UNIVERSITY OF ABLERTA

An Improved MAC Protocol for Reverse Link Data Transmission in Wideband DS-CDMA

By

ROBERT HANG ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

EDMONTON, ALBERTA

FALL 1998

0-612-34373-1

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF THE AUTHOR: Robert Hang

TITLE OF THE THESIS: An Improved MAC Protocol for Reverse Link Data Transmission in Wideband DS-CDMA

DEGREE: Master of Science

YEAR THIS DEGREE GRANTED: 1998

Robert Hang
10 Rue de l'Harmonie
95490 Vauréal
FRANCE

DATE: 1 Oct. 1998

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled *An Improved MAC Protocol for Reverse Link Data Transmission in Wideband DS-CDMA* submitted by Robert Hang in partial fulfillment of the requirements for the degree of Master of Science.

W. Krzymień

Dr. Witold A. Krzymień, Supervisor
Professor of Electrical and Computer Engineering

Dr. Charles L. Despins, Co-supervisor
Associate Professor
INRS-Télécommunications (Université du Québec)

Bruce Cockburn

Dr. Bruce F. Cockburn, Internal Examiner
Associate Professor of Electrical and Computer Engineering

Dr. Pawel Gburzynski, External Examiner
Professor of Computing Science

Date: 1 Oct. 1998

Abstract

This thesis describes a new medium access control (MAC) protocol for packet data transmission on the reverse link of a third generation wideband DS-CDMA system. The protocol is particularly efficient if data transmission involves short packets. The packet data transmission channel (PDTCH) is a time-shared channel used to provide the packet data service. By using the time taken by the receiver to process the previous packet to transmit a new packet, the average packet delay is reduced compared to a conventional MAC scheme. Assuming Poisson distributed arrivals, analysis and computer simulation results consistently show that, for short packets, the new scheme can handle a higher packet arrival rate than the conventional scheme, while achieving a similar throughput. However, this delay improvement decreases as the packet size increases. In the context of a DS-CDMA cellular system, the thesis shows the possibility of the new MAC protocol to have a higher number of packet data users at the expense of having a lower number of voice users compared to a conventional MAC scheme. This flexibility is an asset for network operators who can then dynamically trade off the number of voice users against the number of possible data users.

Dedicated to my parents, sister, brother, and my friend Tetyana for their love, support, and encouragement.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisors, Dr. Witold A. Krzymień and Dr. Charles L. Despins (from INRS-Télécommunications and Microcell Connexions Inc., Montréal, Québec, Canada), for their advice, patience, and support. I am also very thankful to them for the their financial support throughout my graduate program.

I would like to thank Drs. Bruce F. Cockburn and Pawel Gburzynki for reviewing this thesis.

I would like to thank everyone at TRLabs and INRS-Télécommunications for providing such a friendly environment to do research or study. TRLabs and INRS-Télécommunications financial support is also much appreciated.

I would like to thank Songsong Sun for very helpful discussions, and for providing me with some of his Matlab code. I would also like to thank Deepak Sarda and Peng Mok for their in debugging my C programs.

Finally I would like to thank all my family, Lina, Ihor, and Valentyn for their love and support.

Table of Contents

Table of Figures

List of Symbols

λ_{MS}	arrival rate at MSs
$\overline{T}_k$	average delay of the k^{th} packet
N_{ls}	average number of frames needed to successfully receive the leading segment
N_{ts}	average number of frames needed to successfully receive the trailing segment.
$\overline{N}(t)$	average number of packets in the system at time t
$\overline{X_N}$	average packet transmission time of the new MAC scheme
$\overline{\mu}$	average signal to noise ratio over a PN code period
$S(t)$	BPSK modulated signal
f_0	carrier frequency
β	characteristic of pulse shaping filter used
$E_{c,v}^{'}$	chip energy of the PDCCH from the voice users
F_c	chip rate
$E\{X_j \mid N_i\}$	conditional expectation of X_j given N_i
$P_e(u)$	conditional probability of a RS decoder error given u RS symbol errors
C_{N_d}	cost in terms of the number of voice users having N_d data users for the circuit-switched transmission case
$C_{N_d}^{'}$	cost in terms of the number of voice users having N_d data users for the packet-switched transmission case
α_d	data activity factor for data users
α_v	data activity factor for voice users
α	data activity factor of the ARCH/PDTCH channel pair
α_C	data activity factor of the ARCH/PDTCH channel pair under the conventional MAC scheme

α_N	data activity factor of the ARCH/PDTCH channel pair under the new MAC scheme
$x(t)$	data stream
$S_c(t)$	despread signal
$\left(E_b/I_0\right)_d$	energy per bit to the interference power spectral density for packet data users
$\left(E_b/I_0\right)_v$	energy per bit to the interference power spectral density for voice users
$E_{c,ARM}$	energy per chip of ARMs
$E_{c,d}^{'}$	energy per chip of the DPCCH
$E_{c,d}$	energy per chip of the DPDCH
$E_{c,period}$	energy per PN code period
t_c	error correcting capability of the RS code
$\hat{T}$	estimation of the propagation delay
$\overline{X}$	expectation of the random variable X
$r(\tau)$	instantaneous residual service time of a packet at time τ
θ	intercell interference factor
$\lambda_{\max}$	maximum arrival rate tolerable for the ARCH/PDTCH channel pair
$\tilde{N}_d$	maximum number of data users in the circuit-switched transmission case
$\tilde{N}_d^{'}$	maximum number of data users in the packet-switched transmission case
$\tilde{N}_v$	maximum number of voice users in the circuit-switched transmission case
$\tilde{N}_v^{'}$	maximum number of voice users in the packet-switched transmission case
$T_{\max}$	maximum tolerable delay
$\overline{X_C}$	mean packet transmission time of the conventional MAC scheme
n_1	number of bits per codeword of the inner code
k_1	number of data bits per codeword of the inner code

k_2	number of data symbols per codeword of the outer code
$\beta(t)$	number of departures within $[0,t]$
N_k	number of frames that have been transmitted after k TS retransmissions
m	number of MS attached to a ARCH/PDTCH channel pair
$N(t)$	number of packets in the queueing system at time t
$M(t)$	number of packets sent within $[0,t]$
$\alpha(t)$	number of packets which arrived within $[0,t]$
L	number of resolvable paths
n_2	number of symbols per codeword of the outer code
δ	packet departure rate
$\overline{A}$	packet transmission time
A_k	Poisson process
μ	power control degradation factor
I_v	power spectral density of the interference created by voice users
$I_{0,d}$	power spectral density of the interference seen by the DPDCH of a data user
$P(E)$	probability of accepting an erroneous frame
P_{cd}	probability of correct RS decoding
$p_n(t)$	probability of having n packets in the system at time t
P_e	probability of RS decoder error or RS undetected error
P_f	probability of RS decoder failure
G_d	processing gain of data users
G_v	processing gain of voice users
ρ_d	ratio of the data DPCCH chip energy to the data DPDCH chip energy
ρ_v	ratio of the voice DPCCH chip energy to the voice DPDCH chip energy
ρ	roll-off factor of a raised cosine filter
ε	RS symbol-error probability

$\overline{X_C^2}$	second moment of the packet transmission time of the conventional scheme
$\overline{X_N^2}$	second moment of the packet transmission time of the new MAC scheme
$G(t)$	spreading sequence or pseudo-noise sequence
λ	steady-state arrival rate
$\overline{N}$	steady-state average number of packets in the system
T	steady-state average packet delay
p_n	steady-state probability of having n packets in the system
N	steady-state time average or time average of $N(t)$
T	steady-state time average packet delay
$A(t)$	stochastic process taking positive integer values
T_{total}	sum of n_{ts}, n_{pc}, n_s, and n_p times T_f
γ_v	threshold of the SIR that ensures the BER required for voice users
θ_d	threshold used to minimize the synchronizer false acquisition probability
η	throughput efficiency
η_c	throughput efficiency of the conventional MAC scheme
η_N	throughput efficiency of the new MAC scheme
N_t	time average of $N(t)$ up to time t
T_t	time average of packet delay up to time t
λ_t	time average of the arrival rate over the interval $[0,t]$
T_i	time spent in queue by the i^{th} arriving packet
τ	time variable
$S_t(t)$	transmitted signal
b	number of bits per symbol for the non-binary outer code
C_1	binary inner code
C_2	non-binary outer code
E_b/I_o	signal-to-interference ratio

$E_{c,v}$	chip energy of the DPDCH from the voice users
$I_{o,v}$	power spectral density of the interference seen by a voice user
k	number of data symbols per codeword
L	number of short code repetitions in an ARM
n	number of symbols per codeword
N	round trip delay
N_{cl}	constraint length of the convolutional code
N_{code}	length (also called period) of the short code
N_d	number of data users in the system
N_i	number of packets found waiting in queue by the packet no. i upon arrival
n_p	number of frames needed to transmit a packet if no retransmission is required
n_{pc}	average number of frames needed to re-establish closed-loop power control and re-synchronize the PN code after a silent period
n_s	average number of ARMs needed for an MS to acquire synchronization
n_s	average number of ARMs needed to acquire synchronization
N_{slot}	number of slots in the access subframe
n_{ts}	number of frames in the trailing segment
N_v	number of voice users
N_v	number of voice users
P	power of a BPSK modulated signal
P_c	correct acquisition probability
P_{fa}	false acquisition probability
p_k	probability that the TS transmission ends after k retransmission
p_{retx}	probability of codeword retransmission
R	mean residual time
R_i	residual service time seen by the packet no. i
T	mean packet delay
t	particular value of τ
T	propagation delay
T_{access}	duration of the access subframe

T_f duration of a frame

T_k time at which the transmission of packet no. 1 ends after k TS retransmission

T_{slot} duration of a time slot dedicated to a given MS on the ARCH

W mean access delay

W_i waiting time in queue for the packet no. i

X_i service time of the packet no. i

Y steady state average of the sum of MS access periods

Y_i sum of all the MS access periods during which packet no. i must wait before being transmitted

List of Abbreviations

3G	third generation
ACK	acknowledgement
ACTS	Advanced Communications and Technologies Services
AMPS	Advanced Mobile Phone System
ARCH	access request channel
ARIB	Association of Radio Industry Businesses
ARM	access request message
ARQ	automatic repeat request
ATDMA	Advanced Time Division Multiple Access
BCCH	broadcast control channel
BER	bit error rate
BPSK	binary phase shift keying
BS	base station
CCH	common channel
CEPT	Conférence Européenne des Postes et Télécommunications
CLPC	closed-loop power control
CODIT	Code Division Testbed
DCH	dedicated channel
DL	downlink
DLL	data link layer
DPCCH	dedicated physical control channel
DPDCH	dedicated physical data channel
DS-CDMA	direct-sequence code division multiple access
DTX	discontinuous transmission
EC	European Community
ETSI	European Telecommunications Standards Institute
EU	European Union
FACH	forward access channel

FDD	frequency division duplex
FDMA	frequency division multiple access
FEC	forward-error-correction
FER	frame error rate
FM	frequency modulation
FRAMES	Future Radio Wideband Multiple Access System
G	general distribution
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
IMT-2000	International Mobile Telecommunications 2000
IS-54	Interim Standard 54
IS-95	Interim Standard 95
ISDN	Integrated Services Digital Network
ITU	International Telecommunications Union
kbps	kilo bits per second
LAC	link access control
LS	leading segment
M	memoryless distribution
MAC	medium access control
MAI	multiple access interference
Mchips/s	mega chips per second
MS	mobile station
NACK	negative-acknowledgement
NMT	Nordic Mobile Telephone
NTT	Nippon Telephone and Telegraph
OSI	Open System Interconnect
PC	power control
PCH	paging channel
PDC	Pacific Digital Cellular
PDTCH	packet data transmission channel

PDU	protocol data unit
PL	physical layer
PN	pseudo-noise
PRACH	physical random access channel
PSPDN	packet-switched public data network
PSTN	public switched telephone network
RAB	random access burst
RACE	Research for Advanced Communications in Europe
RACH	random access channel
RF	radio frequency
RLC	radio link control
RLP	radio link protocol
RS	Reed-Solomon
RTT	radio transmission technology
RX	receiver
SIR	signal-to-interference ratio
SMS	short message service
TACS	Total Access Communication System
TDD	time division duplex
TDMA	time division multiple access
TIA	Telecommunications Industry Association
TS	trailing segment
TX	transmitter
UL	uplink
UMTS	Universal Mobile Telecommunications System
USDC	United State Digital Cellular
UTRA	UMTS Terrestrial Radio Access

Chapter 1. Introduction

Today, wireless communications is one of the fastest growing fields in telecommunications. It allows users to communicate while they are away from their homes or offices. The success of mobile communications has been better than expected and is an indicator that the need for mobility is becoming more and more important. In the European Union (EU), the growth of cellular subscribers has been steady over the last few years as shown in Figure 1. Also shown in that figure is the penetration rate in the EU. Figure 2 shows the penetration rate by countries in the EU. This figure shows that for Scandinavian countries like Finland, Norway, and Sweden, the penetration rate is greater than 40 percent. In May 1998, the penetration rate of cellular subscribers in the entire EU was still low and was less than 18 percent. It is expected that the subscriber growth in EU will be maintained for several years to reach the penetration rates of Scandinavian countries. In fact, the number of cellular subscribers using the Global System for Mobile Communications (GSM) is expected to reach 100 million by September 1998 and 200 million by the turn of the century worldwide [1]. Japan has also experienced similar growth (as shown by Figure 3), but in an even more dramatic fashion.

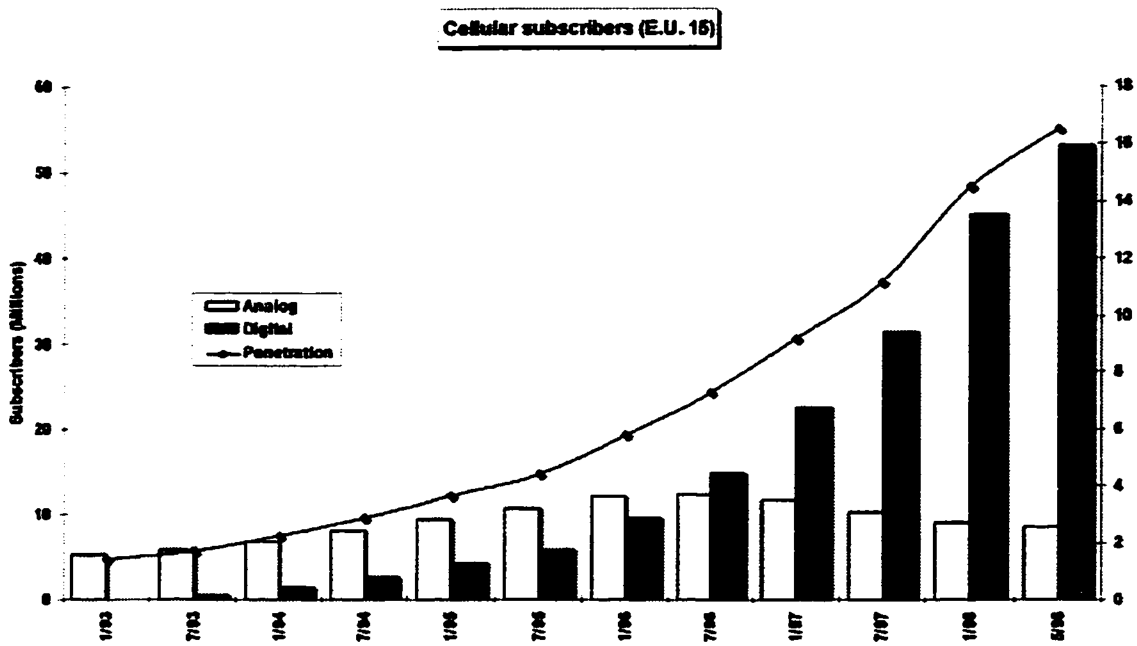

Figure 1. Cellular subscribers growth in the European Union [2].

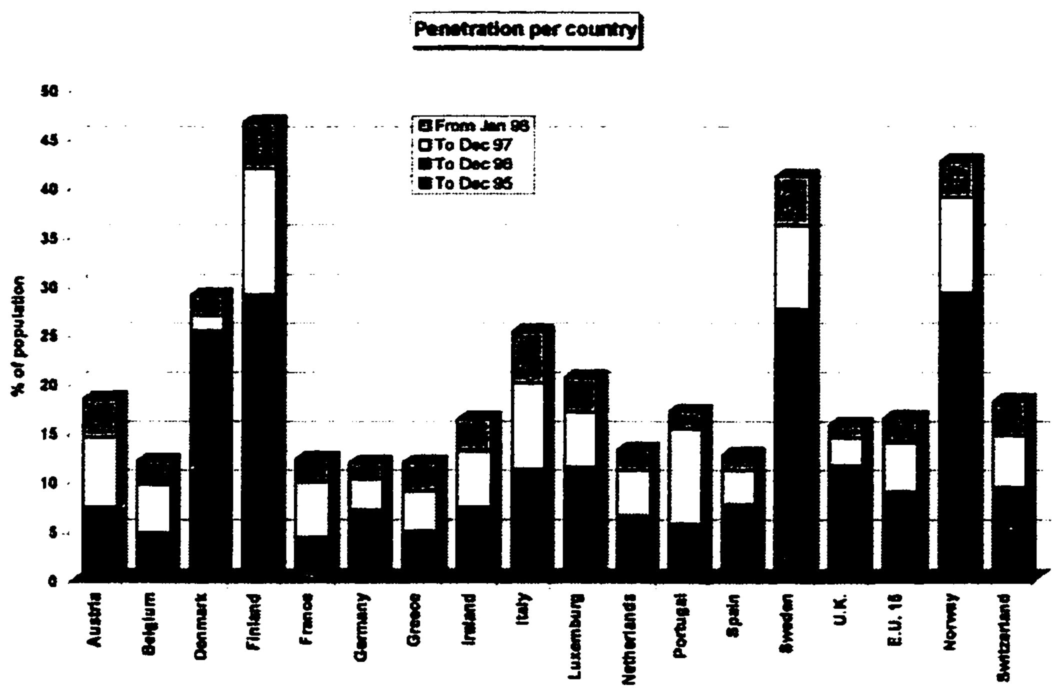

Figure 2. Cellular penetration in Europe [3].

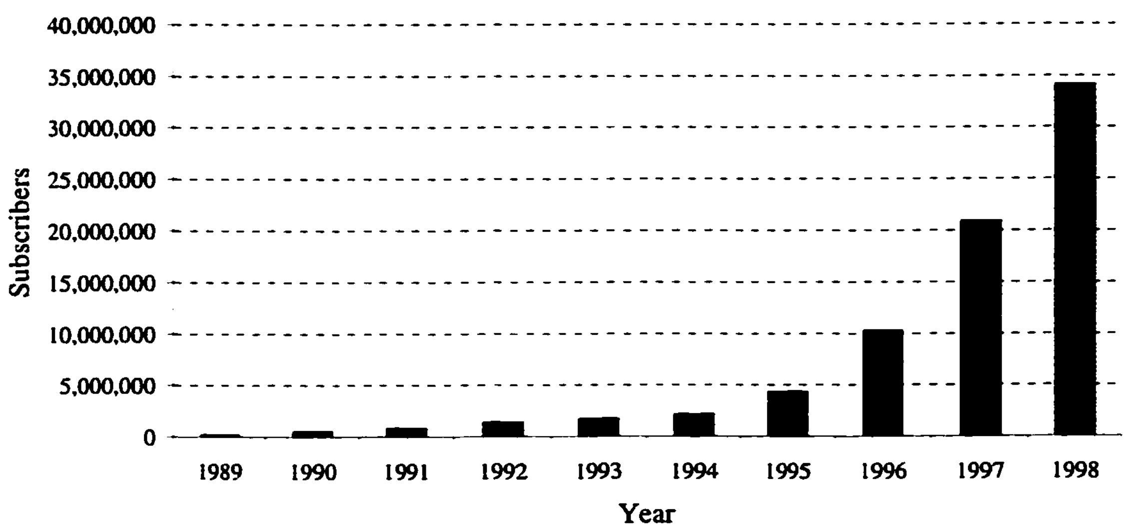

Figure 3. Cellular subscriber growth in Japan [4].

