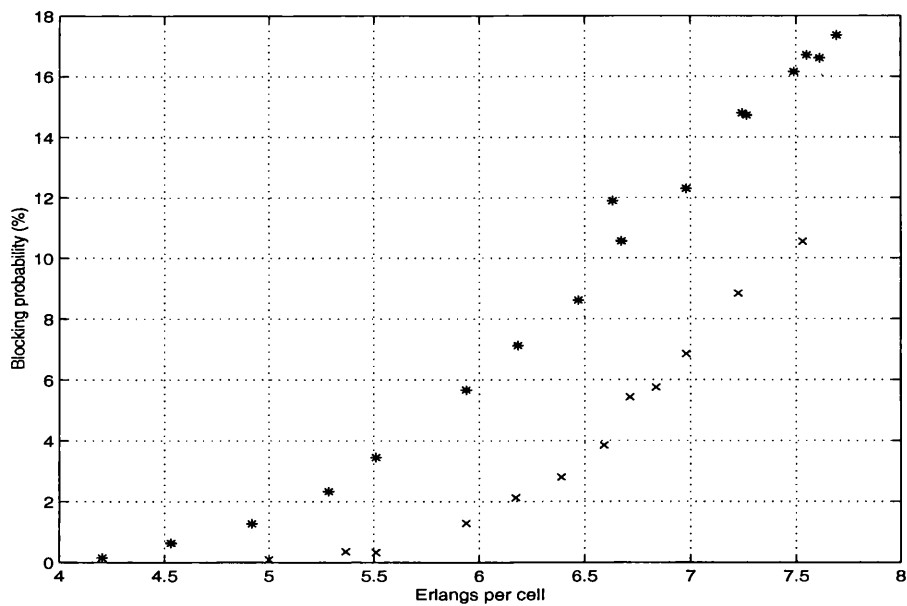
Figure 6.7: Results from the third algorithm with (marked with 'x') and without (marked with '*') the effects of annealing

following the adoption of the third algorithm from section 6.3 with the '*' marks, while the 'x' marks show the effect of annealing on this particular algorithm.

The use of the simulated annealing process produces results that converge for all
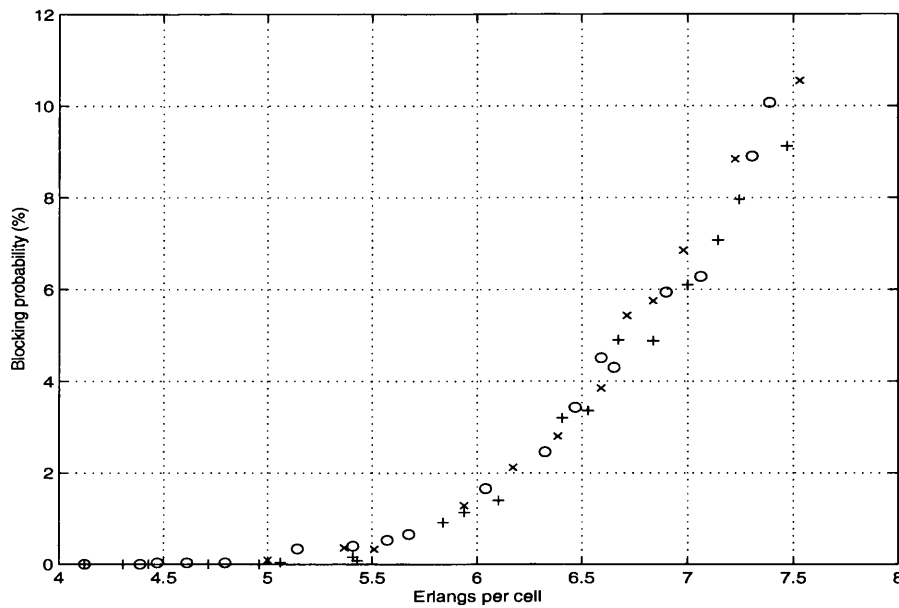
Figure 6.8: Comparison of the results of all three algorithms (first second and third algorithm marked with 'o', '+' and 'x' respectively) including the effects of annealing

three cost functions. This is demonstrated in figure 6.8, which shows the results of the three algorithms, following the application of annealing, with the same symbols as before. It appears therefore, that the choice of a particular algorithm makes no significant difference if we use annealing throughout the process, making the use of the third, simpler, algorithm more attractive.