So far, the main application of mobile systems is mobile telephony, but the use of mobile fax and short message applications is becoming more and more popular. Applications like electronic mail (e-mail), Web browsing, or fund transfers are also foreseen [5]. This trend has led to intensive research on efficient ways to transmit data over mobile radio links. It is the purpose of this thesis to investigate means to improve the efficiency of data transmission over a wireless medium. Before details of the project are presented, it is useful to describe the concepts related to cellular mobile communications and a short history of cellular systems.

1.1. The cellular concept

In order to achieve systems with high capacity under the constraint of a fixed amount of (scarce) spectrum, the concept of frequency reuse in a cellular system structure was developed. In a cellular system, the whole service area is divided into *clusters* of small areas. Clusters themselves consist of *cells* of smaller areas (*e.g.* seven cells per cluster). The entire available spectrum is divided into radio frequency (RF) channels (frequency slots). The channels are distributed among all the cells of a given cluster. Cells that belong to the same cluster do not use the same RF channels, while different clusters re-use the same RF channels. The reason for this RF channel reuse is that the distance that separates cells, in different clusters using the same RF channel, is carefully designed so that co-channel interference is minimized. This is called the *frequency reuse* concept because the same spectrum is reused over the service area as many times as the number of clusters. By decreasing the size of the cells or clusters, the capacity of the cellular system can be increased because frequency will be reused more often [6].

In a cellular system, every user is equipped with a mobile station (MS) which communicates with the base station (BS) of the cell. Design parameters of the link between the MS and the BS is what is called the *air interface*. In order to allow communications with other users who are not necessarily mobile, the BS is connected (maybe through other entities) to other networks such as the public switched telephone network (PSTN). It is useful to introduce a few definitions now. The link from the BS to the MS is called *forward link* or *downlink* (DL) while the link from the MS to the BS is called *reverse link* or *uplink* (UL). The uplink and the downlink can be both active at the same time if two different frequency bands are used for each of them. This is called *frequency division duplexing* (FDD) and it allows simultaneous UL and DL connections. If only one frequency band is available for both directions of transmission then this band will be used some fraction of the time by the UL and the other fraction of the time by the DL. This is called *time division multiplexing* (TDD). *Mobile communications* refers to communications involving subscriber stations that can move. *Wireless communications* refers to communications that involve subscriber stations that are stationary or mobile but nevertheless not wired. There are mainly three types of multiple access schemes for the air interface of cellular systems. In *frequency division multiple access* (FDMA), the bandwidth is divided into subbands that are assigned to users. In *time division multiple access* (TDMA), the available bandwidth is divided into time slots dedicated to users. In *code division multiple access* (CDMA), all the users use the entire available bandwidth at the same time. In CDMA users are separated by the code they use. More detail will be given for CDMA in the rest of the thesis.

1.2. First generation cellular systems

In the late seventies and early eighties, the first commercial mobile cellular systems used analog frequency modulation (FM) to carry voice, frequency shift keying for signaling and FDMA as an accessing scheme. These systems were commercialized more or less at the same time. In North America, the first system introduced was the Advanced Mobile Phone System (AMPS) engineered by Bell Labs. AMPS was deployed in 1983. In AMPS, the total available bandwidth is divided into 30kHz channels and each of them is dedicated to one user in a cell. In Europe, different systems were introduced depending on the country. The Total Access Communication System (TACS) derived from AMPS was introduced in 1985 in the United Kingdom. The Nordic Mobile Telephone (NMT) used in Scandinavia was introduced in 1981. Radicom 2000 was introduced in 1985 in France [6]. In Japan, the Nippon Telephone and Telegraph (NTT) Company introduced in 1979 the NTT mobile system. All these analog systems are called first generation cellular mobile systems because of the analog modulation they used to transmit voice. While the North American AMPS system had good success, the European systems were almost totally incompatible, which limited their success in their own countries. Indeed, European users could not *roam* from one system to another. In other words, they could not use their mobile phone when they traveled to another country. This situation was not profitable in terms of equipment sales and could not last in a Europe that was uniting.

1.3. Second generation cellular systems

In 1982, the Conférence Européenne des Postes et Télécommunications (CEPT) set up a study group called Groupe Spécial Mobile (GSM) in charge of designing a new pan-European system to find a remedy to the incompatibility problem of first generation cellular systems deployed in different European countries. The goals of this system were multiple. The unrestricted roaming capability of the new system was one of the first requirements but emphasis was also put on:

- better speech quality with the use of digital voice coders
- better spectrum efficiency

- lower cost
- integration of speech, data, and signaling on the same channel
- compatibility with the existing Integrated Services Digital Network (ISDN)
- modular structure that would allow an easy extension

In order to meet all of these requirements, a natural choice was to use digital transmission. In 1989, GSM gave responsibilities for the standardization process of the new system to the European Telecommunications Standards Institute (ETSI). ETSI published the first recommendations in 1990 and renamed the system the Global System for Mobile Communications (GSM). The access scheme of GSM is a hybrid FDMA/TDMA. Specifically, the available bandwidth is divided into 200kHz RF channels and each of these channels is itself divided into eight time-slots dedicated to a maximum of eight different users [7].

In North America, the same research was undertaken in order to satisfy the demand for higher capacity in large cities. The United States Digital Cellular System (USDC) was developed concurrently to the development of GSM. USDC also uses digital transmission because of its better spectrum efficiency. USDC uses the same frequency bands as AMPS because it was designed to gradually replace AMPS. Consequently, users must be provided with dual mode handsets that can switch from the digital system to the analog system and vice-versa. USDC has been standardized as Interim Standard 54 (IS-54) in 1990. Similar to GSM, IS-54 also has a FDMA/TDMA structure to access the cellular system. Because of compatibility needs with AMPS, the available bandwidth is divided into 30kHz RF channels with each channel supporting three or six users (depending on rate of the speech coder) in a time division fashion. In Japan, the Pacific Digital Cellular (PDC), which uses digital transmission as well, was introduced in 1991 to satisfy capacity demand. Another standard, Interim Standard 95 (IS-95), was also introduced in North America in 1993 to overcome the problem of capacity shortage. The novelty with IS-95 is that it does not use the traditional multiple access schemes of the previous systems. IS-95 uses an access scheme relying on spread spectrum techniques designed and developed previously for military communications. IS-95 uses direct sequence-code division multiple access (DS-CDMA), in which all users use the entire available bandwidth and transmit at the same time. Users are differentiated by the code that they use. Hence, in DS-CDMA the cluster size is one, which eliminates the need for frequency planning. DS-CDMA promises capacity superior to that of FDMA/TDMA [9].

GSM, IS-54, IS-95, and PDC are what are called second generation cellular systems because they employ digital modulation to carry digitized voice. These systems are primarily designed for voice even if they can provide other services like fax, short message services (SMS), and paging services. The second-generation cellular systems are currently serving millions of users as shown in Figure 1 and Figure 3. In addition, wireless and mobile communications are becoming more and more popular. Therefore, new services are emerging and they require more and more resources that current second-generation cellular systems will not be able to provide.

1.4. The future: third generation systems

There are many factors spurring the demand for performance advances beyond the capability of second-generation technology:

- The rapid advance of component technology.
- The pressure to integrate fixed and mobile networks.
- The desire to have one multi-application, hand-held terminal.
- The increasing scope and sophistication of the multimedia services expected by the customer.

All these pressures will lead to the emergence to what is called third generation (3G) cellular systems [8]. The objectives and the system framework of 3G cellular systems are the following [10]:

- The wide range of services and their increasing demand of bandwidth (384kbits/s for full mobility and 2Mbits/s for limited mobility).
- The high quality of service requirements (e.g. toll quality speech with a bit error rate (BER) of 10^{-3}, and data services with a BER less than 10^{-6} or 10^{-9}, depending on the quality of service required).

- Operation in a mixed-cell scenario (macro cells or radius 35km, micro cells of radius 1km, pico cells of radius 50 m).
- Operation in different environments (indoor/outdoor, business/domestic, cellular/cordless, etc.).
- The required flexibility of frequency and radio resource management, system deployment, and service provision.

At a global level, the International Telecommunication Union (ITU) has named these 3G systems, International Mobile Telecommunications 2000 (IMT-2000). In Europe, they are commonly referred to as Universal Mobile Telecommunications Systems (UMTS). In 1988 the European Community (EC) launched an ambitious research initiative called the Research into Advanced Communications in Europe (RACE) program in order to study the enabling technologies for third generation systems. Two projects were related to the air interface, namely, Advanced Time Division Multiple Access (ATDMA) that studied the potential of TDMA-based UMTS, and Code Division Testbed (CODIT) [10] that explored the potential of DS-CDMA as an access scheme for UMTS. The conclusions and findings of CODIT published in [11] were quite promising with respect to the feasibility of a DS-CDMA-based UMTS. After the end of the RACE program in 1995, another series of projects was launched by the European Union (EU) within the Advanced Communications Technologies and Services (ACTS) program. This program, using the RACE experience, was conceived as a demand-driven research and was geared towards the development of demonstration trials. Within ACTS, the Future Radio Wideband Multiple Access System (FRAMES) project has the goal to define, develop and evaluate a hybrid multiple access scheme from different candidates. The FRAMES project largely benefits from the ATDMA and CODIT experiences. FRAMES aims at defining a UMTS multiple access air interface specifications, which will be used as input towards the standardization process of UMTS. After an intensive simulation campaign and demonstration trials, DS-CDMA proved to be a serious candidate for the multiple access scheme of UMTS. At the ETSI meeting held in Paris on 28-29 January

1998, a consensus agreement was reached concerning the adoption of DS-CDMA as a multiple access scheme for the third generation European UMTS.

Similar activities for 3G system developments have been carried out all over the world. In Japan, the Association of Radio Industries Businesses (ARIB) is actively designing and building demonstration trials for 3G systems [12] [13]. The situation in Japan is even more critical than it is in other parts of the world as shown by the explosive growth in cellular subscriber members in Figure 3. Indeed, Japan is targeting the year 2000 for the deployment of 3G systems while Europe is only aiming at 2002 [14]. In North America, the Telecommunications Industry Association (TIA) is working on evolution of IS-95 to satisfy 3G system requirements called cdma2000 [15]. The US is targeting the year 2005 or even later for 3G system deployment because of the huge first generation legacy and the heavy investment in second generation systems. ITU set June 30, 1998 as the deadline for the IMT-2000 proposals for 3G systems. There are nine different terrestrial radio transmission technology (RTT) proposals coming from Europe, Japan, North America, Korea and China [16]. The commonality between all these proposals is that almost all of them use DS-CDMA as an access technique.

1.5. Packet-switched transmission

One of the key requirements for third generation systems is the provision for efficient packet switched transmission. This type of transmission differentiates itself from the traditional view of circuit-switched transmission in first and second generation cellular systems. As a matter of fact, in *circuit-switched transmission*, users are given a channel for the entire duration of the communication. In *packet-switched transmission*, users are only using the channel when they need to transmit, and then release the channel when transmission is completed. As a result, MS transmissions are discontinuous. This allows users to be always "on line" without occupying a dedicated channel. Packet services make it possible for users to pay only for the amount of data transmitted and not the connection time. From a network operator's point of view, packet-switched transmission is attractive because it allows many users to share one transmission medium, since each of these users only transmits a fraction of the time. This type of mapping, involving many users on one medium, is called *statistical multiplexing*. Packet-switched transmission makes a better use of the resources of a network [17].

Packet-switched transmission services in TDMA systems such as GSM have been standardized by ETSI as the General Packet Radio Service (GPRS) [5]. The focus of the present thesis is on packet-switched transmission in 3G DS-CDMA cellular systems.

1.6. Scope of the thesis

This thesis describes an improved medium access control (MAC) protocol that allows for efficient uplink transmission of short packets over a packet data transmission channel (PDTCH). This PDTCH is time-shared by different mobile users in a multi-service DS-CDMA cellular system. Chapter 2 gives background information such as a short description of DS-CDMA and the issues related to discontinuous transmission. Means to provide reliable data transmission over a wireless medium such as forward error correction and automatic repeat request will be discussed. Next, a description of the radio interface structure of the European 3G RTT proposal, UTRA, will be given. Then issues related to packet switched data transmission in the European 3G systems will be covered. Finally some background on queueing theory will be given. Chapter 3 describes the new MAC protocol proposed and its advantages over other proposed MAC protocols. Chapter 4 presents the performance analysis of the new MAC protocol as well as the conventional one. Chapter 5 gives the results obtained through analysis and simulations. Chapter 6 shows the improvement in capacity obtained in using such a MAC protocol in cellular systems. In conclusion, Chapter 7 gives a summary of the thesis and presents some directions for future work.

Chapter 2. Background

This chapter gives background information on the concepts and issues related to DS-CDMA systems and packet-switched data transmission within these systems.

2.1. Direct-sequence code division multiple access

The goal of this section is to give a simple introduction to the issues related to discontinuous transmission (DTX) used in packet-switched transmission in the context of DS-CDMA cellular systems. It will be shown that for DTX, the problem of pseudo-noise (PN) code synchronization and power control must be carefully dealt with. It is not the purpose of the present section to give an in-depth description of DS-CDMA and spread-spectrum in general. The following description is similar to the one used in [9].

2.1.1. DS-CDMA in general

A binary phase shift keying (BPSK) modulated signal will be considered, $S(t)$, upconverted to the carrier frequency f_0 with power P

$$S(t) = \sqrt{2P}x(t)\cos(2\pi f_0 t) \tag{1}$$

where $x(t) = \pm 1$ is the data stream. If BPSK spreading is chosen, this signal is multiplied by the PN (or spreading) sequence $G(t) = \pm 1$ prior to transmission. The transmitted signal is therefore

$$S_t(t) = \sqrt{2P}x(t)G(t)\cos(2\pi f_0 t) \tag{2}$$

Note that by multiplying $S(t)$ by $G(t)$, a signal with a much wider bandwidth than $S(t)$, the resulting signal, $S_t(t)$, has the same bandwidth as $G(t)$. The spectrum of $S(t)$ has been *spread* over the bandwidth of $G(t)$. The *processing gain* of a DS-CDMA system is the ratio of the bandwidth of $S_t(t)$ to the bandwidth of $S(t)$. $S_t(t)$ is going to be sent and will arrive at the receiver after T seconds of propagation delay. First the receiver is going to correlate the received signal with a local replica of the same spreading sequence offset by $\hat{T}$, $G(t-\hat{T})$. The signal coming from the correlator is

$$S_c(t) = \sqrt{2P}x(t-T)G(t-T)G(t-\hat{T})\cos(2\pi f_0(t-T)) \quad (3)$$

The receiver will have to synchronize its locally generated spreading sequence with the one received in order to allow the energy to be despread. If the delay estimation is correct, $\hat{T} = T$, then $G(t-T)G(t-\hat{T}) = 1$ because $G(t) = \pm 1$. Thus the signal after correlation is $\sqrt{2P}x(t-T)\cos(2\pi f_0 t)$ and can be conveniently phase demodulated [23]. If the local spreading sequence is not synchronized, $\hat{T} \neq T$, or if the local spreading sequence is different from the one sent, then despreading of the signal will not occur and the received signal will look like noise to the demodulator [18]. This is why in DS-CDMA different MSs with different spreading sequences can transmit at the same time and in the same RF channel. Indeed, the spreading sequences are chosen so that their cross-correlations are very small and the signal at the output of the correlator will consist of the correctly despread signal plus a sum of unintended signals that look like noise. However, the sum of the unintended signals can become large and, the interference they create can significantly degrade performance of the demodulation of the intended signal. Thus, the performance of CDMA systems is interference-limited [18]. The sum of the unintended signals is *called multiple access interference* (MAI). MAI is the primary limiting factor of CDMA systems. Note that there is a trade-off between capacity and quality of transmission. Capacity of CDMA systems is not limited by the number of RF channels (like in FDMA) or the number of time-slots (like in TDMA). CDMA system capacity is limited by the quality of transmission users can tolerate. These systems exhibit what is called *soft capacity*. In [19], it is shown that capacity is proportional to the processing gain and the inverse of the ratio of the bit energy to interference power spectral density denoted by E_b/I_o. This ratio is also called signal-to-interference ratio (SIR).

2.1.2. Synchronization

For continuous transmission, acquisition of synchronization of the local spreading sequence is only required once: at the beginning of the communication. Since the connection is torn down only at the end of the communication, the time taken for the synchronization process is usually negligible compared to the total duration of the call.

However in a DTX approach, the connection is released after each packet is transmitted. Therefore every time a packet needs to be transmitted, the acquisition step must resume because the transmitting MS could have moved during the silent period, changing the propagation delay of the signal. The overhead of the synchronization time can be significant compared to the actual transmission time. This overhead is particularly significant in the case of short packets. Therefore, a prerequisite for efficient DTX in DS-CDMA is a short preamble before packet transmission resumes in order to allow the receiver to re-synchronize the PN code. A fast synchronizer is also needed to avoid too much overhead in setting up connections [20].

2.1.3. Power control

It was previously stated that the receiver only despreads the user's signal with the same spreading sequence, while the other MS signals look just like noise after the despreading. However, this latter statement must be taken carefully. The following situation will be considered. On one hand, there can be a transmitting MS that is far from the BS, therefore the signal that the BS receives from the MS is weak. On the other hand, there can be one MS very close to the BS, and the MAI that this MS creates may be too much interference for the correct demodulation of the first MS. The problem of having a MS (close to BS) transmitting too strongly is called the *near/far* problem. In order to remedy this situation the power of the MSs must be carefully controlled so that their signals arrive at the BS at the same power level. In this manner, the effect of interfering MSs can be minimized and the capacity of CDMA cellular systems maximized [19]. For a good description of the power control mechanism the reader is invited to refer to [21].

The establishment of the correct power level for transmission is crucial to the good operation of a DS-CDMA cellular system. With too little power, the intended signal will not be correctly demodulated and with too high a power used, the MS signal might deteriorate the quality of other MS communications (near/far problem). The determination of the correct power level is first established in an open-loop fashion at the beginning of a communication. In an open-loop power control, MSs first measures the received power from the BS to MS link (downlink). Then, the power used to transmit on the link from the MS to the BS (uplink) is determined according to that measured power. However, open-loop power control does not give accurate short-term average power levels. It only compensates for the path loss and very slow small scale fading. Another type of power control method is needed to refine the power level estimation. As soon as communication has been established between the BS and the MS, the BS will start measuring the power received and send power control commands to the MS so that the MS adjusts its own power level to a more accurate value. This type of power control involving both the MS and the BS is called *closed-loop power control* (CLPC). CLPC compensates for faster small scale fading and inaccuracies in open loop power control.

In a continuous transmission, the problem of determining the correct level to transmit at is mainly faced at the beginning of the communication. Once this level is found and assuming that the CLPC algorithm can track the channel variations, the communication can take place in good conditions. In DTX, the link between MS and BS is constantly torn down, so CLPC needs to be re-started for every packet. This situation makes CLPC implementation hard and less efficient because continuous transmission of power commands is not possible. Therefore, it is necessary to have a short period of power level adjustment every time a transmission resumes.

2.2. Forward error correction and automatic repeat request

Mobile communication systems use various signal processing techniques that make communications possible in the hostile mobile radio environment. One technique that is essential to data transmission over a mobile channel is *channel coding*. Channel coding improves mobile communication link performance by adding redundant bits in the message to be transmitted [7]. The channel coder maps a digital message into a coded sequence that will be modulated and transmitted. At the receiver, after demodulation, the decoder is able to detect and/or correct some or all of the errors caused by the mobile channel. The coding of the initial message is done following precise rules that depend on the code used. The receiver uses these rules to detect and/or correct the error introduced by the channel. Channel codes that are used to detect errors are called *error detecting codes*, while those used to correct errors are called *error correcting codes.* In general, these codes are called *error control codes.* Some codes can achieve both detection and correction.

Error correcting/detecting codes perform best when the channel is memoryless [23]. That is, the errors caused by the channel are not correlated. However, a mobile channel causes errors to be correlated [18]. Therefore, in order to improve performance, messages are always interleaved before being transmitted. *Interleaving* is basically a way to shuffle the data. At the receiver, the messages are deinterleaved. This decorrelates the errors, which then look independent. This way the performance of the error detecting/correcting codes is improved. The drawback of interleaving is that it introduces delay in the transmission.

In communications that can only be accomplished in one way, or in communications, for which the use of a return channel is not appropriate, only the error correcting capabilities of the codes are used. The receiver only corrects errors, using what is called *forward error correction* (FEC). When a return channel can be used, then error control mechanisms involving the use of error detection and retransmission of erroneous messages can be used. This mechanism is called *automatic repeat request* (ARQ). An instance of service that uses FEC is voice communication. The real-time aspect of voice communications precludes the use of ARQ because of the delay introduced by retransmissions. In contrast, when data transmission for which real-time is not an issue, FEC combined with ARQ greatly improves performance [22].

2.2.1. Forward error correction

2.2.1.1. Description

There are two different types of error control codes: block codes and convolutional codes [22]. A block code encoder takes k bits and appends $n-k$ bits (with $n>k$) to it so that the resulting block is n bits long, this resulting block is called a *codeword.* The code is referred to as an (n,k) block code.

With convolutional codes, a continuous sequence of data bits is mapped into another continuous sequence of encoded bits. This mapping is highly structured and requires a totally different decoding approach in comparison to block coding. A widely used method of decoding convolutional code is the Viterbi algorithm [23]. A convolutional code is generated by passing the data bit sequence through a finite state shift register. In general, the shift register contains N_{cl} (k bit) stages and n linear algebraic function generators. The input data is shifted into and along the shift register k bits at a time. The number of output bits for each k bit input data sequence is n bits. Each encoded sequence of n bits depends not only on the k-bit sequence at the same time, but also on the N previous k-bit sequences. The code is referred to as an (n,k) convolutional code.

A method to improve the performance of codes is *concatenation* [22]. Concatenation involves successively coding a data stream by two distinct codes. The two codes used may be a binary (n_1,k_1) code C_1 and a non-binary (n_2,k_2,b) block code C_2. The non-binary (n_2,k_2,b) block code encodes k groups of b bits into n groups of b bits. The data stream is first encoded by the non-binary (n_2,k_2,b) block code, called the *outer code*, and then by the binary (n_1,k_1) code, called the *inner code*. Note that the inner code can be a block code or a convolutional code. The resulting doubly encoded data stream is then sent. At the receiver, decoding following the rules of the inner code C_1 is performed first followed by the decoding of the outer code C_2. If designed carefully, a concatenated code can be very effective against bursts of errors [22]. This is particularly interesting when the inner code is a convolutional code and Viterbi decoding is used to decode it. The errors at the output of the Viterbi decoder are grouped in bursts so the use of a non-binary block code to correct these bursts is most appropriate [24].

2.2.1.2. Error control codes and DS-CDMA

The *rate* of either a block or convolutional code is defined as $R = k/n$. Note that the bit rate of the resulting data stream is multiplied by $1/R$. Since $R < 1$, the resulting bit rate is actually increased. However, in DS-CDMA systems for which the bandwidth of the transmitted signal is determined by the bandwidth of the spreading code (see Section 2.1.1), the increase of the data rate is not as important an issue as it is in TDMA. Given the same bit error rate, channel coding enables the modulator/demodulator to operate at a lower signal to interference ratio (E_b/I_o parameter defined in 2.1.1) compared to a transmission without channel coding. The difference between these signal-to-interference ratios is called *coding gain*. Since the capacity of CDMA systems is inversely proportional to E_b/I_o, being able to transmit at a low E_b/I_o through coding contributes to increasing the capacity. The coding gain depends on several parameters like the bit error rate, the type of codes and their characteristics, the modulation used, the decoding method used, etc. A typical coding gain value for a rate 1/2 convolutional code (with constraint length $N_{cl} = 7$) operating at BER of 10^{-3} with Viterbi decoding is 4dB [23].

2.2.2. Automatic repeat request

Despite the improvement in bit error rate obtained by FEC, data communications require even higher reliability than what FEC can provide under reasonable operating conditions. ARQ can provide the desirable extra reliability. In an ARQ system, a code with a good error detecting capability is used. Either convolutional or block codes can be used to detect errors [22]. However, the focus will be on the use of block codes since they are more widely used.

2.2.2.1. Description

In an ARQ system, every received codeword is checked for errors by computing its syndrome. The *syndrome* of a received codeword is a vector of $n-k$ bits computed according to rules characterizing the block codes used. If the syndrome is the zero-vector, the received vector is assumed to be free of errors. If the syndrome is different from the zero-vector, then the received codeword contains errors. The received codeword will be discarded and the receiver will ask for the retransmission of this codeword through the return channel. The transmitter will send this codeword until it is successfully received. Therefore, codewords are delivered to the data sink only if the receiver fails to detect the errors. Using a proper block code, the probability of undetected errors can be made very small and very high reliability can be obtained using ARQ. This is why ARQ mechanisms are very attractive for providing reliable data communications [22]. A performance measure for ARQ schemes is the *throughput efficiency*, which is defined as the number of codewords successfully received over the total number of codewords sent.

2.2.2.2. ARQ schemes

There are mainly three types of ARQ schemes: stop-and-wait, go-back-N, and selective repeat ARQ.

In stop-and-wait ARQ, the transmitter sends a codeword and waits for a notification from the receiver. If the codeword is received error-free, the receiver will send an acknowledgement (ACK) back to the transmitter. If the receiver detects errors in the received codeword, it sends a negative acknowledgment (NACK) back to the transmitter to ask for a retransmission. In stop-and-wait ARQ, once the transmitter has sent the codeword, it stops and waits for an ACK or a NACK. This ARQ scheme is inherently inefficient since the transmitter sits idle most of the time waiting for a notification.

In go-back-N, the transmitter sends codewords continuously. It does not wait for an ACK or a NACK after sending the codeword. The transmitter keeps sending codewords until a NACK is received. The acknowledgement for a codeword arrives after a round trip delay. The *round trip delay* is the time interval between the transmission of a codeword and the receipt of its acknowledgement. During this interval, there has been $N-1$ other codewords transmitted. When a NACK for a codeword is received, the transmitter goes back and resends that codeword and the succeeding $N-1$ codewords that were transmitted during the round trip delay. In other words, the transmitter goes back to the previous N codewords and resumes transmission from there. In this operation the transmitter has resent N codewords. In order to be able to resend these N codewords, the transmitter must be provided with buffers to store these codewords. On the other side, the receiver does not need to have any buffer. A codeword received in error will trigger the retransmission of itself and all codewords that have followed. At the receiver, the $N-1$ codewords that follow an erroneously received codeword are discarded regardless of whether they are error-free or not. Go-back-N

ARQ provides better throughput efficiency than stop-and-wait ARQ. However, for channels with high data rate and large round trip delay, go-back-N ARQ is not appropriate. The retransmission of many possibly error-free codewords (after a codeword has been detected in error) is not efficient [22].

The last ARQ scheme that is going to be discussed is selective repeat ARQ. In selective repeat ARQ, codewords are transmitted continuously. However, as opposed to go-back-N, the transmitter only resends codewords that are received in error. Figure 4 illustrates the mechanism of selective-repeat ARQ with a round trip delay of 3 codewords. In order to deliver codewords in the same order that they arrive, selective-repeat ARQ requires that the transmitter as well as the receiver be provided with buffers to store the packets. The receiver must store the error-free codewords that are received after an erroneously received codeword. When the first negatively acknowledged codeword is successfully received, the receiver releases all the error-free codewords in consecutive order until the next erroneous vector is encountered. Sufficient buffer size at the transmitters and also at the receiver must be provided in order to avoid buffer overflow and loss of data. Selective-repeat ARQ is the most efficient ARQ scheme but also the most complex one to implement. Selective-repeat ARQ has the highest throughput efficiency among all three ARQ schemes described. The throughput of ideal (with infinite buffer size) selective-repeat ARQ, η, is [22]

$$\eta = 1 - p_{retx} \tag{4}$$

where p_{retx} is the probability of codeword retransmission, which depends on the code used and the channel conditions. (4) shows that η decreases as p_{retx} increases.

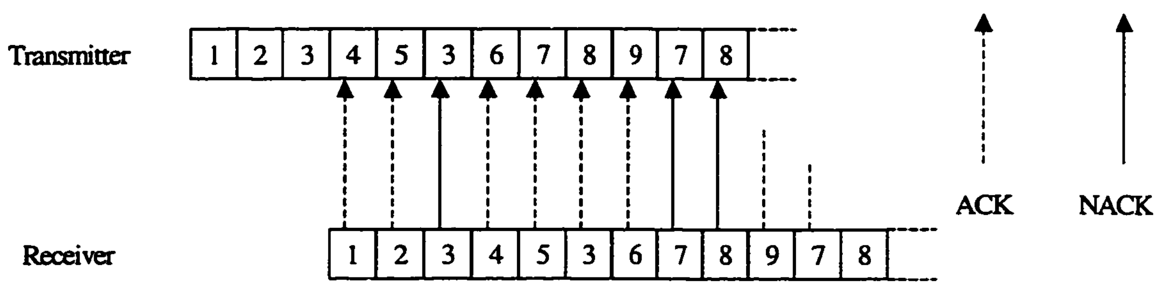

Figure 4. Selective-repeat ARQ.

A variation of selective-repeat ARQ involving only sending NACKs can also be used [25]. This reduces the amount of information exchanged between the transmitter and the receiver. In a DS-CDMA context where interference limits the system capacity, this approach is attractive. In NACK-based ARQ, the receiver does not send ACKs back to the transmitter. When the transmitter does not receive any notification back from the receiver, it assumes that codewords are received correctly and keeps sending codewords. In order to prevent the situation where the retransmission request does not reach the transmitter, a timer is used at the receiver. If a requested frame does not reach the receiver within an interval of time set by the timer, another request will be sent. This process will be repeated until the codeword is obtained. In order to keep track of the codeword order, each of the codewords has a sequence number. Details of this scheme are given in [25].

2.2.2.3. Hybrid ARQ schemes

ARQ can provide very simple and very reliable communications. However, ARQ systems have a severe drawback, in that their throughput falls rapidly with increasing channel error rate [22]. FEC systems maintain a constant throughput equal to one since the number of codewords successfully received is the same as the number of received codewords. However, if a codeword is incorrectly decoded, the codeword will still be delivered to the user. Moreover, to achieve reliable communication, very long codes must be used. These long codes make decoding very hard and expensive.

A solution combining the advantages of both FEC and ARQ, called *hybrid ARQ*, is more efficient. In a hybrid ARQ system, the function of the FEC portion is to correct the most frequent errors in order to reduce the frequency of retransmission. Reducing the frequency of retransmission leads to a higher throughput for the ARQ portion. If the FEC portion cannot correct these errors the ARQ portion should be able to detect them and ask for a retransmission. The FEC provides a higher throughput and the ARQ portion provides a higher reliability. This type of hybrid scheme is *called type-I hybrid ARQ*. There is another form of hybrid ARQ scheme [22] but it will not be discussed since it is not used in the present thesis.

2.3. Radio interface structure of 3G cellular systems

Since the present thesis deals with MAC protocols for packet data transmission; it is necessary to explain shortly what a MAC protocol is. The description given here is in accordance with the Open System Interconnect (OSI) reference model shown in [17]. The air interface (also called the *radio interface*) is layered into three protocol layers: the physical layer (PL), the data link layer (DLL), and the network layer (NL). The DLL is in turn split into two sublayers: the link access control (LAC) and the medium access control (MAC). The focus of the present thesis in on the MAC sublayer, hence the following description will be restricted to the MAC sublayer and the physical layer. The following description will focus on the uplink. The radio interface protocol architecture will be illustrated in the context of the proposed European UMTS Terrestrial Radio Access (UTRA) [26] in order to make the presentation more concrete. Other radio interface protocol architectures are similar to UTRA [16].

2.3.1. The physical layer

2.3.1.1. Physical layer services

The physical layer offers services to the MAC and higher layers. Transport channels of the PL are used to provide these services. They are characterized by the way data is transferred over the radio interface. These transport channels are classified in two groups: common channels (CCH) and dedicated channels (DCH). The first category refers to channels that are used to identify MSs when a particular MS is addressed. The second category refers to channels where the MSs are identified by the physical channel (*i.e.* code and frequency).

Common transport channel types include the random access channel (RACH), the forward access channel (FACH), the broadcast control channel (BCCH), and the paging channel (PCH). The RACH is used by MSs to access the BS and exists in the uplink only. The BCCH provides information to all MSs in a cell on how the RACH should be accessed. The BCCH is a downlink channel. The FACH is used by the BS to carry control information to MSs. The FACH is also a downlink. The PCH has the same function as a FACH except that is used when the system does not know the location cell of the MS being called.

Dedicated channels (DCH) are used for the direct addressing of MSs. Several DCHs can be multiplexed into a data stream before coding, modulation and spreading. The resulting composite stream is then mapped into one or several dedicated physical data channel(s) (DPDCH). DPDCHs are divided into frames of duration 10ms and parameters of the transmission (*e.g.* bit rate) can be varied from one frame to another. An extra physical channel provides information on some of the following: how many DPDCHs are used, how the multiplexing of DCHs is performed, what bit rate is used etc. This type of information is called the *transport format*. This channel is the dedicated physical control channel (DPCCH). It carries a transport format indicator as well as power control commands, and pilot symbols in order to allow the receiver to perform channel estimation for coherent detection.

2.3.1.2. Physical layer functions

The PL functions include the functions described in Section 2.1. These functions are FEC encoding and decoding of the transport channels, mapping of composite channels on physical channels, modulation and demodulation of the physical channels, spreading and despreading of physical channels, closed-loop power control, PN code synchronization, etc.

2.3.2. MAC sublayer

This section will cover the services and the functions of the MAC sublayer. Also included in this section is another sublayer, the radio link control (RLC) sublayer, for which the question of being included in the MAC sublayer is an open issue [26].

2.3.2.1. MAC services

The main responsibility of the MAC sublayer is to handle access to the physical layer. The MAC sublayer deals with the mapping or/and multiplexing of user information and control signaling into transport channels. The MAC sublayer provides the following services to the LAC sublayer: establishment and release of connections, and peer-to-peer transportation of LAC protocol data units (PDU).

2.3.2.2. MAC functions

The functions of the MAC sublayer include: multiplexing of higher layer PDUs into transport blocks delivered to the physical layer on transport channels (and the inverse operations), selection of the transport format within the available transport format set, priority handling, dynamic scheduling, contention resolution on RACH, and identification of MSs on common transport channels. Here, the focus will be on the two last functions because they have a profound impact on the performance of packet-switched services. Note that this list of functions may not be complete [26].

Identifying MSs will be shortly discussed first. When a MS requests access on the common channel RACH, there is a need for an MS identifier in the access process. The same problem is raised on the common channel FACH.

Now the contention resolution problem on RACH will be considered. Cellular systems are multi-access systems, since many MSs in a cell may want to access the BS, which has a limited number of channels to offer to the MSs. The role of the MAC layer in those systems is to ensure that access requests to these channels are properly handled. In order to illustrate this point, the following example is going to be taken: a cellular system, in which only one channel is available for the MSs to make the access requests. Then, the role of the MAC protocol is to ensure that two simultaneous requests (this situation is called a *collision*) are properly dealt with and that these requests are eventually both satisfied. A channel on which users compete to gain access is called a *contention-based* channel.

2.3.2.3. RLC sublayer

The RLC sublayer provides the LAC sublayer with the following services: assured mode operation, unassured mode operation, and transparent mode operation. In the assured mode operation, a reliable link is achieved through the use of ARQ (see Section 2.2.2.2). In the unassured mode of operation, no ARQ is used. The previous two modes allow variable bit rates. In the transparent mode, the data stream will pass through the RLC sublayer without any processing, without appending overhead to the data stream.

The RLC functions are the following: segmentation and assembly of LAC protocol data units and ARQ. Either selective repeat request or go-back-N ARQ is proposed for ARQ schemes in [26].

The RLC sublayer is basically a layer that makes the unreliable wireless channel more suitable for data transmission. Communications protocols involving data need low error rate channels for correct operation. The goal of the RLC sublayer is similar to that of the radio link protocol (RLP) proposed for data transmission in IS-95-based cellular systems [27].

2.4. Packet-switched data service in DS-CDMA

For data networks in general, as seen in Section 1.5, statistical multiplexing provided by packet-switched transmission is an efficient way to achieve high capacity. It was also stated in Section 2.1.1 that, in DS-CDMA cellular systems, MAI limits the capacity. Therefore, the DTX implied by packet-switched data service is very interesting for providing high capacity. However, provision of DTX is a difficult issue as explained in Section 2.1.2 and 2.1.3. Also, DTX raises the problem of accessing the system after a period of silence. The next section presents the DTX scheme used in the European CODIT system that finds solutions to the problems stated previously, but that also raises new problems.

2.4.1. Discontinuous transmission

2.4.1.1. CODIT approach

In the European 3G cellular system proposal, UTRA (described in Section 2.3.1.1), every uplink communication is associated with actually two physical channels: the DPDCH and the DPCCH. The DPDCH is used to carry data generated by the data link layer and the upper layers, while the PDCCH is used to carry physical layer control information for correct demodulation of the DPDCH. One approach to provide DTX could be the one described in [11]. Following this approach, the DPCCH is transmitted continuously in order to allow continuous closed-loop power control, and to keep the synchronization of the spreading codes. This approach allows an easy DTX operation for the DPDCH. The DPDCH can just stop and resume transmission at will since the DPCCH ensures continuous synchronization and power control. This method is illustrated in Figure 5.

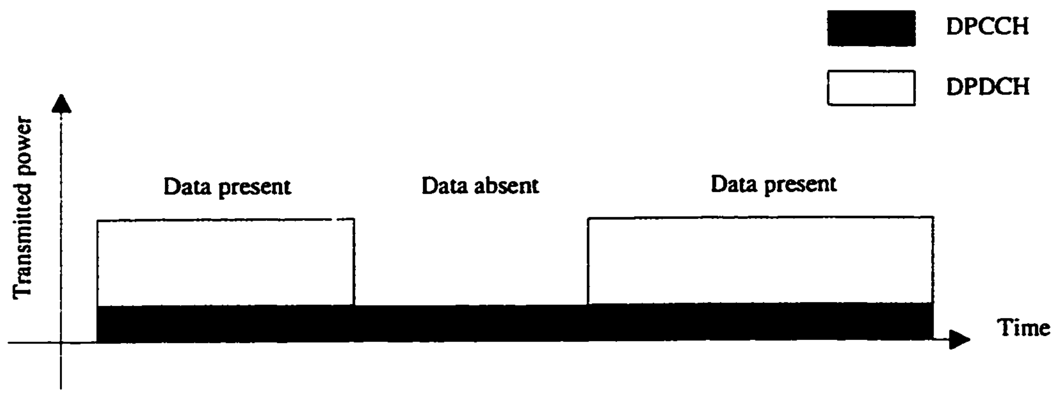

Figure 5. Parallel transmission of DPDCH and DPCCH for DTX.