ii: $\frac{K}{M} > 1$, $M \geq 1$

In an attempt to minimise the signalling even further we increased K and performed studies for different values of M. These studies were executed for the second and third algorithms of section 6.3, since our target was to evaluate the performance of the most complex and the simplest algorithms under the same conditions. It was observed that the results were similar for every pair of M and K that had an integer ratio. In figures 6.9 and 6.10 we present the results for the second and the third

algorithms respectively. The results marked with '$*$' are for $(K = 1, M = 1)$ for the system without the effect of annealing respectively, while the results marked with 'x' do not include the effect of annealing. They are both included for comparison purposes. The results marked with '$+$' and 'o' are for $\frac{K}{M} = 2$ and $\frac{K}{M} = 5$ respectively. It can be observed from both figures that the system performance is degrading when $\frac{K}{M}$ increases. In both cases it approaches the system performance obtained without the effect of annealing. The degradation is more distinct in figure 6.10.
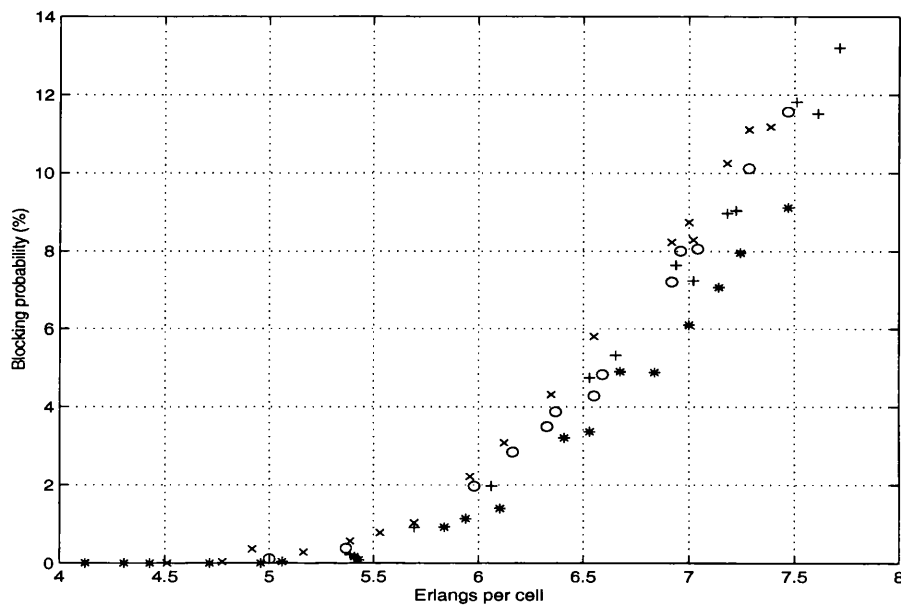
Figure 6.9: The blocking probability for the second algorithm from section 6.3 without annealing ('x') and with annealing for $\frac{K}{M} = 5$ ('o'), $\frac{K}{M} = 2$ ('$+$') and $(K = 1, M = 1)$ ('$*$').

iii: $M > 1, K > 1, \frac{M}{K} = 1$

Since the results from subsection ii were not totally satisfying the next attempt for minimising the signalling was to maintain $\frac{M}{K} = 1$ and increase both M and K. The main idea was to employ some kind of broadcast message to simultaneously inform more than one user of the reassignments every T_p seconds, instead of sending
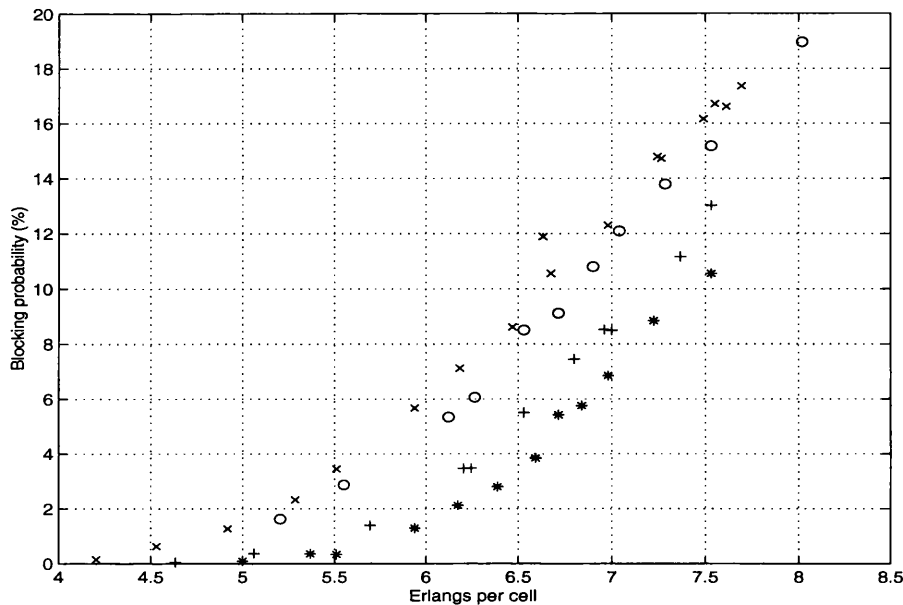
Figure 6.10: The blocking probability for the third algorithm from section 6.3 without annealing ('x') and with annealing for $\frac{K}{M} = 5$ ('o'), $\frac{K}{M} = 2$ ('+') and ($K = 1, M = 1$) ('*').

the reassignment message to single mobiles every T_b seconds. The results not only decreased the amount of signalling, but also slightly improved the system performance. Specifically, figures 6.11 and 6.12 illustrate the results for the second and third algorithms respectively. The results marked with '*' are for ($K = 1, M = 1$) and are again included for comparison purposes. The results marked with '+' and 'o' are for ($K = 3, M = 3$) and ($K = 7, M = 7$) respectively.

iv: $\frac{K}{M} < 1$

We then increased M so that $\frac{K}{M} < 1$. In that way we could perform more reassignments but still use a broadcast message for informing the users of the changes. Figures 6.13 and 6.14 illustrate the results for the second and the third algorithms respectively. The results marked with '*' are for ($K = 1, M = 1$) and are again included for comparison purposes. The results marked with 'o' and '+' are for

Figure 6.11: The blocking probability for the second algorithm from section 6.3 for ($K = 7, M = 7$) ('o'), ($K = 3, M = 3$) ('+') and ($K = 1, M = 1$) ('*').
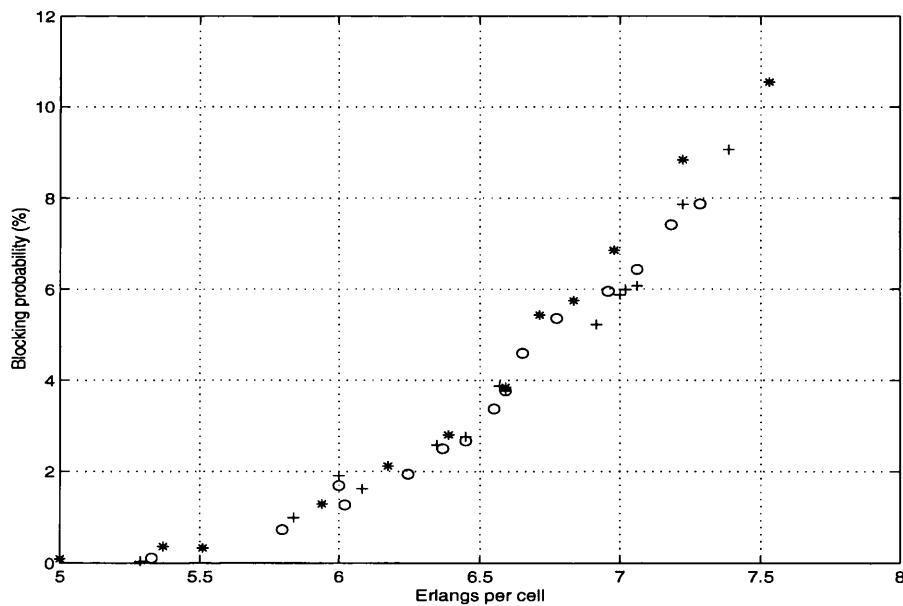
Figure 6.12: The blocking probability for the third algorithm from section 6.3 for ($K = 7, M = 7$) ('o'), ($K = 3, M = 3$) ('+') and ($K = 1, M = 1$) ('*').