This method of DTX is simple and therefore relatively easy to implement. This method solves the problem of re-synchronizing the PN codes and re-establishing power control after an interruption. Moreover, the MS does not need to send any access requests after the initial access. This type DTX service is adequate for data sources for which the mean silent period is short. It is also suitable for data sources with very long packets. This DTX approach follows a circuit-switched approach because the channel is not released after each transmitted packet. This approach will be referred to as the *CODIT approach* for data transmission in the rest of this thesis.

However, from an interference point of view, the previous method is not the most efficient and this will be shown in Section 6.3. In fact, the continuous transmission of the DPCCH requires an extra amount of power that is not necessary if the data source goes off for long periods of time. Depending on the quality of service to provide, the DPCCH requires from 12 to 25 percent of the power of an active DPDCH [26]. This fraction of the power is quite significant. It was mentioned in Section 2.1.1 that the capacity of DS-CDMA systems is interference-limited. Therefore, in terms of maximizing capacity, this form of DTX is not optimal. A DTX approach where both the DPCCH and DPDCH go off is preferable.

Moreover, the CODIT approach does not save BS equipment since the PN code used by the DPDCH is not released when the data source goes off. On the one hand, the number of MSs using packet-switched transmission can be large because of statistical multiplexing (see Section 1.5). On the other hand, the number of codes available in a cell is limited (256 Very Large Kasami codes in [26]). Therefore, it is more efficient to share one code to provide packet-switched transmission to several MSs than giving each of them one code that they do not use most of the time. In conclusion, it is preferable to share a code for provision of packet-switched data and totally stop transmission (including control information) if no packets need to be sent.

2.4.1.2. A example of FRAMES approach

Another proposal based on the same concept as before has also been proposed in FRAMES mode 2 [33]. This proposal aims at minimizing the level of interference caused by the DPCCH and improves the sharing of resources among MSs. In this proposal, MSs go through a regular call setup in order to register to a packet service. After the registration, MSs go to an idle mode until they have a packet to send. When a MS has something to send, it sends an access request on the RACH. Then, it waits until the BS echoes an access grant on the FACH. Only when the access is granted can the requesting MSs start sending its packet on the dedicated channel granted. After the MS has sent its packet, it will stop transmitting on the DPDCH but the DPCCH will be maintained over a link time-out period in case the MS has another packet to send within that period. In order to acknowledge the correct reception of the packet, the BS will send an acknowledgement (ACK) message on the downlink of the DCH. The choice of the time-out is a trade-off between frequent random access and radio resource plus hardware spending. This protocol is exemplified in Figure 6.

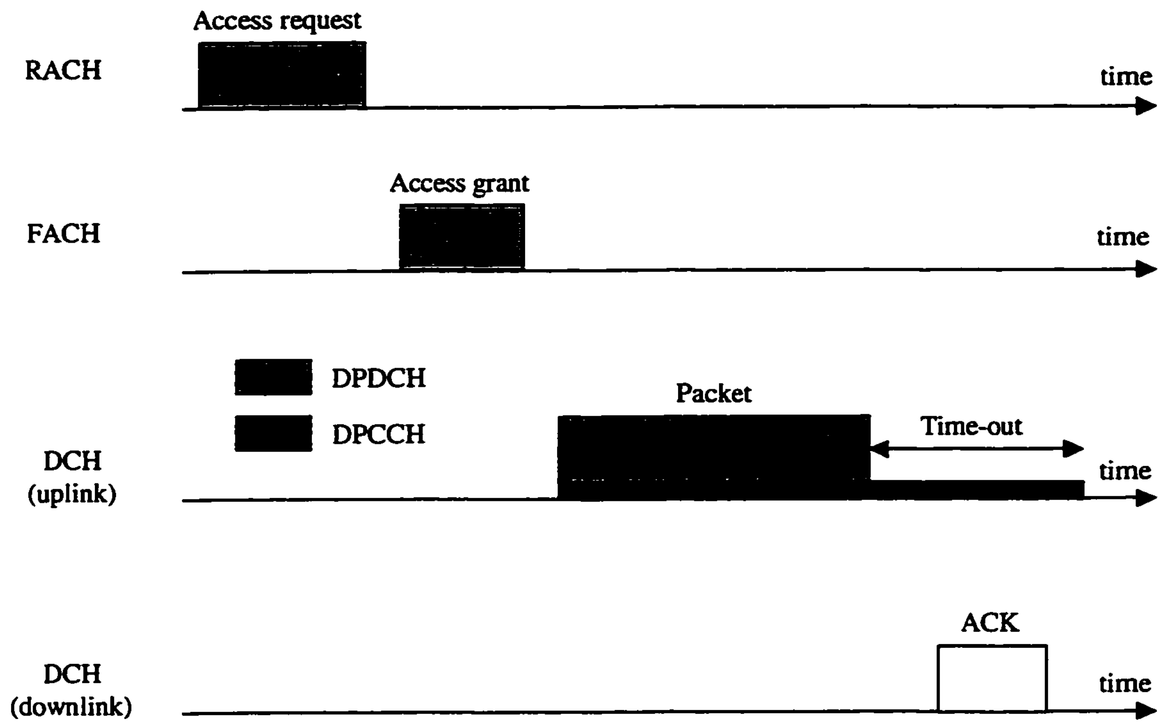

Figure 6. Discontinuous transmission with time-out [33].

Note that this protocol for providing packet-switched service is similar to that of cdma2000, the proposal of the 3G system based on an evolved version of IS-95 [28]. In the cdma2000 proposal, after a period of inactivity, a MS will transition from an "active" state to a "control hold" state in which the dedicated traffic channel is released for other MSs to use.

This method reduces the level of interference caused by the control channel (DPCCH in [26] and the dedicated control channel in cdma2000) compared to CODIT approach. However, this scheme still leads to unnecessary interference during the time-out period. Furthermore, it is not adapted to the case where many MSs are requesting access to the dedicated channel while this channel is still in the time-out period. This time-out period has a very bad effect on the access delay of other MS packets. This effect will be pointed out in the main part of the present thesis (Chapter 4). The time-out period leads to an underutilization of the dedicated channel when the aggregate traffic of all MSs is high.

2.4.1.3. A conventional approach

In order to keep this underutilization to a minimum, the time-out period could disappear. Indeed, right after the ACK has been received from the BS (see Figure 6), the BS can allocate the common channel for another packet-switched data user. This method will be referred to as the *conventional* MAC protocol approach for packet-switched data transmission. This approach is described in [29].

2.4.2. Random access

So far, the discussion on packet-switched transmission focused on methods for providing DTX. Another important aspect of this type of data transmission is the random part that will be covered in this section.

In the context of circuit-switched transmission in DS-CDMA cellular systems, the main task of the MAC sublayer is to make sure that MS initial access requests are satisfied. Once these initial access requests are satisfied, MSs are granted different spreading codes that allow them to transmit on their respective dedicated physical channels whenever they want to without worrying about accessing the access channel again. Thus, the main task of a MAC protocol is to handle MS initial access requests. The MAC sublayer does not need to actively manage communications once they are set up. One interesting characteristic of DS-CDMA cellular systems is that the coordination needed between MSs is very low.

In the context of packet-switched transmission, and in order to save hardware, it was previously argued that it is preferable to share a channel (or equivalently, a code) among several MSs and release that channel after each transmitted packet. Then, access to the code becomes a more serious issue than it is with circuit-switched transmission because each packet that needs to be transmitted will require an access through the access channel. This might lead to an overload of the access channel and throughput degradation. The problem of accessing a channel has been widely addressed in [17]. The number of users using packet-switched services is expected to be quite significant in 3G cellular systems. These growing packet-switched services will raise the challenge of providing fast and efficient random access schemes that are robust under heavy loads.

The random access scheme proposed in UTRA [26] is described next. In UTRA, the RACH is mapped into the physical random access channel (PRACH). The PRACH is spread by a cell-specific PN Gold code [18]. This channel is based on a slotted Aloha approach. It is a contention-based access channel (see Section 2.3.2.2) and MSs start the transmission on the PRACH at a number of well-defined time-offsets relative to the BCCH (see Section 2.3.1.1). The different time offsets are called *access slots*, and are spaced 1.25ms apart (eight slots per frame of duration 10ms). Information on the available access slots is given by the BCCH. MSs send random access bursts (RAB) that consist of a preamble part (of 1ms) and a message part (of 10ms). The preamble part consists of a randomly selected signature sequence (16 possibilities). First, the BS monitors the received power level from the BS to adjust its transmitting power in a open loop power control fashion (see Section 2.1.3). Then the MS sends a RAB after randomly choosing an access slot among all available slots and a signature sequence. At the BS, the RAB detector is designed such that RABs will collide only if they are sent in the same access slot and have the same signature sequence in the preamble. After sending the RAB, the MS waits for a certain period of time and if nothing is received (because of a collision or because too low a power is used) it sends another RAB. The message part is for service negotiation and a short packet can be included in it. The message part will not be discussed because the focus here is not on quality of service provision. This random access scheme is used for initial access as well as access for packet transmission. However, using the same channel for packet random access and initial access (for voice users as well as data users) might not be the best solution in terms of optimizing system resources. Besides, Aloha-based random access channels are known to provide low throughput under high load [17] [32]. This low throughput is mainly due to the detrimental effect of random access collisions. Aloha also requires a stabilization mechanism so that throughput does not become null [17]. This stabilization mechanism leads to complex MAC protocols.

In cdma2000, the mechanism for re-accessing a packet data channel after an idle period is not executed on the RACH. As stated previously in cdma2000 (see 2.4.1) after a period of inactivity, MSs switch to a control hold state. While MSs are in this state of inactivity, fast reassignment of a traffic channel is still possible through the use of an extra dedicated MAC channel. The dedicated MAC channel is mapped onto a dedicated control channel. However, this form of random access leads to the same MAI problem as maintaining a continuous DPCCH (see the CODIT approach in Section 2.4.1.1).

2.5. Queueing models

One measure of performance, which is of primary importance in packet data transmission system, is the average time required to deliver a packet from origin to destination [17]. For example, for a real-time service such as video-conferencing, knowing the average delay that packets will encounter is valuable information to assess the feasibility of such a service within the system.

Queueing theory permits an analysis of network delay. It often requires simplifying assumptions in order to make analysis possible. Nevertheless, these assumptions often lead to reasonably good delay approximations as well as valuable insights into the different issues involved in the design of packet data services in cellular systems.

2.5.1. Scope of the work

In the following, the emphasis will be on the delay from MS to BS. The delay caused by the network is beyond the scope of the present thesis because it involves routing and flow control of packets [17]. The delay that is of interest is therefore the time between the arrival of the packet at the MS and the end of its successful delivery at the BS. This delay will be called the *MS to BS packet delay*. Figure 7 shows the MS to BS packet delay and *the network packet delay*, which is the time between the reception of the packet at the BS and its successful delivery to the destination host (shown as a computer on the right side of Figure 7). Note that the network reference model used in Figure 7 is that of CODIT [11]. The radio network controller (RNC) manages macro diversity. The mobile control node (MCN) is mainly a gateway that allows the interconnection of the UMTS network to fixed networks such as the public switched telephone network (PSTN), the integrated services digital network (ISDN), and the packet switched public data network (PSPDN).

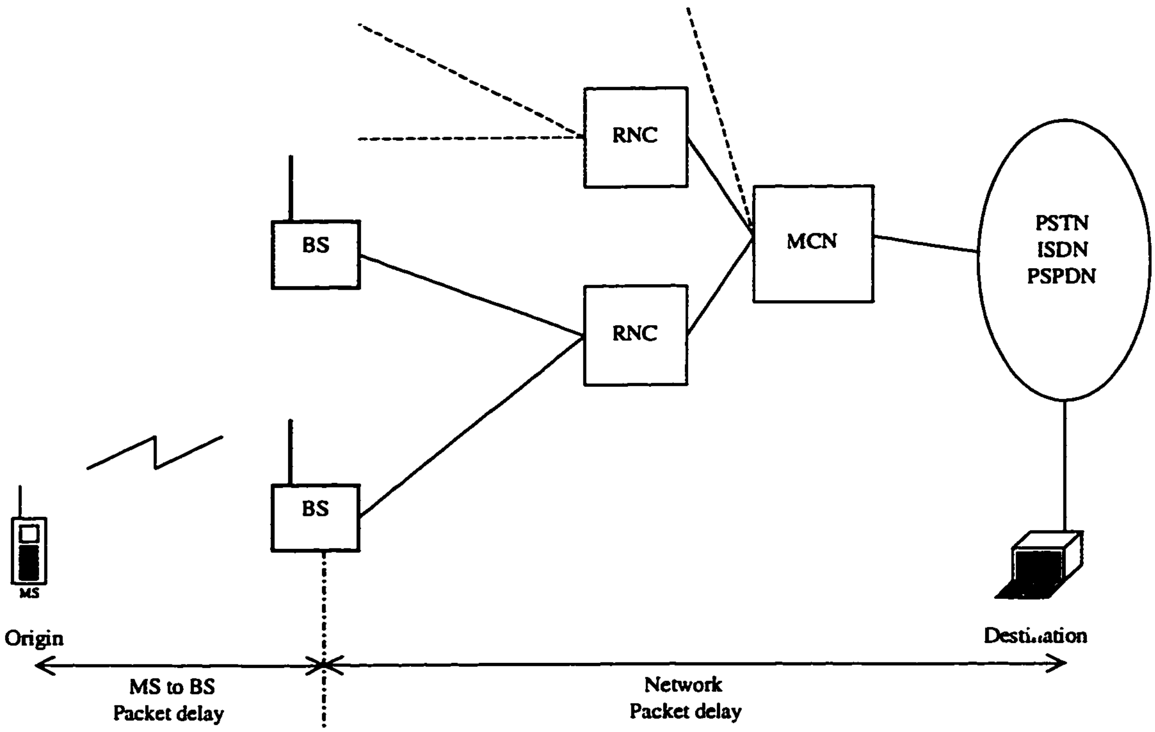

Figure 7. Air interface delay and network delay.

In the rest of the thesis, only the MS to BS packet delay will be discussed, so this delay will just be called *packet delay*. This delay is in turn composed of four different components, namely: the processing delay, the queueing delay, the transmission delay, and the propagation delay [17]. These delays are going to be described one at a time in the context of DS-CDMA cellular systems.

The *processing delay* is the time taken by the receiver to demodulate, de-interleave, and decode a received packet. This delay is deterministic and depends on the structure of the receiver, the error control codes used, and the interleaving depth used. This delay will be assumed to be independent of the computing power of the BS and therefore has a fixed value.

The *queueing delay*, also called the *packet access delay*, is the delay between the arrival of a packet and the beginning of its transmission. During this time, the packet must wait for other packets in the queue(s) to be transmitted. This delay is a random variable because of the random nature of the packet arrivals.

The *transmission delay* of a packet is the time between the start of transmission of that packet and its successful reception. In this thesis the use of ARQ will be considered (see Section 2.2.2). Because of the random nature of packet retransmissions due to errors caused by the mobile channel, the transmission delay is also a random variable.

The *propagation delay* is the time taken by the packet to propagate from MS to BS. In a terrestrial cellular system, the propagation delay is negligible. Indeed, if the maximum radius of a macro cell is 35km [11], the propagation delay is roughly 0.117ms, which is negligible compared to the size of the frames (*e.g.* 10ms). Therefore, in the rest of the thesis, this delay will be ignored.

2.5.2. Little's theorem

In this section, Little's theorem, which is very useful in analyzing delay models, will be reviewed. Little's theorem is general and can be applied to almost all queueing systems that reach a steady state. The following description is taken from [17], which is a very good reference on delay models. However the presentation has been modified in order to focus more on packet transmission.

Before stating what Little's theorem is, a few terms need to be defined. Assume that what is recorded, from time $t=0$ up to an indefinite future, are the values of the following parameters as time progresses: $N(t)$, the number of packets in the system at time t; $\alpha(t)$ the number of packets which arrived within $[0,t]$; and T_i, the time spent in the queue by the i^{th} arriving packet.

The *time average of* $N(t)$ *up to time* t is the following quantity

$$N_t = \frac{1}{t}\int_0^t N(\tau)d\tau \qquad (5)$$

it is also sometimes called the "typical" number of packets in the system observed up to time t. Of course N_t changes with time. However for systems of interest it tends to a steady-state value, N, as t increases. Therefore

$$N = \lim_{t \to \infty} N_t \tag{6}$$

In this case, N is called the *steady-state time average* (or simply *time average*) of $N(t)$.

The *time average of the arrival rate* over the interval $[0,t]$, is the following quantity

$$\lambda_t = \frac{\alpha(t)}{t} \tag{7}$$

Just as before, this quantity can tend to a steady-state value called *steady-state arrival rate* defined as (assuming that the limit exists)

$$\lambda = \lim_{t \to \infty} \lambda_t \tag{8}$$

The *time average of the packet delay up to time t* is similarly defined as

$$T_t = \frac{\sum_{i=0}^{\alpha(t)} T_i}{\alpha(t)} \tag{9}$$

which is the average time spent in the system per packet up to time t. The steady-state time average packet delay is defined as (assuming that the limit exists)

$$T = \lim_{t \to \infty} T_t \tag{10}$$

Little's theorem is the following simple equation that relates N, λ, and T:

$$N = \lambda T. \tag{11}$$

This theorem expresses the natural idea that crowded systems (large N) are associated with long customer delays (large T) and conversely. For example, on a rainy day, traffic during rush hour moves slower than average (large T), while the streets are more crowded (large N). Using the same analogy, a fast-food restaurant (small T) needs a smaller waiting room (small N) than a regular restaurant for the same customer arrival rate.

The following is a graphical proof of the theorem under some simplifying assumptions. Assuming that the system is initially empty and that the packets depart from the system in the same order that they arrive. $\beta(t)$ is the number of departures within $[0,t]$. Figure 8 shows the staircase behaviour of $\alpha(t)$ and $\beta(t)$. Note that the difference $\alpha(t)$- $\beta(t)$ is the number of packets present in the system at time t, $N(t)$.

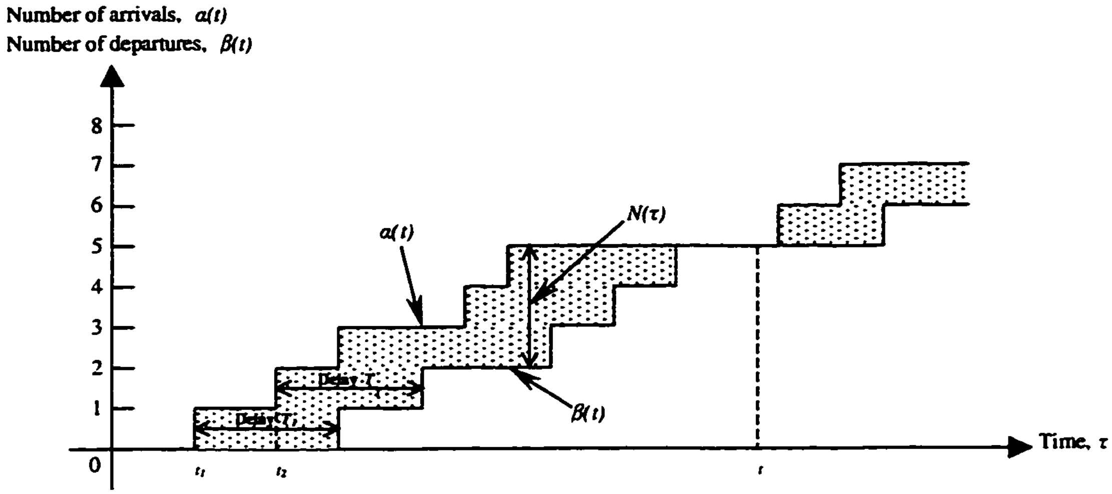

Figure 8. Graphical proof of Little's Theorem [17].