($K = 4, M = 15$) and ($K = 10, M = 40$) respectively. As expected the system performance is improving when $\frac{K}{M} < 1$. The improvement is more obvious in figure 6.14.
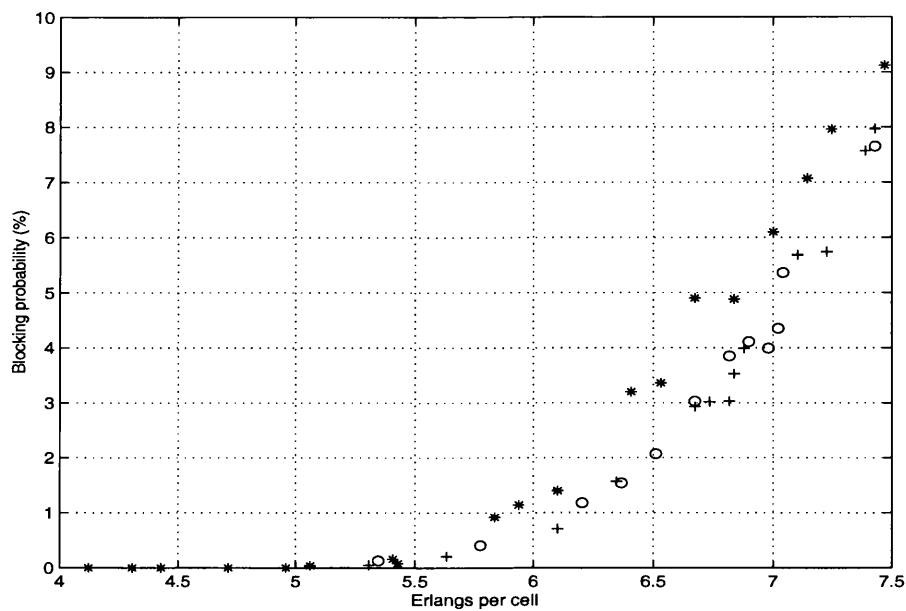
Figure 6.13: The blocking probability for the second algorithm from section 6.3 for ($K = 4, M = 15$) ('o'), ($K = 10, M = 40$) ('+') and ($K = 1, M = 1$) ('*').
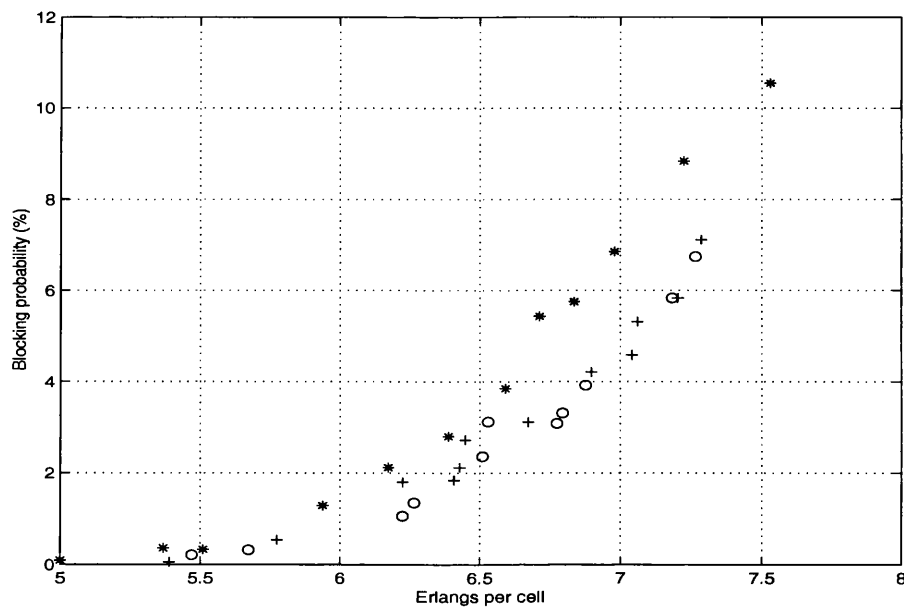
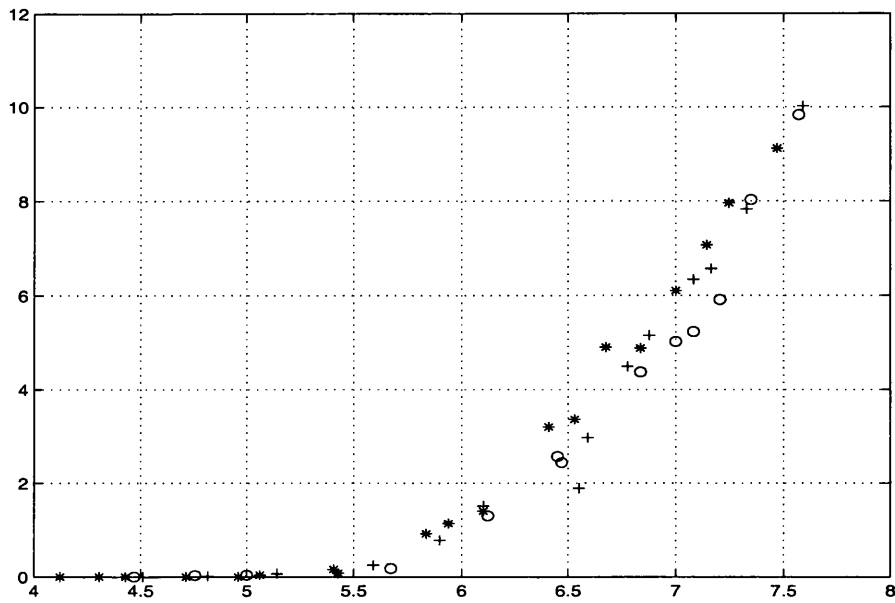Figure 6.14: The blocking probability for the third algorithm from section 6.3 for ($K = 4, M = 15$) ('o'), ($K = 10, M = 40$) ('+') and ($K = 1, M = 1$) ('*').

A conclusion that can be drawn from all the above trials is that the third simple

algorithm performed equally well, or even better, when compared to the complex

one, in cases i, iii and iv, which were actually the cases where an improvement in

system performance was achieved.

A general comment that can be made regarding the specific simulated annealing algorithm and its combination with DCA systems, is that it is applicable to both cell-based as well as non cell-based systems (as is the case with our model). When applied to a cell-based system, the distance parameters refer to distances between cells. We would expect the application of this algorithm to a cell-based system to offer similar levels of improvement as observed in this section. Therefore, further investigation of the above option is suggested as further work.

Before concluding with this section, it is appropriate to comment on some of the practical implications of the techniques discussed in this chapter. The model used in this research introduces complexity implications when compared with a conventional multibeam system. Although it provides a power advantage since less power is required per channel, it also requires a very complex beam-forming network due to the requirements of control for the switched beams, the dynamic frequency-to-beam assignment and the number of beams themselves. This would require, amongst others, extensive look-up tables or the use of analytical techniques for modelling the phase of each beam, on-board switchable filter banks or processors, as well as advanced packaging technologies to simultaneously achieve light weight, low volume and low cost. These practical implications will have to be outweighed against the advantages in order to assess whether such a system would be cost effective for a particular network design. Additionally, as very accurate beam pointing capabilities are required, sophisticated and very accurate location terminal capabilities have to be deployed. The sensitivity required by the beam control systems is greater for

MEO networks than for LEO networks, which might make these approaches more applicable to LEO rather than MEO satellite networks.

An important issue is the location of the controlling software running the optimisation and resource allocation algorithms. An essential requirement is that the controlling software has to have a view of the current resource allocation on a network-wide or satellite-neighbourhood basis. A first option would be to locate this on central points such as earth stations. This would reduce the power required on board as well as the software complexity, although it might have the disadvantage of long propagation delays. Alternatively, the software could reside on each satellite, in which case either intersatellite links would be required to enable the satellites to exchange frequency reuse pattern information, or the satellites could exchange the resource management information through the earth stations.

6.7 Conclusions

In this chapter we have presented DCA channel assignment strategies for mobile satellite communication systems. One of the main novelties of the algorithms lies in their use of interference between the mobiles rather than between cells for the allocation of resources, exactly as in chapter 5. This way, a better 'packing' of channels can be obtained and consequently, the blocking probability of the system can be significantly reduced. Furthermore, a method of incorporating the technique of simulated annealing into the algorithms used in our dynamic model was devised. Although simulated annealing is conventionally used in 'static' problems in order

to approach their global optimum, we applied it to a dynamic problem by using a constant temperature for all the 'micro–states' of our system. Several scenarios were examined in which we were changing the rate and the maximum number of allowed transitions for each simulated annealing step. The goal was to identify the scenarios under which, signalling overheads that might be introduced due to the outcome of the annealing process, are minimised without, however, compromising the system performance.

The results achieved through the use of the DCA techniques led to a better blocking probability system performance than what could be achieved in 'traditional', both FCA and DCA-based, systems. When simulated annealing was introduced, the system performance improved even further. Additionally, the introduction of the simulated annealing allowed the use of very low complexity DCA algorithms, without imposing any performance degradation at all.

The use of our algorithms together with constant temperature annealing can therefore offer major system capacity benefits over existing algorithms. However, some practical implications, as highlighted in this chapter, have to be taken into consideration. These trade-offs have to be analysed on a case-by-case basis. Such an analysis is suggested for further work.

The following chapter brings the thesis to an end by reviewing the main contributions and identifying areas for further work.

Chapter 7

Conclusions

In this thesis timing synchronisation and resource allocation issues in future mobile satellite systems have been examined. Specifically, in the third chapter the timing synchronisation problem generated by the integration of a GSM-like system and a satellite one was addressed. In the fourth chapter the RLC/MAC layer of the GPRS protocol, which provides for resource allocation amongst packet mobile users, was examined in a MEO satellite environment. In chapter five, resource allocation algorithms were examined under different perspectives and were addressed through a combination of DCA techniques and heuristic algorithms. In the sixth chapter the process introduced in chapter five was adjusted to fit a time-variant system.

7.1   Main Contributions

The main contributions of this thesis are as follows:

In the third chapter:

- A novel approach, based on the division of the satellite footprint into determined patterns, was introduced in order to find solutions to the timing synchronisation problems for an integrated GSM and satellite environment. This is considered to be an important contribution from a theoretical point of view, since it addresses and aids the research in one of the most important problems of such a future integrated system. The research work treated the problem as a theoretical approach that could form the basis of a feasibility study, and some implications such as the ones affecting the satellite architecture and system capacity were also discussed. Detailed study of the impact of these implications necessitates further work, and therefore they are discussed further in the section 'Suggestions for further work'.

- In addition to the application of the above approach to a 'fixed-satellite cell' satellite system, the approach was applied to 'fixed-earth cell' satellite systems, emphasising the futuristic nature of such a system. For the latter case we concluded that the system is unfeasible, and showed that the former approach is more practical, while highlighting some of the main design considerations.

In the fourth chapter:

- A detailed simulation platform was developed in order to evaluate the performance of the GPRS RLC/MAC protocol over a MEO satellite system. GPRS is expected to address the mobile world's increased demand for the support of packet-switched services in a mobile environment. The integration of such a protocol with the satellite world could satisfy this demand in a seamless way.

The specific research evaluated the performance of this system and identified

areas in which further enhancements could be introduced.

In the fifth chapter:

- In a limited spectrum mobile world, an efficient distribution of channel re-

  sources amongst the users is of high importance. In this chapter, algorithms

  that enable an optimised 'packing' of channels were identified. Additionally,

  the algorithms were used in conjunction with a novel approach according to

  which the channels were distributed to the users and not to cells. As a result,

  significant performance enhancements were achieved.

- Further investigation, included the identification of the most efficient heuristic

  technique to be used in such environments. Therefore, in this chapter specific

  optimisation techniques were tested for their applicability to the problem at

  hand, as well as different algorithms for their efficiency. These methods and

  algorithms were applied in the next chapter for evaluation in time variant

  systems.