The dotted area between $\alpha(t)$ and $\beta(t)$ on Figure 8, is given by: $\int_0^t N(\tau)d\tau$. For any time, t, for which the system is empty, $N(t)=0$. That dotted area is also equal to $\sum_{i=1}^{\alpha(t)} T_i$. When these two expressions are divided by t

$$\frac{1}{t}\int_0^t N(\tau)d\tau = \frac{1}{t}\sum_{i=1}^{\alpha(t)} T_i = \frac{\alpha(t)}{t}\frac{\sum_{i=1}^{\alpha(t)} T_i}{\alpha(t)} \tag{12}$$

Using the definitions given in (5), (7), and (9), the previous equation can be rewritten as

$$N_t = \lambda_t T_t \tag{13}$$

Assuming that the limits defined in (6), (8), and (10) exist and that the system becomes empty infinitely often, Little's theorem is obtained, as shown in (11), by taking the limit of the previous equation as t goes to infinity.

The previous assumptions can be relaxed while having the theorem remains valid. It is assumed that the limits in (8), (10), and that the following limit, which is the definition of the departure rate

$$\delta = \lim_{t\to\infty}\frac{\beta(t)}{t} \tag{14}$$

all exist. If one further assumes that $\lambda = \delta$, then Little's theorem is still valid [17]. In particular, it is not necessary that the packets be sent in the order they arrive, or that the system be initially empty.

2.5.3. Probabilistic form of Little's theorem

Provided that time averages can be replaced with statistical averages, Little's theorem also has a probabilistic interpretation. The previous discussion dealt only with a single realization (or event) of a time-dependent variable. Now the probabilities of many sample functions are considered and therefore stochastic processes corresponding to the time-dependent variable of the previous section will also be considered.

Denote by $p_n(t)$, the probability of having n packets (in the queue or in service) in the system at time t. If enough statistical information such as the initial probabilities, $p_n(0)$, is given, it is possible to determine $p_n(t)$ for all times t. In such a case, it is possible to compute the average number of packets in the system at time t:

$$\overline{N}(t) = \sum_{n=0}^{\infty} n p_n(t) \tag{15}$$

Note that in (15), all the variables depend on t. However, all the queueing systems that are considered typically reach a steady state. That is, the limits, $\lim_{t\to\infty} p_n(t)$ for $n = 0,1,\ldots$, exist and they will be denoted, p_n for $n = 0,1,\ldots$. In such a steady-state, the average number of packets in the system is given by:

$$\overline{N} = \sum_{n=0}^{\infty} n p_n \tag{16}$$

And typically

$$\overline{N} = \lim_{t\to\infty} \overline{N}(t) \tag{17}$$

Regarding the average delay per packet, enough statistical information is usually given to determine, in principle, the probability distribution of delay of each individual packet (*i.e.* the first, the second, etc.). This is a similar characterization as that of $\overline{N}$ previously discussed. From this information, it is possible to determine the average delay of each packet. Let $\overline{T}_k$ be the average delay of the k^{th} packet. $\overline{T}_k$ typically converges as $k \to \infty$ to a steady-state value:

$$\overline{T} = \lim_{k \to \infty} \overline{T}_k \tag{18}$$

Figure 9 illustrates the previous limit. Recall that for the k^{th} packet, $\overline{T}_k$ is the average delay over all the events and not the time-average delay of a packet as the time goes to infinity.

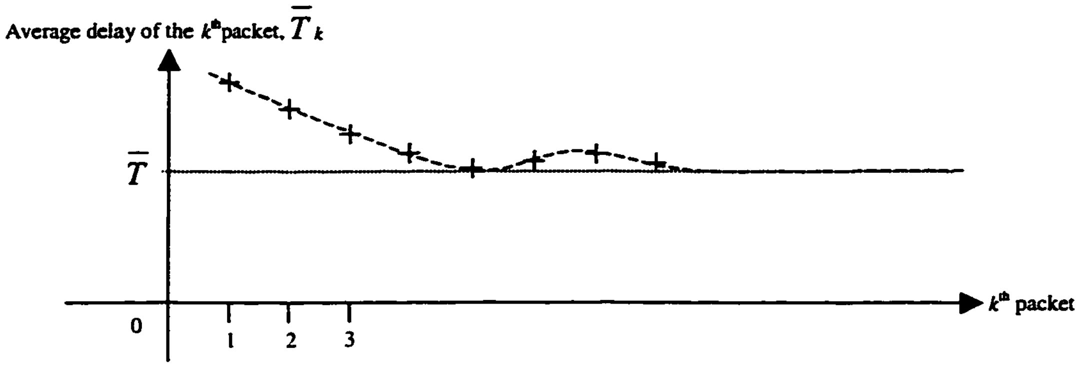

Figure 9. Average delay of the k^{th} packet versus the k^{th} packet.

Now the connection with the time averages that were presented in Section 2.5.2 can be made. Note that all the systems of interest are ergodic, that is, the time average of a single realization of the stochastic process (also called the *sample function*), $N = \lim_{t \to \infty} N_t$ (6), is equal (with probability one) to the steady-state average, $\overline{N} = \lim_{t \to \infty} \overline{N}(t)$ (17). Therefore in the rest of the thesis, it will be assumed that the following equalities hold:

$$N = \lim_{t \to \infty} N_t = \lim_{t \to \infty} \overline{N}(t) = \overline{N} \tag{19}$$

Similarly, for systems of interest, the time average of customer delay T is also equal (with probability one) to the steady-state average delay $\overline{T}$, that is,

$$T = \lim_{k \to \infty} \frac{1}{k} \sum_{i=1}^{k} T_i = \lim_{k \to \infty} \overline{T}_k = \overline{T} \tag{20}$$

Let $\overline{\alpha(t)}$ be the expected number of arrivals in the interval $[0,t]$, then

$$\lambda = \lim_{t \to \infty} \frac{\overline{\alpha(t)}}{t} \tag{21}$$

Using these notations, it is possible to reformulate Little's theorem $N = \lambda T$, with *N* and *T* being the stochastic averages (or also called ***ensemble averages*** of the stochastic processes) defined in (17) and (18) and λ defined in (21). Note that the equality of long-term time average and ensemble average of various stochastic processes will be used throughout the thesis. Therefore the time average notations, *N* and *T*, will often be used for the ensemble average notations, $\overline{N}$ and $\overline{T}$, assuming that the equations (19) and (20) are true.

2.5.4. The Poisson process

Now the Poisson process will be reviewed. This process is a fundamental process in delay models. The Poisson process is generally considered to be a good model for the aggregate traffic of a large number of similar and independent users. Once again the following description can be found in [17].

A stochastic process, $\{A(t) \mid t \geq 0\}$, taking nonnegative integer values is said to be a *Poisson process* with rate λ if:

1. $A(t)$ is a counting process that represents the total number of arrivals that have occurred from 0 to time t [i.e. $A(0) = 0$], and for $s < t$, $A(t) - A(s)$ equals the numbers of arrivals in the interval $(s,t]$.
2. The numbers of arrivals that occur in disjoint time intervals are independent.
3. The number of arrivals in any interval of length τ is Poisson distributed with parameter $\lambda\tau$. That is, for all t, $\tau > 0$,

$$P\{A(t+\tau) - A(t) = n\} = e^{-\lambda\tau}\frac{(\lambda\tau)^n}{n!}, \text{ for } n = 0,1,... \tag{22}$$

Using (22), it can be easily shown that the average number of arrivals within an interval τ is $\lambda\tau$. This leads to the interpretation of λ as an arrival rate as defined in (7) in Section 2.5.2 (average number of arrivals, $\lambda\tau$, divided by the time duration, τ).

Poisson processes have many interesting properties. The reader is invited to refer to [17] for them. Just two of them are going to be presented because they will be used in the analysis of the system in Chapter 4. The first one says that if $k \geq 2$ independent Poisson processes $A_1,...,A_k$ are merged into a single process $A = A_1 + A_2 + ... + A_k$ then the latter is Poisson with rate equal to the sum of the rates of its components. The second property presented states that the interarrival times of packets are independent and are exponentially distributed with parameter λ. The *interarrival interval* is defined to be the interval of time between the arrivals of two successive packets. Then the second property implies that the probability per unit time of an interarrival interval ending is constant, and therefore independent of the current interarrival interval length. This is equivalent to saying that the process is *memoryless*.

2.6. Conclusion

This chapter introduced DS-CDMA and raised issues related to DTX and the provision of packet-switched data service. FEC and ARQ were introduced as means to provide reliable data transmission. The structure of the radio interface used in the European UTRA proposal was also discussed. Different proposals for packet-switched data transmission and the problems that these proposals raised were described. Finally, some background of queuing theory was given for the analysis of Chapter 5.

Chapter 3. The proposed MAC protocol

In this chapter, the new MAC protocol will to be described. The rationale of the MAC protocol design is given.

3.1. Preliminary considerations

As explained in Section 2.4.1, the provision of packet-switched data services is most efficient when MSs that have registered to the services share one spreading (also called channel) code. This channel code sharing saves BS equipment and keeps the level of MAI to a minimum. In Section 2.4.2, it was also stated that a contention-based RACH channel has limited throughput. Therefore, this naturally leads to a choice of an access method that avoids contention. As a result, in the system under consideration, requests from mobile stations to obtain access to the packet-switched channel are made over a non-contention-based access request channel (ARCH) as proposed in [32]. The channel corresponding to the code used for packet data transmission will be called packet data transmission channel (PDTCH)

In order to use the packet switched service, MSs must first register to the service. The procedure for registration to the packet service is similar to the connection to the regular dedicated control channel (DCCH) described in [11][29]. A MS wishing to use the packet-switched service first sends an initial access request through the RACH (see Section 2.4.2). The message part of the MS random access burst contains the indication for the wish to use the packet data service and the bit rate. The BS echoes the access reply (indicating which spreading code to use on the PDTCH) back to the requesting MS on the FACH. The MS can then switch its transmission devices to that code and start using the packet data service. This registration is simple, relatively fast, and avoids too much overhead in setting up connections.

In the next section, a description of the ARCH will be given followed by a description of the PDTCH.

3.1.1. Access request channel

3.1.1.1. Channel structure

This ARCH is an uplink channel that can be used by mobiles which have registered to the packet service. The ARCH is organized into 10 ms frames which are further divided into two subframes of 5 ms each. The first subframe is called the *access request* subframe, while the second is the *acknowledgement subframe*. In the access request subframe, each of the MSs is assigned a slot, in which the access request message (ARM) is sent. Figure 10 illustrates the ARCH structure. Since MSs are assigned their own slots, there are no collisions of ARMs and a MS identifier is not required for MSs either. These slots eliminate the MAC functions required for identifying MSs and for resolving collisions on the access channel (see Section 2.3.2.2). If the number of MSs wishing to register for the packet service exceeds the number of available time slots, the BS can decide to establish another pair of ARCH/PDTCH. An ARM consists of several repetitions, L, of a short PN code, as in [20]. The access attempts of the MSs consist of ARM sequences containing ARMs of increasing power levels.

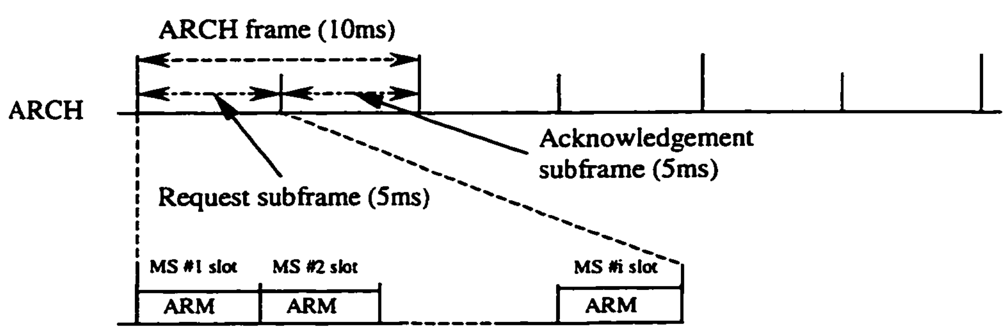

Figure 10. ARCH structure.

3.1.1.2. ARM structure

The ARM structure is obtained using the design guidelines provided in [20]. Let T_{slot} be the duration of a time slot dedicated to a given MS, T_{access} the duration of the access subframe, and N_{slot} the number of slots in the access subframe. These parameters are related by

$$T_{slot} = \frac{T_{access}}{N_{slot}} \tag{23}$$

L denotes the number of short code repetitions in an ARM. Let N_{code} be the length of the short code used (also called the period of the PN code), and F_c-the chip rate used. Note that in the European 3G RTT proposal [26], F_c is equal to 4.096Mchip/s. These quantities are related by

$$T_{slot} = \frac{L \cdot N_{code}}{F_c} \tag{24}$$

Using (24) and (23), L can be obtained as a function of the other terms. Knowing L, it is possible to determine the threshold, θ_d, of the decision device that will minimize the synchronizer false acquisition probability, P_{fa}, and at the same time, maximize the correct acquisition probability, P_c. The probability of false acquisition is given by [20]

$$P_{fa} = e^{-\theta} \sum_{k=0}^{L-1} \frac{\theta_d^k}{k!} \tag{25}$$

θ_d is obtained by setting a maximum value for P_{fa}. In order to neglect P_{fa}, it is reasonable to take P_{fa} smaller than 10^{-6} as suggested in [20]. Once θ_d is found it is possible to determine the probability of correct acquisition in a Rayleigh fading channel, which is given by

$$P_c = e^{-\theta_d/1+\bar{\mu}} \sum_{k=0}^{L-1} \frac{(\theta_d/(1+\bar{\mu}))^k}{k!} \tag{26}$$

Where $\bar{\mu}$ is the average signal to noise ratio. This average signal-to-noise ratio is defined over a entire period of the PN code of N_{code} chips.

It is possible to determine an upper bound on the average number of ARMs needed to acquire synchronization, n_s. This upper bound is obtained by assuming that the probability of acquisition stays the same while the ARM power levels increase. This upper bound is given by [20]

$$n_s < \sum_{i=1}^{\infty} i(1-P_c)^{i-1} P_c = \frac{1}{P_c} \tag{27}$$

3.1.2. Packet data transmission channel

The PDTCH is just a regular traffic channel as found in the European 3G RTT proposal [26]. It consists of a data channel, the DPDCH, and a control channel, the DPCCH (see Section 2.3.1.1). These channels are divided into frames of duration 10ms.

In the proposed scheme, packets are broken up into segments of one frame duration. These segments (also called frames) are encoded and interleaved in order to protect them against channel errors. As discussed in Section 2.5.1, the processing delay of a frame is the sum of the interleaving/de-interleaving, coherent detection, and coding/decoding delays. Figure 11 shows the processing delay for a frame of eight packets.

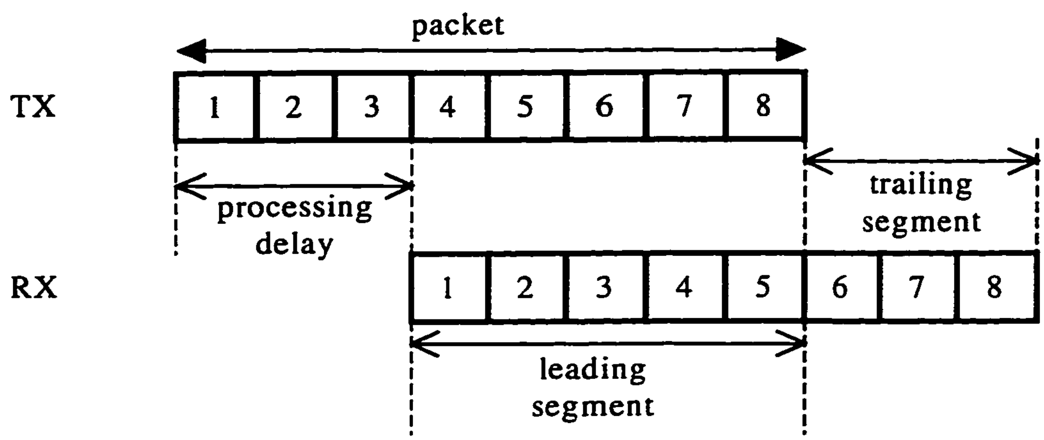

Figure 11. Packet structure and processing delay.

A concatenated scheme (see Section 2.2.1.1), consisting of an inner convolutional code and an outer Reed-Solomon (RS) code, is used for error control [36]. Each frame is encoded by an (n,k,b) RS code defined over $GF(2^b)$ for which k is the number of data symbols per code word, n the length of the code word, and b the number of bits per RS symbol. Such a RS code can correct up to $t_c = \lfloor (n-k)/2 \rfloor$ errors [30]. The RS decoder used is a bounded distance decoder [30], which means that it searches a code word that is within distance t_c of the received word. If such a codeword exists, the decoder finds it; otherwise, the decoder declares a *failure*. Hence, the RS code is capable of both error correction and error detection. Together with the inner error-correcting convolutional code, the decoder operates within a type-I hybrid ARQ scheme with selective-repeat ARQ as described in Section 2.2.2.3. That is, forward-error-correction is performed on each received frame and if errors cannot be corrected, retransmission of this frame will be requested. A copy of each frame is stored in the transmitter buffer before transmission in order to allow its immediate retransmission if necessary. Figure 12 illustrates the concatenated coding scheme and the buffering of frames. Above the data source and sink are higher-level protocol entities. On the left side, the transmitter (TX) structure is shown and on the right side, there is the receiver (RX) structure. In order to minimize interference, NACKs are used to request retransmission (see Section 2.2.2.2). Finally, in order to account for a lost NACKs, retransmission timers are used by the receiver to make sure that frames that need retransmission are eventually sent, as in [25].

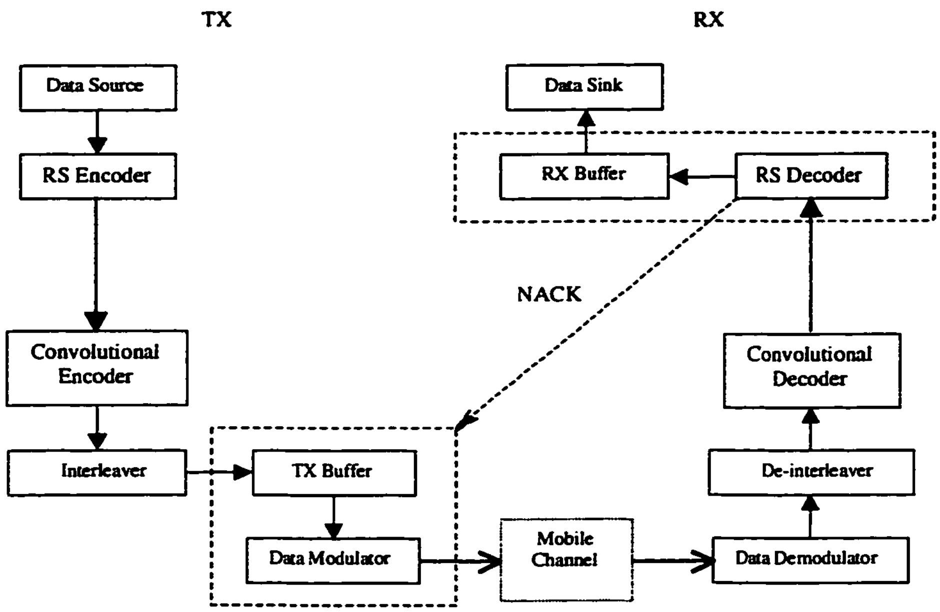

Figure 12. Type-I ARQ scheme with the concatenated error control codes.