In the sixth chapter:

- The techniques that were used in the fifth chapter were applied to a dynamic

  system, in order to achieve a real-time optimised system performance. The

  simulated annealing technique was incorporated into the algorithms used in

  our dynamic model. Although simulated annealing is conventionally used in

  'static' problems in order to approach their global optimum, we applied it

to a dynamic problem by using a constant temperature for all the 'micro-states' of our system. The results achieved after the introduction of simulated annealing were significantly improved. The use of our algorithms together with constant temperature annealing can therefore offer major system capacity benefits over existing algorithms. We expect that our model of constant temperature annealing, in particular, would offer major benefits also when applied to cell-based DCA systems. We expect that it would improve the system blocking probability while allowing the use of low complexity algorithms for the DCA part. Therefore, this option is further discussed in the 'Suggestions for further work' section. Finally, this chapter also discussed some practical implications of such an approach on satellite and system architecture, as well as the complexity of the controlling hardware and software.

7.2 Suggestions for further work

- In the third chapter we are proposing a new footprint cell-division method in order to mitigate GSM timing synchronisation problems over a satellite system. The cell sizes that are proposed could be small. Therefore there is a trade-off between cell sizes and TDMA timeslot duration. The decrease in cell size will result in an increase in signalling messages required, for location area updates, routing area updates as well as for handovers. The increase in signalling with respect to the cell size should be explored further. Additionally, since the application of this technique to different channel environments impacts on the

satellite architecture and system design, a feasibility study would be suggested. Such a study would compare the application of this theoretical approach to different architectures and investigate cost and practical trade-offs. Finally, cell structures other than the 'traditional' hexagonal cell patterns, could be looked into, in order to attain better suitability of the model. Some suggestions for such patterns are included in the thesis. Complexity issues when looking at such new cell patterns could include appropriate frequency re-use patterns.

- In the fourth chapter the same research could be applied to different altitude constellation systems, such as LEO, and the impact of the altitude on the GPRS system loads that can be supported in this new environment should be evaluated.

- In the fifth and sixth chapters, different optimisation techniques, such as exhaustive search, that could produce more optimal results, but could also introduce significant increase in the execution time and processing capacity required can be applied. Their efficiency could be compared with that of the heuristics used in this work, and then a comparison on the basis of the two previously mentioned limiting factors should take place. Additionally, Genetic Algorithms (GA) techniques are becoming increasingly popular and have proved to offer very efficient alternatives to other optimisation techniques. A further work item could be to evaluate the applicability of GA algorithms to the problem presented in chapters 5 and 6.

- In the fifth and sixth chapters, a further step that can be implemented is to

estimate the average system interference between mobiles with respect to the level of 'packing' that has been achieved through the optimisation process.

- When simulated annealing is used in the time variant system, different simulation sets were executed in order to achieve a balance between improved performance and number of signalling messages sent. A more detailed study in this direction could follow, resulting in a system design which would take into account detailed signalling messages, and the estimation of extra capacity required.

- The system in chapter 6 requires a very complex beam-forming network due to the requirements of control for the switched beams, the dynamic frequency-to-beam assignment and the number of beams themselves, while also requiring advanced packaging technologies to simultaneously achieve light weight, low volume and low cost satellite systems. These practical considerations will have to be outweighed against the advantages that such a system would introduce, in order to assess whether such a system would be cost effective for a particular network design.

- The idea of channel reshuffling between resource allocation events in a DCA system using our proposed, in chapter 6, constant temperature simulated annealing algorithm, could be applicable to both cell-based as well as non cell-based systems (as is the case with our model). When applied to a cell-based system, the distance parameters refer to distances between cells. We would expect the application of this algorithm to a DCA cell-based system to offer

similar levels of improvement as observed in chapter 6. Therefore, further investigation of the above option is suggested as further work.

In summary, this thesis has addressed a range of issues of importance to satellite communication systems and has presented and evaluated a number of innovative solutions.

Bibliography

[1] Eberspacher J., and Vogel H. J., "GSM Switching, Services and Protocols", John Wiley & Sons Ltd., 1999.

[2] Rappaport T. S., "Wireless Communications/ Principles and Practices", Prentice Hall PTR, 1996.

[3] Feldmann M., Rissen J. P., "GSM Network Systems and Overall System Integration", *Electrical Communication*, pp.141-154, 2nd Quarter 1993.

[4] Varin J., Bezler M., Hofmans R., Van den Bossche K., "GSM Base Station System", *Electrical Communication*, pp. 115-163, , 2nd Quarter 1993.

[5] Schmid E. H., Kahler M., "GSM Operation and Maintenance", *Electrical Communication*, pp. 164-171, 2nd Quarter 1993.

[6] No.245 Colloquium (Digest), pp. 37-46, 1994.

[7] GSM 03.64, "Digital cellular telecommunications system (Phase 2+); Stage 2 Service description of GPRS".

[8] Vanttinen V., "The role of mobile satellites in extending terrestrial cellular networks", *Helsinki University of Technology, Department of Computer Science and Engineering*, MSC Thesis, 1998.

[9] GSM 03.60, "Digital cellular telecommunications system (Phase 2+); Overall description of the GPRS radio interface; Stage 2".

[10] Turina D., Ludwig R., "Link Layer Analysis of the General Packet Radio Service for GSM", *ICUPC'97*, Volume 2, pp 525-530, 1997.

[11] Brugel E. W., "On-board processing and future satellite communication services", *Space Communications*, vol.12, iss.3-4, pp.121-174, 1994.

[12] Abrishamkar F., Siveski Z., "PCS global mobile satellites", *IEEE Communications magazine*, pp.132-136, September 1996.

[13] www.ico.com/system

[14] Argenti F., Cappelletti L., Del Re E., and Ferro A., "Integration of satellites into GSM:Signalling flow analysis", *IEEE International Conference on Universal Personal Communications*, Vol.2, pp.755-759, Cambridge, Sep29-Oct2 1996.

[15] Del Re E. and Iannucci P., "The GSM procedures in an integrated cellular/satellite system", *IEEE Journal on Selected Areas in Communications*, Vol.13, No.2, pp.421-430, 1995.

[16] Zhao W., Tafazolli R. and Evans B.G., "Positioning assisted inter-segment handover for integrated GSM-satellite mobile system", *IEE conference publication No.424*, pp.124-128, May 13-15, London, 1996.

[17] Zhao W., Tafazolli R. and Evans B.G., "An integrated system architecture for GSM and Satellite PCN", *IEE Colloquium*, pp.3/1-3/9, London, 1997.

[18] Zhao W., Tafazolli R. and Evans B.G., "Inter-network Signalling Load Analysis for GSM-satellite Integration", *Second European Workshop on Mobile/Personal Satcoms (EMPS'96)*, pp.97-107, October 9-11, Rome, 1996.

[19] Zhao W., Tafazolli R. and Evans B.G., "System Architecture for GSM-Satellite Integration", *The Seventh IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC'96)*, pp.953-957, October, Tapei, Taiwan, 1996.

[20] Restrepo J, Maral G., "Constellation sizing for non-GEO 'earth-fixed cell' satellite systems", *AIAA Int. Com. Sys. Conf.*, pp.768-778, 1996.

[21] Wyatt-Millington R.A., Sheriff R.E., Mort R.J. et al, "Performance Evaluation of Satellite-UMTS Trials over Integrated Testbeds", *Proceedings of the IST Mobile Communications Summit 2000*, pp. 771-776, Galway, Ireland, 1-4 October 2000.

[22] 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects, Iu Principles, *3GPP TR 23.930*

[23] www.infowin.org/ACTS/RUS/PROJECTS/prtiti.htm

[24] www.estec.esa.nl/xewww/cost227

[25] www.infowin.org/ACTS/ANALYSYS/CONCERTATION/MOBILITY/raceproj.htm

[26] Mokhtar A., Azizoglu M., "On the Downlink Throughput of a Broadband LEO

Satellite Network with Hopping Beams", *IEEE Communications Letters*, Vol.4,

No.12, December 2000.

[27] Sherman K.N., "Phased array shaped multi-beam optimisation for LEO satel-

lite communications using a genetic algorithm" *IEEE International Conference*

on Phased Array Systems and Technology, pp.501-504, 2000.

[28] Ovum reports:"Third Generation Mobile Systems",Ovum 1997.

[29] Itkonen J., "Satellite based mobile communications",

Nokia Internal report, November 1998.

[30] He X., Tafazolli R., Evans B. G., "An optimised signalling architecture for a

MEO/ICO satellite system", *AIAA Int. Com. Sat. Sys. Conf.*, pp. 443-450,

1996.

[31] Taaghol P., Tafazolli R., "GPRS for mobile satellite communication systems",

AIAA Int. Com. Sat. Sys. Conf., 1996.

[32] Werner M., Jahn A., Lutz E., Bottcher A., "Analysis of system parameters for

LEO/ICO-satellite communication networks", *J. on Sel. Areas in Com.*, 1995,

13, (2), pp. 371-381

[33] Johannsen K. G., "Mobile p-service satellite system comparison", *Int. J. of Sat. Com.*, **13**, (12), pp. 453-471, 1995.

[34] Mouly M., Pautet M.B., "The GSM system for mobile communications", ISBN 2-9507190-0-7, France, 1992

[35] Jamalipour A., "Low earth orbital satellites for personal communication networks", *Artech House*, 1997

[36] www.skypub.com/sights/satellites/iridium.html

[37] www.mil3.com/products/modeler/method.html

[38] Ween A., Qureshi A., Kraetz M., Rossiter M., "Dynamic resource allocation for multi-service packet based LEO satellite communications", *Global Telecommunications Conference* Vol. 5 , pp 2954 -2959 , 1998.

[39] Einola H., "Satellite Access in Personal Communications Networks", Helsinki University of Technology, Espoo October 17, 1997.

[40] GSM 04.60, "Digital cellular telecommunications system (Phase 2+);General Packet Radio Service (GPRS); Mobile Station (MS) - Base Station System (BSS) interface; Radio Link Control/ Medium Access Control (RLC/MAC) protocol".