3.2. The improved MAC protocol and its advantages

In the proposed MAC scheme, mobile stations are allowed to send access requests on the ARCH only after their home base station notifies them on the PDTCH that it is clear for transmission. Otherwise, mobile users are not allowed to request access and must wait in order to do so. Packets arriving at MSs that cannot be transmitted upon arrival will be stored in MS buffers. In other words, those packets will be queued until they can be transmitted.

When the PDTCH is clear for transmission, MSs having packets in their queues will send ARMs. After an ARM has been detected in the access request subframe, the base station broadcasts a notification on the downlink DPCCH of the PDTCH to all mobiles allowing one of them to transmit. This notification is sent at the beginning of the acknowledgement subframe of the ARCH. The period of time dedicated to sending ARMs will be called the *mobile station (MS) access period.* Once a mobile station (e.g. no. i) is granted access to the PDTCH, the base station activates closed loop power control and also refines the synchronization just acquired on ARCH during the remainder of the acknowledgement subframe. Transmission of a packet starts at the beginning of the next frame. The time between the start of transmission of a packet and the end of it will be called the *packet transmission time.*

The packets are sent in segments in consecutive frames, and some of them will require retransmission. After the i-th mobile station has sent the last frame of a given packet, the base station makes the PDTCH immediately available to other mobile users, so that transmission of the next packet may start without waiting to verify if the most recently received frames need retransmission. The most recently received frames are called the *trailing segment* of a packet while the other packets are the *leading segment.* Thus, the time needed to process the trailing segment of a packet is used by other mobile users for transmission, and hence it does not increase the packet access delay. Figure 11 shows the leading segment and the trailing segment of an eight-frame long packet. Note that from this figure, it is clear that the duration of a trailing segment is the same as the processing delay.

In the event that a negative acknowledgment is received for frames belonging to the trailing segment of the packet transmitted by the i-th mobile station, while another mobile user (e.g. no. j) is transmitting, the PDTCH is given back to the i-th user after the j-th user has completed transmission of its packet. In this manner, the frames of the trailing segment of the i-th user's packet are not lost, but they will incur an extra delay due to the transmission of the j-th user's packet. This MAC protocol can be viewed as a sliding window protocol also called go-back-N protocol [17] for packets belonging to different users. While a user is waiting for possible negative acknowledgments of the trailing segment of its packet, the channel does not stay idle but is assigned to another user. The channel will be reassigned back to the first user if there is an error. This way the time corresponding to the processing delay of the last frame is used to admit another user. Figure 13 points out the difference between a conventional MAC protocol (see Section 2.4.1.3) and the new one. On this figure, the leading segment is denoted as LS, the trailing segment is denoted as TS, and the sequence of ARMs needed to request access to the PDTCH is denoted ARMs. In order to give good insight into the efficiency of the new MAC scheme, the following assumptions are going to be made. On one hand, packet no. 2 on that figure is already in queue when packet no. 1 is transmitted, and on the other hand, the probability of frame retransmission is negligible (in order to ignore TS retransmissions). It can be seen from that figure, for the new MAC protocol, packet no. 2 waits a shorter time before being transmitted than it does for a conventional MAC scheme.

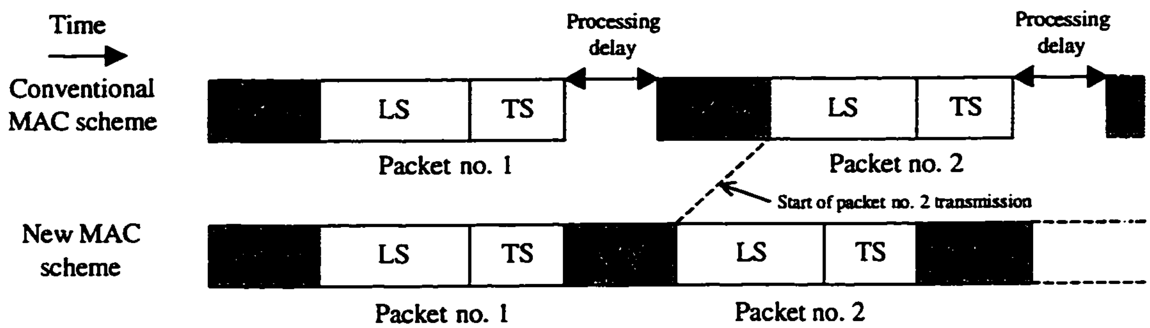

Figure 13. Comparison of a conventional and the new MAC protocol.

If retransmissions are needed, then the previous user will be given access to the PDTCH again. Figure 14 illustrates how the transmission procedure is performed under the new MAC. This figure exemplifies the case where a negative acknowledgement (NACK) is received for the trailing segment of packet no. 1. The MS which has transmitted packet no. 1 must restart a closed-loop power control procedure (denoted "P.C." on Figure 14) in order to transmit at the right power level, and it will also have to resynchronize the pseudo-noise code due to the silent period (see Section 2.1). After this "P.C." procedure the MS can transmit the TS.

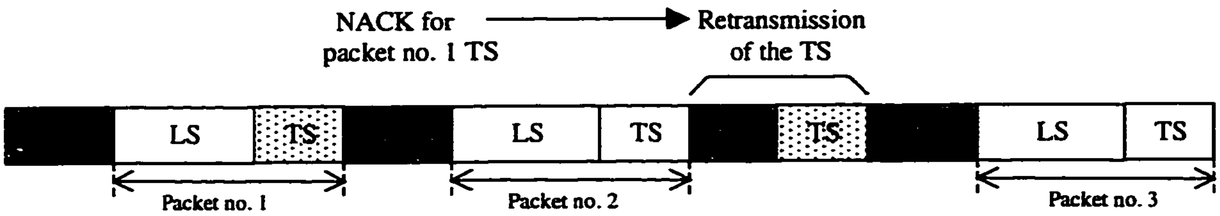

Figure 14. Transmission procedure under the new MAC protocol.

The new MAC protocol handles access to the PDTCH in a very straight forward manner since the ARCH has a slotted structure. This protocol also deals with ARQ. Therefore the new protocol assumes that the RLC sublayer (discussed in Section 2.3.2.3) is included in the MAC sublayer. This is essential to the new MAC protocol since this protocol makes use of the latency introduced by the ARQ scheme.

3.3. Conclusion

This chapter presented the new MAC protocol for packet-switched data transmission. The new MAC protocol makes use of a separate access channel, the access request channel (ARCH). The ARCH is a time-slotted channel that is used by MSs to access the common traffic channel used for packet data transmission. The ARCH is spread by a short code to allow its easy synchronization by a fast synchronizer. The common traffic channel used by MSs is called the packet data transmission channel. It is a time-shared channel that can be used by MSs to transmit their packets. Once packets have been transmitted, the channel is released for other MSs to use. A concatenated coding scheme is used on the PDTCH to provide reliable communications. This type of coding and the necessary interleaving lead to a relatively important processing delay that is used by the new MAC scheme to allow more users into the system. This is what differentiates the new MAC scheme from to a conventional scheme, which does not utilize this processing delay.

Chapter 4. Performance analysis

4.1. Error detection

In this section, we analyze the error detection performance of the (n,k,b) RS code used. If the number of errors is smaller than, or equal to t_c, the decoder can correct all errors in the codeword (correct decoding is accomplished). If the number of errors is larger than t_c, two things may happen. Either the decoder fails to find any correct codeword within distance t_c of the received codeword (a decoder failure), or it will find a codeword different than the transmitted one (a decoder error). When a decoder error occurs, the RS decoder will accept an erroneous codeword, which is why this event is usually called an undetected error. Let P_{cd} be the probability of correct decoding, P_f the probability of decoder failure, and P_e the probability of decoder error (or undetected error). Figure 15 shows the different events that can arise after the RS decoding of an arriving frame. Either the RS decoder is not able to correct the errors (event with probability P_f) then a NACK will be sent back to the transmitter or the RS finds a codeword within distance t and therefore accepts the frame. In the latter case, either the frame has been correctly decoded (with probability P_{cd}) or it contains an undetectable error pattern (with probability P_e).

Figure 15. Events associated to RS decoding.

Obviously

$$P_{cd} + P_f + P_e = 1 \tag{28}$$

An analytical expression for P_e is available for RS codes [37], but its numerical complexity becomes impractical for large n (e.g. $n \geq 100$). Therefore, the approximations given in [38] are going to be used. Let $P_e(u)$ be the conditional probability of a decoder error given u symbol errors, and P_u the probability of having u symbol errors in the received word. Then,

$$P_e = \sum_{u=0}^{n} P_e(u) \cdot P_u \tag{29}$$

For $u \leq t_c$, always $P_e(u) = 0$, and [38] gives the following approximation:

$$P_e(u) \leq \frac{1}{t_c!} \text{ for all } u > t_c \tag{30}$$

If ideal outer interleaving is assumed, then the symbol errors are independent. Consequently,

$$P_u = \binom{n}{u} \cdot \varepsilon^u (1-\varepsilon)^{n-u} \tag{31}$$

where: ε is the RS symbol-error probability.

Combining (30) and (31) gives the probability of undetected error, P_e, as a function of the symbol error rate, ε. However, ε is not the common parameter used when it comes to compare different systems. A more meaningful and more commonly used parameter is the ratio of the energy per bit, E_b, to the noise-plus-interference power spectral density, I_o. This parameter was defined in Section 2.1.1 and was denoted E_b/I_o. Third generation CDMA cellular systems are of interest and CODIT is one example of proposed system for which a testbed has been realized [10][24]. An upper bound on ε as a function of E_b/I_o is given [24] for that system. This upper bound takes into account convolutional decoding using a soft-decision Viterbi decoder and assumes a Ricean fading multipath, which becomes Rayleigh fading in the worst case. Great detail of the physical layer will not be given because it is beyond the scope of this thesis. The following will be mentioned: the modulation used is coherent binary phase shift keying (BPSK), and coherent combining is performed by a RAKE receiver which knows perfectly the complex channel amplitudes and has exactly the same number of fingers as the number of paths in the multipath channel [24]. The bound on symbol errors at the output of the Viterbi decoder is given by equation (29) in [24]. This bound is shown in Figure 16 for a Rayleigh fading multipath channel with four paths. The convolutional code used is the NASA code with $k = 1$, $n = 2$, and $N_{cl} = 6$ [34]. The more interested reader should refer to [24] for more details about the physical layer that has been assumed. Finally, the analysis presented here does not take into account antenna diversity or the effect of power control. Therefore, the performance shown is conservative and corresponds to the case of fast small scale fading for which power control is usually ineffective.

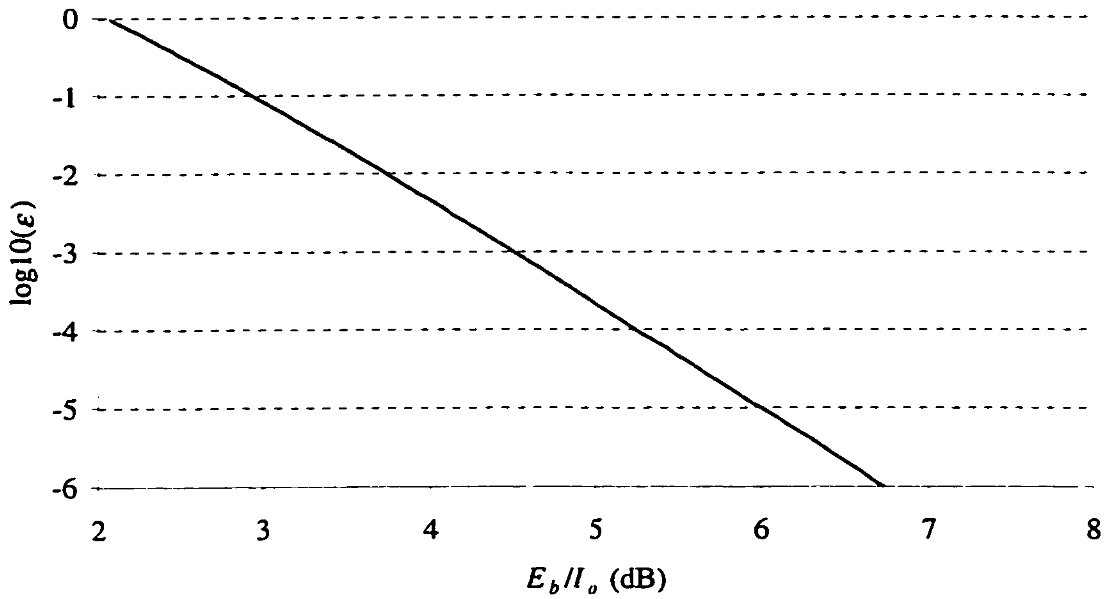

Figure 16. Upper bound on the symbol error probability for Rayleigh fading multipath (4 paths) channel [24].

Using the previous bound on ε, a bound on P_e can also be given as a function of the more meaningful E_b/I_o parameter, since P_e depends on ε. For the code used in the study, (n=100, k=80, b=8), the results shown in Figure 17 for P_e are obtained. As can be seen, the probability of undetected error, P_e, is very low; it is smaller than 10^{-6} for the entire range of E_b/I_o, and it decreases drastically as E_b/I_o increases above 2.75 dB.

Therefore, in the following, the delay analysis and the throughput analysis will not take into account P_e.

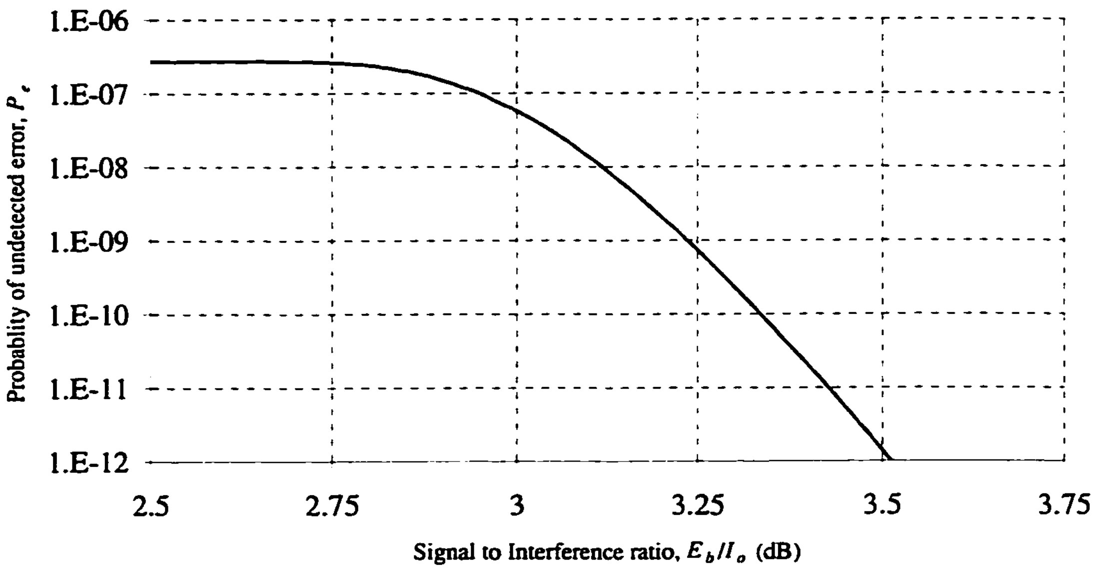

Figure 17. Probability of undetected error versus RS code symbol error rate.

Knowing that the RS code can correct up to t errors, the computation of P_{cd} is straightforward:

$$P_{cd} = \sum_{u=0}^{t} \binom{n}{u} \cdot \varepsilon^u (1-\varepsilon)^{n-u} \tag{32}$$

Once P_e and P_{cd} are known, P_f can be determined using (28). Using the same previous bound on ε as a function of E_b/I_o, a bound on P_f as a function of E_b/I_o can therefore be obtained. Figure 18 shows P_f versus E_b/I_o. Note that P_f is also the probability of frame retransmission, $p_{retx} = P_f$, because there is only one codeword per frame [11][33]. It can be seen from Figure 18 that for values of E_b/I_o greater than 3.2 dB (a very moderate value of SIR), the probability of frame retransmission is less than 10^{-2}. This low probability ensures a good throughput performance of the ARQ scheme as well as a good delay performance because retransmissions do not happen too frequently.

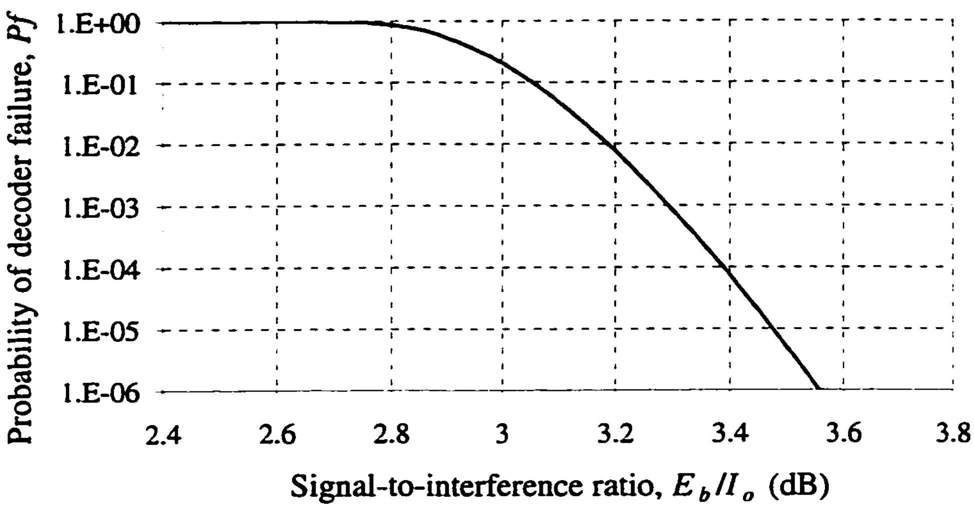

Figure 18. Probability of decoder failure versus RS code symbol error rate.

Finally, the reliability of a packet service is measured by the probability that a frame is accepted and contains an error, $P(E)$. This is given by [22]:

$$P(E) = P_e + P_e P_f + P_e P_f^2 + P_e P_f^3 + \ldots = \frac{P_e}{1 - P_f} \tag{33}$$

This probability can be used to find a bound on the BER, P_b. For accepted frames containing an error, the two extreme cases are either there is only one bit in error (out of kb bits), or all the kb bits of the frame are in error. Therefore, from [30]

$$\frac{1}{kb} P(E) \leq P_b \leq P(E) \tag{34}$$

As shown by Figure 19, for a value of E_b/I_o greater than 3.25dB, $P(E)$ is smaller than 10^{-9}. Therefore, according to (34), the BER is also smaller than 10^{-9}. Hence, a very high quality of data transmission service is ensured with a very moderate value of E_b/I_o.