[41] Brasche G., Walke B., "Concepts, Services, and Protocols of the new GSM Phase 2+ General Packet Radio Service", *IEEE Communications Magazine*, Vol.35, No.8, August '97.

[42] Cai J., Goodman D. J., "General Packet Radio Service in GSM", *IEEE Communications Magazine*, Vol.35, No.10, October '97.

[43] Reeves C. R., "Introduction" in *C. R. Reeves ed. "Modern Heuristic Techniques for Combinatorial Problems"*, Blackwell Scientific Publications, 1993.

[44] Kirkpatrick S., Gelatt Jr. C. D. and Vecchi M. P., "Optimisation by simulated annealing", *Science*, 220, pp.671–680, 13 May 1983.

[45] Cerny V., "A thermodynamical approach to the travelling salesman problem: An efficient simulation algorithm", *J. Optim. Theory Appl.*, 45, pp.41–51, 1985.

[46] Borndörfer R., Eisenblätter A., Grötschel M. and Martin A., "Frequency assignment in cellular phone networks", *Annals of Operations Research*, vol. 76, pp. 73–93, 1998.

[47] Haas Z. J., Winters J. H. and Johnson D. S., "Simulation results of the capacity of cellular systems", *IEEE Transaction on Vehicular Technology*, vol. 46, no. 4, pp. 805–817, Nov. 1997.

[48] Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., and Teller E., "Equation of state calculations by fast computing machines", *Journal of Chemical Physics*, vol. 21, pp. 1087–1092, June 1953.

[49] Duque-Anton M., Kunz D. and Rüber B., "Channel Assignment for Cellular Radio Using Simulated Annealing", *IEEE Transaction on Vehicular Technology*, vol. 42, no. 1, pp. 14–21, Feb. 1993.

[50] Ortigoza–Guerrero L. and Lara–Rodriguez D., "Dynamic channel assignment strategy for mobile cellular networks based on compact patterns with maximised channel borrowing", *Electronics Letters*, vol. 32, no. 15, pp. 1342–1343, July 1996.

[51] Kwok K. F., Liu P. C. K. and Li K. C., "Comparisons of channel assignment strategies in LEO satellite land mobile communication systems", *IEEE International Conference on Communication Systems and IEEE International Workshop on Intelligent Signal Processing & Communication SYSTEMS*, pp. 1274–1278, Nov. 1996.

[52] Del Re E., Fantacci R. and Giambene G., "Performance comparison of different dynamic channel allocation techniques for mobile satellite systems", *European Transaction on Telecommunications*, vol. 8, no. 6, pp. 609–621, Nov.–Dec. 1997.

[53] Godara L.C., "Applications of the Antenna Arrays to Mobile Communications, Part I: Performance Improvement, Feasibility, and System Considerations", *Proceedings of the IEEE*, Vol.85, No.7, pp.1031-1060, July 1997.

[54] Gebauer T., Gucklich H.G., "Channel-Individual Adaptive Beamforming for Mobile Satellite Communications", *IEEE Journal on Selected areas in Communications*, Vol.13, No.2, February 1995, pp.439-448.

[55] Del Re E., Fantacci R. and Giambene G., "Performance comparison of a dynamic channel allocation technique for satellite mobile cellular networks", *International Journal of Satellite Communications*, Vol.12, pp.25-32, 1994.

[56] Connolly D. T., "An improved annealing scheme for the QAP", *European Journal of Operational Research*, 46, pp. 93–100, 1990.

[57] Del Re E., Fantacci R., and Giambene G., "Efficient Dynamic Channel Allocation Techniques with Handover Queueing for Mobile Satellite Networks", *IEEE Journal on Selected Areas in Communications*, Vol.13, No.2, pp.397-405, February 1995.

[58] Obradovic V., Kanth L. M., "Improvement of the Capacity of the LEO Mobile Satellite Systems Using Channel Reshuffling", *IEEE Global Telecommunications Conference*, pp.1166-1170, 2000.

[59] Boukhatem L., Beylot A. L., et al, "Performance Analysis of Dynamic and Fixed Channel Allocation Techniques in a LEO constellation with an "Earth-Fixed Cell" System", *IEEE Global Telecommunications Conference*, Volume 2, pp1145-1148, 2000.

[60] Del Re E., Fantacci R., et al. " Dynamic Resource Allocation for LEO Satellite Systems with QoS provision for an IPv6 Network", *IEEE Global Telecommunications Conference*, General Conference (Part A), pp.21-25, 1999.

[61] Mohorcic M., Kandus G., et al. " Performance Study of an Integrated Satellite/Terrestrial Mobile Communication System", *International Journal of Satellite Communications* , Vol.14, pp.413-425, 1996.

[62] Hu Y. F., Sheriff R. E., Der Re E., et al. " Satellite UMTS Traffic Dimensioning and Resource Management Technique Analysis", *IEEE Transactions on Vehicular Technology* , Vol.47, No.4, pp.1329-1341, 1998.