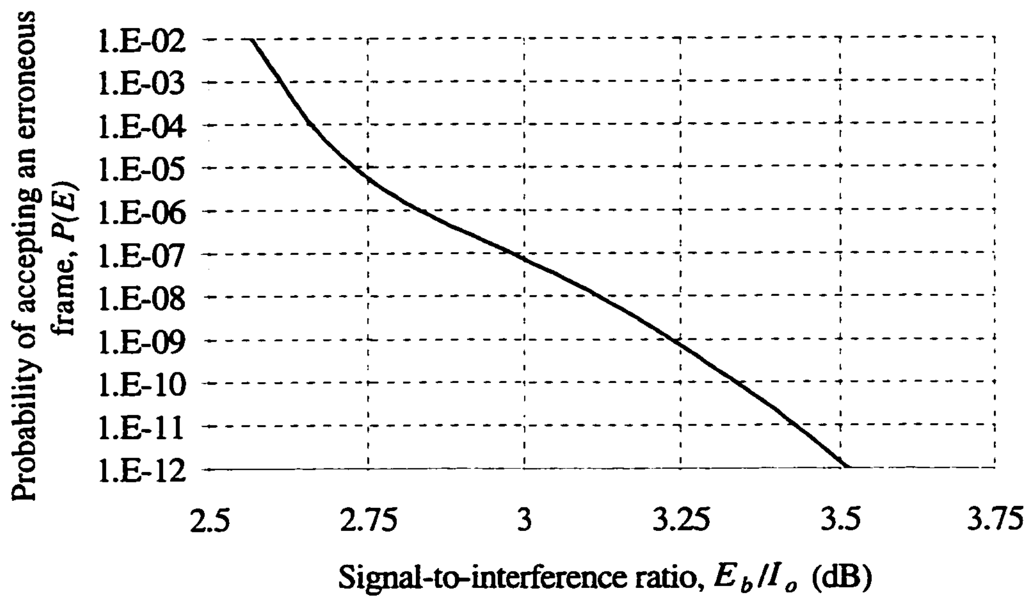

Figure 19. Probability of accepting an erroneous frame versus signal-to-interference ratio.

4.2. Delay analysis

The following delay analysis is the same as the one presented in [39]. The analysis is valid for the proposed new MAC scheme and the *conventional* scheme described in Section 2.4.1.3, in which transmission of the next packet is not allowed when the processing of the current one is still under way. The only difference in the analysis of the two MAC protocols are the expressions for the first and second moments of the respective transmission times. A comparison between the performance of the proposed MAC scheme and the conventional one is presented below.

4.2.1. The system under consideration

Before the analysis is presented, the system under consideration is going to be clearly defined. Let m be the number of MSs that are attached to the ARCH/PDTCH channel pair for packet transmission. Any arriving packets will be stored in MSs if they cannot be sent immediately upon arrival. Therefore every MS registered to the packet service will have its own queue in order to store waiting packets. It will be assumed that packets arrive at MSs according to a Poisson process, as defined in Section 2.5.4. Then if it is assumed that these processes are independent and have the same arrival rate, λ/m, then according to the property of Poisson processes mentioned in Section 2.5.4, the sum of the processes is also Poisson and has an overall rate of λ. Figure 20 shows the different queues of MSs, with arrival rate λ/m, and the overall queue, with rate λ.

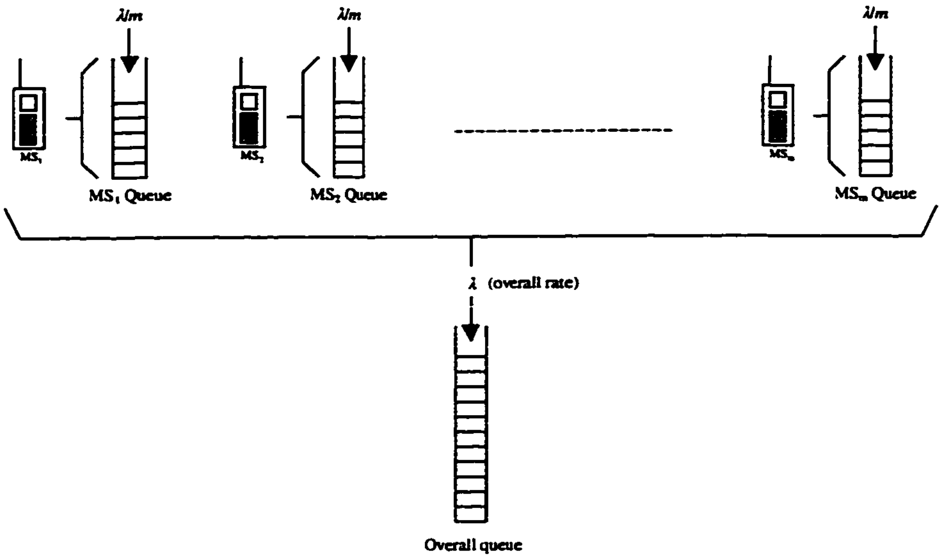

Figure 20. MS queues and overall queue.

The service time (or transmission time) of packets follows what is called a general distribution, as opposed to the Poisson distribution seen in Section 2.5.4. Due to the errors caused by the mobile channel, the transmission time will be random. The mean and second moment of this time will be of interest for the delay analysis.

Finally, only one server will be considered, the ARCH/PDTCH channel pair. In queueing theory, a *server* is an entity capable of performing the service required by the customer. In a packet-switched transmission case, the customers are the packets.

According to the previous description, the system under consideration is an *M*/*G*/1 system with reservation intervals [17]. The first letter indicates the nature of the arrival process. *M* stands for a memoryless (Poisson) process. The second letter indicates the nature of the probability distribution of the service time. It is *G* because the service time follows a general distribution as opposed to a memoryless distribution or a deterministic one. Finally the last letter indicates the number of servers. Recall that in the following analysis, there is only one ARCH/PDTCH channel pair.

4.2.2. Data period and access period

In order to organize transmission from several packet streams into the PDTCH, requests to access this channel are made by sending ARMs in a time-slotted fashion on the ARCH. Due to this time slotted structure, there could possibly be ARMs detected from several different MSs during the access subframe (Section 3.1.1). If such a case arises, then the BS chooses randomly which MS is to transmit and notifies its on the downlink of the PDTCH. The *packet transmission time or delay* is the interval of time corresponding to the actual transmission of a packet, as opposed to the MS access period defined in Section 3.2.

4.2.3. Access delay

Previously, the system was presented and now a few terms used for the analysis are going to be introduced. Let W_i be the waiting time in queue for packet no. i, X_i-the packet transmission time of the packet no. i, and N_i the number of packets found waiting in queue by the packet no. i upon arrival. R_i will be the residual service time seen by the packet no. i. This means that if packet no. j is already being served when packet no. i arrives, R_i is the remaining time until packet no. j's service is complete. Note that for this precise definition, the service time will include the access period. Finally, let Y_i be the sum of all the reservation intervals (or MS access periods) during which packet no. i must wait before being transmitted. Figure 21 illustrates all the previous parameters.

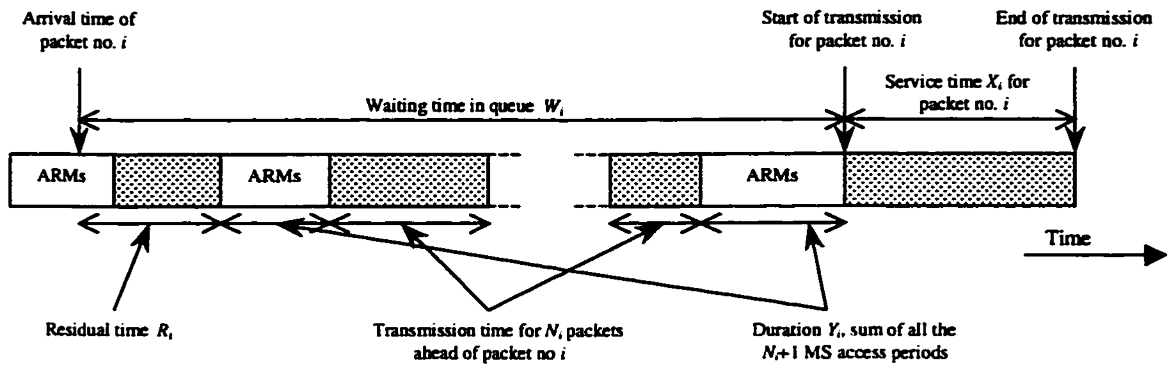

Figure 21. Derivation of the waiting time W_i.

Figure 21 shows that the waiting time, W_i, for packet i is:

$$W_i = R_i + \sum_{j=i-N_i}^{i-1} X_j + Y_i \tag{35}$$

The steady-state behaviour of these quantities is of interest. Therefore, taking the expectations of these quantities leads to

$$E\{W_i\} = E\{R_i\} + E\left\{\sum_{j=i-N_i}^{i-1} E\{X_j \mid N_i\}\right\} + E\{Y_i\}, \tag{36}$$

where $E\{X_j \mid N_i\}$ is the expectation of the random variable X_j given the occurrence of the random variable N_i (conditional expectation of X_j given N_i). If it is assumed that the variables N_i and $X_{i-N},\ldots,X_{i-1}$ are independent, then $E\{X_j \mid N_i\} = E\{X_j\}$ for $j \in [i-N_i, i-1]$. Therefore, the sum on the right hand side of (36) is given by: $\sum_{j=i-N_i}^{i-1} E\{X_j \mid N_i\} = \sum_{j=i-N_i}^{i-1} E\{X_j\}$. If it is now assumed that the random variables $(X_1, X_2, \ldots)$ are identically distributed and mutually independent, then the last sum becomes $\sum_{j=i-N_i}^{i-1} E\{X_j\} = \sum_{j=i-N_i}^{i-1} E\{X\}$. At this stage, the elements of the sum do not depend on N_i anymore, therefore $\sum_{j=i-N_i}^{i-1} E\{X\} = N_i E\{X\}$. Consequently, the sum becomes $E\left\{\sum_{j=i-N_i}^{i-1} E\{X_j \mid N_i\}\right\} = E\{X\}E\{N_i\}$ and therefore (36) becomes:

$$E\{W_i\} = E\{R_i\} + E\{X\}E\{N_i\} + E\{Y_i\} \tag{37}$$

Denote by $\overline{X}$ the expectation of the random variable X, $\overline{X} = E\{X\}$. As mentioned in Section 2.5.2, systems that reach a steady state are of interest. Therefore, it will be assumed that the following limits exist: R, the *mean residual time* defined as $R = \lim_{i \to \infty} E\{R_i\}$; W, *the mean access delay* (also called the mean waiting time in queue) defined as $W = \lim_{i \to \infty} E\{W_i\}$; N, the mean number of packet in queue defined as $N = \lim_{i \to \infty} E\{N_i\}$; and Y, the mean duration of the entire sum of the MS access periods defined as $Y = \lim_{i \to \infty} E\{Y_i\}$. Because of the Poisson character of the arrival process, the residual time, R, and the average number of packets in queue, N, seen by an arriving packet are equal to the mean residual time and mean number of packets in queue [17]. Therefore, taking the limit as i goes to infinity in (37), and using the previous definitions, leads to

$$W = R + N\overline{X} + Y \tag{38}$$

R will be determined using a graphical argument as it is done in [17]. The instantaneous residual service time of a packet, $r(\tau)$, is going to be used. This time is the remaining time for completion of the packet being sent at the time τ. Figure 22 illustrates the residual service time as a function or time, τ.

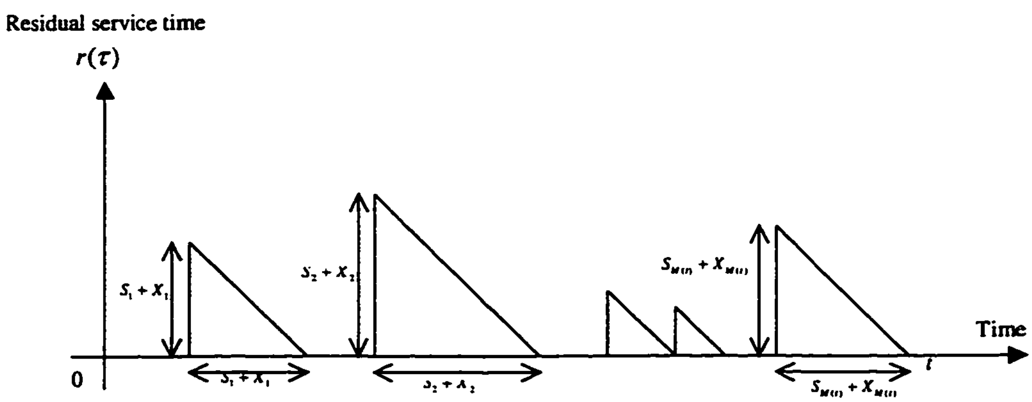

Figure 22. Residual service time of packets.

Note that when transmission of a new packet of duration X starts, $r(\tau)$ has the value $S+X$, where S is the access period duration needed for the transmission. The instantaneous residual time decays then linearly for $S+X$ time units. If a time t for which $r(t)=0$ is considered, the time average of the residual time in the interval $[0,t]$ is

$$\frac{1}{t}\int_0^t r(\tau)d\tau = \frac{1}{t}\sum_{k=1}^{M(t)} \tfrac{1}{2}(S_k + X_k)^2 \tag{39}$$

where $M(t)$ is the number of packets sent within $[0,t]$, S_k is the access period of packet no. k, and X_k is the transmission time of packet no. k. The previous equation can also be rewritten as:

$$\frac{1}{t}\int_0^t r(\tau)d\tau = \frac{1}{2}\frac{M(t)}{t}\frac{\sum_{k=1}^{M(t)}(S_k + X_k)^2}{M(t)} \tag{40}$$

If the sum of squares on the right side of (40) is expanded, if the limit as t goes to infinity is taken, and if it is assumed that the limits below exist, this equation becomes

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t r(\tau)d\tau = \frac{1}{2}\lim_{t\to\infty}\frac{M(t)}{t}\cdot\lim_{t\to\infty}\left(\sum_{k=1}^{M(t)}\frac{S_k^2}{M(t)} + \sum_{k=1}^{M(t)}\frac{X_k^2}{M(t)} + 2\sum_{k=1}^{M(t)}\frac{S_k\cdot X_k}{M(t)}\right) \tag{41}$$

Define the mean residual time as: $R = \lim_{i\to\infty} E[R_i]$ where R_i is the residual time seen by the packet no. k. It is assumed, as in [17], that the long-term time average of the residual time, which is the left side of (41), is equal to the mean residual time, R. Therefore, the left side of (41) becomes $\lim_{t\to\infty}\frac{1}{t}\int_0^t r(\tau)d\tau = R$. As for the right side of (41), it can be seen that the first limit is equal to the departure rate as defined in (14), $\delta = \lim_{t\to\infty}\frac{M(t)}{t}$. However, for a system in equilibrium, the departure rate equals the arrival rate, that is $\delta = \lambda$, therefore $\lim_{t\to\infty}\frac{M(t)}{t} = \lambda$. For the other sums, the ergodicity of the stochastic processes mentioned in Section 2.5.3 leads to $\lim_{t\to\infty}\sum_{k=1}^{M(t)}\frac{S_k^2}{M(t)} = \overline{S^2}$, which is the second moment of the MS access period, and leads to $\lim_{t\to\infty}\sum_{k=1}^{M(t)}\frac{X_k^2}{M(t)} = \overline{X^2}$, which is the second moment of the transmission time (or service time). Finally, if it is assumed that MS access periods and transmission times are independent, $\lim_{t\to\infty}\sum_{k=1}^{M(t)}\frac{S_k \cdot X_k}{M(t)} = \overline{S}\cdot\overline{X}$, where $\overline{S}$ is the average MS access periods, and $\overline{X}$ is the average transmission time of packets. Substituting the above identities into (41) gives

$$R = \tfrac{1}{2}\lambda(\overline{X^2} + \overline{S^2} + 2\cdot\overline{X}\cdot\overline{S}). \tag{42}$$

Note that $\overline{S}$ and $\overline{S^2}$ depend on the type of synchronizer used.

According to Little's theorem from queuing theory (in Section 2.5.2), $N = \lambda W$. Even if the number of packets that a packet arriving at a given instant must wait for is not equal to the number of packets already awaiting service at different mobiles at that time, the arriving packet sees an average number of transmitted packets before it given by Little's theorem. This is because the order in which packets are served is independent of their transmission times. Thus,

$$Y = \lambda W \overline{S} \tag{43}$$

Note that in this equation the last MS access period has been neglected, $N = \lambda W >> 1$, because it is negligible for high arrival rates, which are of interest. All the previous relations are substituted into (38) in order to solve (38) for *W*. W will be given by

$$W = \frac{\tfrac{1}{2}\lambda(\overline{X^2} + \overline{S^2} + 2\cdot\overline{X}\cdot\overline{S})}{1-\lambda(\overline{X}+\overline{S})} \text{ valid for } \lambda \le \frac{1}{\overline{X}+\overline{S}} \tag{44}$$

The validity range of the previous equation is due to the fact that a separate MS access period of average length $\overline{S}$ is needed for each packet transmission, thereby effectively increasing the average time dedicated to a packet from $\overline{X}$ to $\overline{X}+\overline{S}$.

The mean delay of a packet, T, is the sum of its mean access delay, W, and its mean transmission time, $\overline{X}$. Thus

$$T = \overline{X} + \frac{\tfrac{1}{2}\lambda(\overline{X^2} + \overline{S^2} + 2\cdot\overline{X}\cdot\overline{S})}{1-\lambda(\overline{X}+\overline{S})} \tag{45}$$

4.2.4. Transmission time

Now, expressions for the means and the second moments of the packet transmission time for the conventional and the proposed new MAC scheme need to be found. Let p_{retx} be the probability of frame retransmission, T_f the duration of a frame (10 ms), n_{ts} the number of frames in the trailing segment (corresponding to the processing delay), n_{pc} the average number of frames needed to re-establish closed-loop power control and re-synchronize the PN code after a silent period, n_s the average number of ARMs needed for an MS to acquire synchronization ($\overline{S} = n_s \cdot T_f$), and n_p the number of frames needed to transmit a packet if no retransmission is required.

4.2.4.1. Conventional scheme

In order to simplify the analysis, it will be assumed that, in the conventional scheme, all the frames of a given packet are delivered without retransmissions. In other words, the ARQ scheme will not be taken into account for the analysis of the conventional scheme as it is done in [29]. Therefore the packet transmission time is deterministic, and $\overline{X_C}$, the mean packet transmission time of the conventional MAC scheme is

$$\overline{X_C} = (n_p + n_{ts})T_f \tag{46}$$

Using this deterministic argument, it is also possible to find $\overline{X_C^2}$, the second moment of the packet transmission time of the conventional scheme, which is equal to the square of $\overline{X_C}$, *i.e.*

$$\overline{X_C^2} = (n_p + n_{ts})^2 T_f^2 . \tag{47}$$

4.2.4.2. New scheme

For the proposed new scheme, the same assumption as before is kept for the leading segment, that is, frames of the leading segment of a packet do not need retransmissions. However, for trailing segments, possible retransmissions will be taken into account. If a trailing segment (TS) needs to be retransmitted, it is only retransmitted after another entire packet is transmitted (see Figure 14). Therefore, the delay introduced by that packet's transmission can be significant and should be taken into account in the analysis. Let $\overline{X_N}$ be the average packet transmission time and $\overline{X_N^2}$ -the second moment of the packet transmission time under the new MAC scheme. In order to simplify the analysis, it will be assumed that whenever there are errors in the trailing segment, the whole segment is retransmitted, and not just the frames containing errors.

The analysis is made easier if it is assumed that packets have a fixed size (n_p is constant). Assuming that frame retransmissions are independent, the probability that a trailing segment needs a retransmission is

$$p_{ts} = 1-(1-p_{retx})^{n_{ts}}. \tag{48}$$

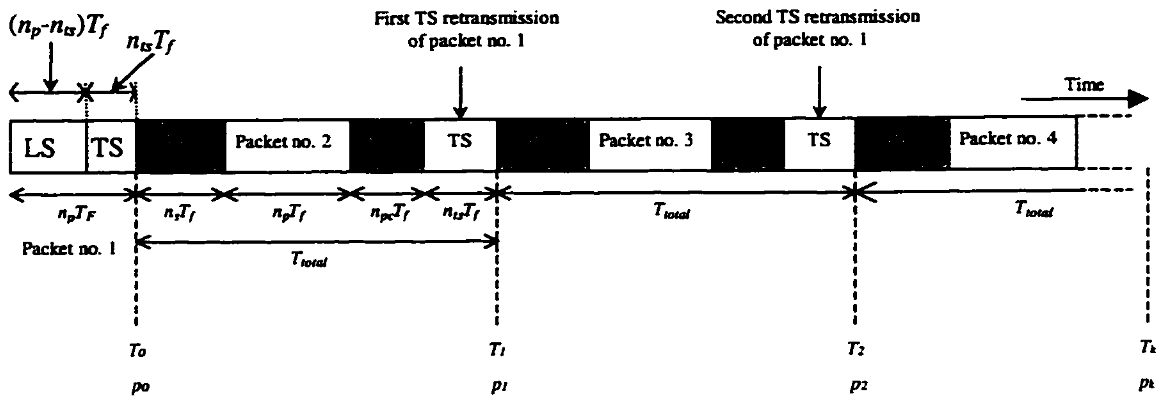

Figure 23. Computation of the average transmission time.

The average transmission time will be derived according to the example in Figure 23. The packet under consideration is packet no. 1. Let T_k be the time at which the transmission of packet no. 1 ends with probability p_k for $k = 0,1,\ldots$. Using Figure 23, it can be seen that the time T_k can be expressed as

$$T_k = n_pT_f + kT_{total}, \tag{49}$$

where $T_{total} = (n_p + n_s + n_{pc} + n_{ts})T_f$. And the following expression is obtained for p_k:

$$p_k = (1-p_{ts})p_{ts}^k \tag{50}$$

where p_{ts} is given by (48). Using the previous quantities, the average transmission time $\overline{X_N} = \sum_{k=0}^{\infty} T_k p_k$ of the new MAC scheme can be computed. It is given by

$$\overline{X_N} = (1-p_{ts})\left[n_p T_f \sum_{k=0}^{\infty} p_{ts}^k + T_{total}\sum_{k=0}^{\infty} k p_{ts}^k\right]. \qquad (51)$$

Using the following well-known results; $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$ and $\sum_{k=0}^{\infty} k x^k = \frac{x}{(1-x)^2}$, $\overline{X_N}$ becomes

$$\overline{X_N} = n_p T_f + T_{total}\frac{p_{ts}}{1-p_{ts}}. \qquad (52)$$

The second moment of the transmission time is defined as $\overline{X_N^2} = \sum_{k=0}^{\infty} T_k^2 p_k$. Using (49), (50), and the identity $\sum_{k=0}^{\infty} k^2 x^k = \frac{x+x^2}{(1-x)^3}$ (because of the square in the definition of $\overline{X_N^2}$), one can easily show that

$$\overline{X_N^2} = (n_p T_f)^2 + 2(n_p T_f)T_{total}\frac{p_{ts}}{1-p_{ts}} + T_{total}^2\frac{p_{ts}(1+p_{ts})}{(1-p_{ts})^2} \qquad (53)$$

4.3. Throughput analysis

The throughput efficiency of a MAC scheme is defined as the ratio of the number of frames successfully delivered to the overall number of transmitted frames. In a conventional MAC scheme, the overall number of transmitted frames includes first transmissions and possible retransmissions, while for the new scheme proposed, it also includes the frames used to restart the power control procedure for each of the trailing segment retransmissions. This is shown on Figure 23: for each of the TS retransmissions of packet no. 1, a "P.C" procedure of n_{pc} frames is required. In the following analysis, the following will be assumed: a receiver buffer (at the BS) of infinite length and a negligible probability of undetected error in the RS decoder (consistently with the analysis in Section 4.1).

The throughput efficiency of the conventional MAC scheme, η_c, is the same as that of the selective repeat ARQ scheme described in [22]. It is given by

$$\eta_c = 1 - p_{retx}. \qquad (54)$$

For the proposed new MAC scheme, the average number of transmitted frames used to successfully receive the leading segment and the trailing segment must be determined. The number of frames in the leading segment is $n_p - n_{ts}$, and in the trailing segment the number is n_{ts}. Note that the frames of the leading segment follow a selective repeat ARQ scheme, and therefore, the average number of frames needed to successfully receive the leading segment, N_{ls}, is given by [22]:

$$N_{ls} = \frac{n_p - n_{ts}}{1 - p_{retx}} \tag{55}$$

For the trailing segment, the same assumptions as the ones used for the computation of the average transmission time will be made: whenever there are errors in the trailing segment, the whole segment (containing n_{ts} frames) must be retransmitted, and not just the frames containing errors. This assumption will lead to a conservative value of the throughput, and therefore it will be used as an upper bound. Using Figure 23, it is possible to count how many frames have been transmitted every time a TS needs to be retransmitted. It can be seen that at time T_k for $k = 0,1,\ldots$, the number of frames that have been transmitted is

$$N_k = (k+1)n_{ts} + kn_{pc}. \tag{56}$$

The average number of frames needed to successfully receive the trailing segment is $N_{ts} = \sum_{k=0}^{\infty} N_k p_k$. Substituting the values of (50) and (56) into the previous definition of N_{ts} leads to

$$N_{ts} = (1 - p_{ts})\left[n_{ts} \sum_{k=0}^{\infty} (k+1) p_{ts}^k + n_{ps} \sum_{k=0}^{\infty} k p_{ts}^k \right] \tag{57}$$

By changing the starting indices of the first sum of the right hand side and using the identity $\sum_{k=0}^{\infty} kx^k = \frac{x}{(1-x)^2}$ for both of those summations, the following result can be derived:

$$N_{ts} = \frac{n_{ts} + n_{pc} p_{ts}}{1 - p_{ts}} \tag{58}$$

Using (55) and (58), the throughput efficiency of the overall scheme, η_N, can be computed. It is given by:

$$\eta_N = \frac{n_p}{N_{ls} + N_{ts}} \tag{59}$$

4.4. Data activity factor

In interference-limited systems, such as DS-CDMA systems, it is useful to know how much interference the pair PDTCH/ARCH causes in order to determine its effect on the capacity of the whole system. The data activity factor of a channel is defined as the fraction of time during which it is active. A data transmission channel is active when data transmission is under way. Hence the concept of data activity factor is similar to that of voice activity factor described in [19]. Note that when the ARCH is active, the PDTCH is not and vice-versa. Therefore, from an interference point of view, the following analysis treats the pair PDTCH/ARCH as only one channel.

Once again, Little's theorem will be used to determine the data activity factor like it is done in [17]. The left side of Figure 24 shows how Little's theorem is applied to the queueing system within the dotted ellipse. The system consists of a queue with an average of N packets waiting in queue for W units of time with packets arriving at a rate λ. This theorem can also be applied to the PDTCH/ARCH pair within the dotted ellipse on the right side of this figure. On one hand, the departure rate, δ, is equal to the arrival rate of the queueing part, λ (see Section 2.5.2) and on the other hand, the departure rate for the queueing part is equal to the arrival rate of the PDTCH. The transmission time of packets, $\bar{A}$, is equal to the time that packets stay in the PDTCH. Therefore, this time can be seen as the time packets need to wait for before leaving the PDTCH. Thus Little's theorem can be applied to that part of the system as shown on Figure 24 by the other dotted ellipse. The quantity $\alpha = \lambda \cdot \bar{A}$ is therefore the average number of packets in the ARCH/PDTCH.

α is what is called the *data activity factor* of the pair ARCH/PDTCH (similar to N in the previous queue).

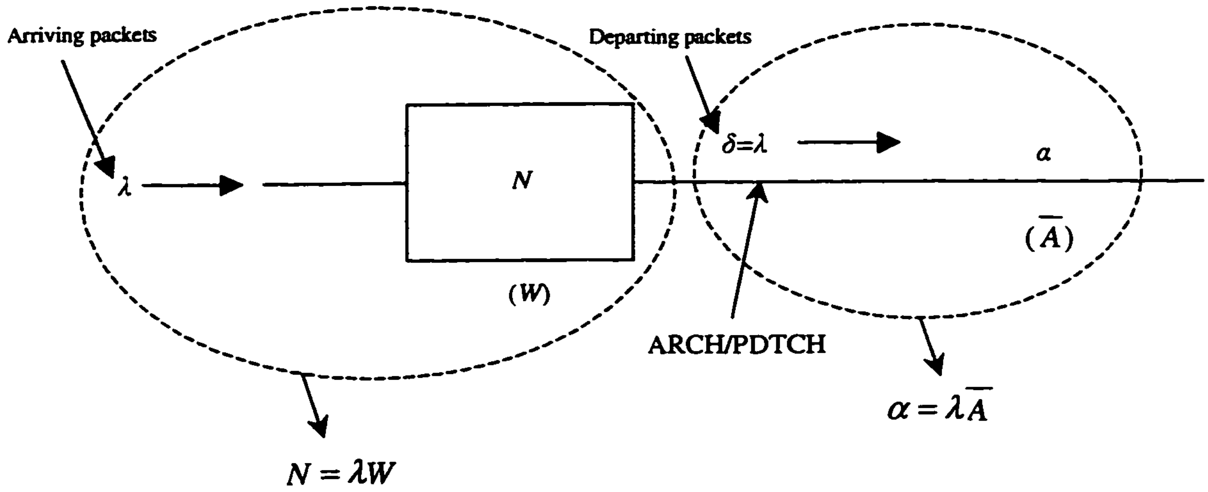

Figure 24. Little's theorem applied to the PTDCH.

The theorem gives the data activity factor of this pair as

$$\alpha=\lambda\overline{A} \tag{60}$$

For the proposed new scheme, the average duration of packet transmission on the pair PDTCH/ARCH is the sum of the mean packet transmission time, $\overline{X_N}$, and the MS access period, $\overline{S}$. Therefore the data activity factor of the new scheme is

$$\alpha_N=\lambda(\overline{S}+\overline{X_N}) \text{ valid for } \lambda\leq\frac{1}{\overline{X_N}+\overline{S}} \tag{61}$$

For the conventional scheme, the channel is not active all the time, even when it is dedicated to a MS. Indeed, the channel is not active when MSs wait for the time corresponding to the acknowledgements from the BS, $n_{ts}T_f$. In this case, the data activity factor, α_C is

$$\alpha_C=\lambda(\overline{S}+\overline{X_C}-n_{ts}T_f) \text{ valid for } \lambda\leq\frac{1}{\overline{X_C}+\overline{S}} \tag{62}$$

4.5. Conclusion

Performance analysis of the new MAC scheme was presented. First, the error detection performance of the concatenated coding scheme was given. This performance was obtained using tight upper bound on the RS symbol error probability at the output of a Viterbi decoder discussed in [24]. Then the delay analysis of the new and the conventional MAC scheme was presented. Throughput as well as data activity factor analyses of the pair ARCH/PDTCH were finally given.

Chapter 5. Simulation results

5.1. Framework of the simulations

In order to check the accuracy of the previous analytical models, the two MAC protocols were simulated. MSs attached to an ARCH/PDTCH pair were simulated. Packets arrive at MSs following a Poisson distribution. Each of the MSs is assumed to have an infinite size buffer in order to store these packets. Storage of these packets before transmission will lead to queues in the MSs that are called *MS queues*. At MS queues, packets are sent on a first-come first-served basis. Packets have either a fixed size or their size varies according to some particular distribution (*e.g.* email packet size distribution), depending on the type of simulation. As soon as a given MS has a packet in its queue, and when it is allowed by the BS, it will try to access the PDTCH by sending requests on the ARCH.

System-level simulations have been performed with the time divided into frame-length units. The system starts with an idle state for which all the MS queues are empty. Then after a certain period of time, packets will arrive at MS queues. MSs with packets in the queue will send ARMs on the ARCH. The system is said to be in the *access state*. After one or several ARM(s) is (are) detected and a MS has been chosen to transmit over the PDTCH, all other MSs that have packets in their queues must wait until the actual transmission is over in order to compete for the PDTCH. The system is said to be in the *transmission state*. Therefore, there are three possible states for the system during the simulations. There is an alternation of access states followed by transmission states and possibly *idle states* if no packets are present in MS queue.

As far as the link is concerned, packets are not converted to bits that are sent on a fading channel. Packets are just broken into frames and then "transmitted". The physical layer causes frames to be received in error and introduces delay for negative acknowledgements (NACKs) of frames corresponding to the processing delay of frames. For the probability of frame retransmission, p_{retx} (equal to the frame error rate, FER), the value taken is 0.01 (see Section 4.1). This value is justified by the analysis included in Section 4.1. The BER target is 10^{-9} and a value of E_b/I_o that leads to this is 3.25dB according to Figure 19. In Figure 18, a value of E_b/I_o equal to 3.25dB leads to a frame retransmission probability less than 0.01. The number of frames corresponding to a processing delay is 4: one for coherent detection described in [10], two for interleaving/de-interleaving of each frames, and one for both convolutional decoding and RS decoding [22]. Although the error control coding (convolutional code and RS code) has been programmed, the coding of the frame bits has not been performed in order to save simulation time. Moreover the probability of undetected errors for the RS decoder is so low that roughly 10^7 frames would have to be simulated in order to see a single undetected error occur (see section 4.1).

The return channel used by the NACKs is assumed to be error-free. The receiver and transmitter buffers for the ARQ scheme are also assumed to be infinite. It is possible to use this scheme with finite buffers as described in [22]. Then the scheme becomes more complex because of the need to handle buffer overflows and it requires an extra buffer (of the same length) per PDTCH/ARCH pair set up at the BS in order to keep track of the trailing segment.

Regarding the ARM detection, it was previously assumed that a fast synchronizer as described in [20] is used. If F_c is equal to 4.096Mchips/s, T_{access} is equal to 5ms, and N_{slot} is equal to 16, then L is equal to 10. In order to have the probability of false synchronization negligible, P_f can be taken to be smaller than 10^{-6}. Using a numerical analysis in (25), θ_d is found to be equal to 32.711. Once θ_d is found, it is then possible to get the probability of correct acquisition in a Rayleigh fading environment using (26).

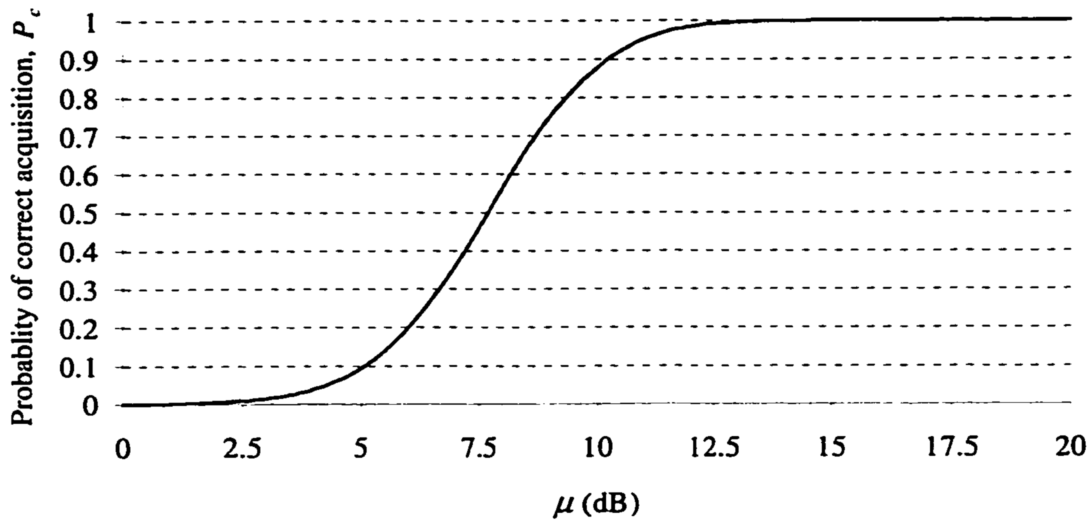

Figure 25. Probability of correct acquisition versus signal-to-noise ratio.

According to Figure 25, in order to have a probability of correct acquisition of a least 0.99, the average value of $\overline{\mu}$ must be greater than 12.4 dB. Using this value, it is possible to determine the upper bound on the average number of ARMs needed to acquire synchronization as given by (27)

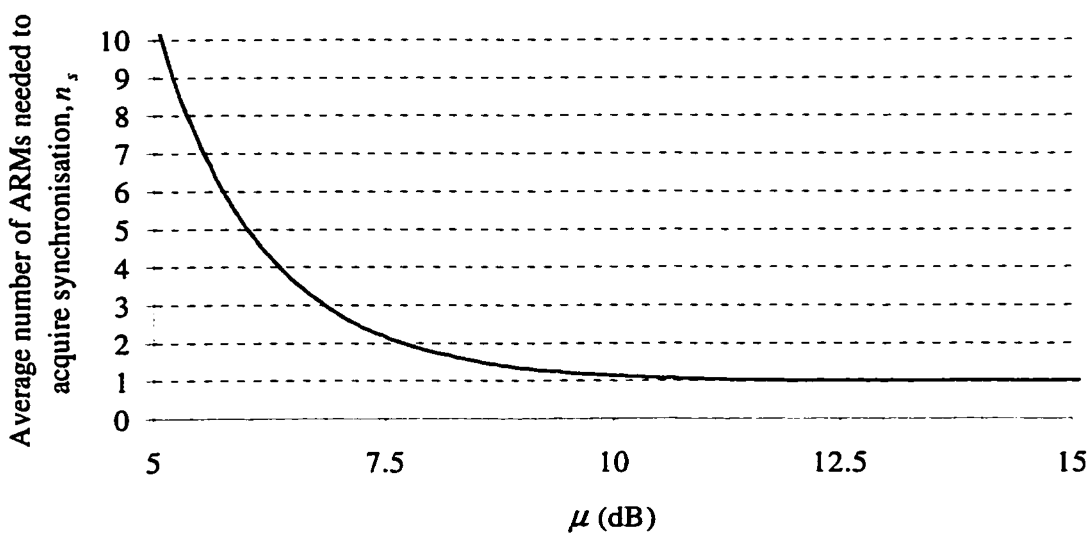

Figure 26. Upper bound on the average number of ARMs needed for PN code synchronization.

From Figure 26, one can see that the number of ARMs needed for code synchronization is less than $1/P_c \cong 1.01$ at $\overline{\mu} = 12.4$ dB. Considering this high probability, the number of frames that are needed to re-synchronize the PN code and re-adjust the power level for a TS retransmission, n_{pc}, was assumed to be one.

The bit rate used in the simulation is 64kbps, and 16 MSs are attached to the PDTCH. 10000 packets have been simulated in order to obtained credible results.

5.2. Mean packet delay

5.2.1. Effect of packet size

Figure 27 shows the mean packet delays of the conventional scheme and the new scheme. The figure demonstrates that the analysis and the simulation results are close for a FER equal to 0.01. Given a mean packet delay, Figure 27 shows that the new MAC scheme allows the system to handle a higher arrival rate than a conventional scheme for packet size of 0.5kbyte and 1kbyte. This improvement increases with decreasing packet size. Figure 28 shows that, for a given mean packet delay, the improvement in arrival rate is small for 2kbyte packets and almost none for 4 kbyte packets. This is not surprising, since for longer packets the processing time at the receiver is negligible compared to the transmission time. However, for shorter packets the processing delay is comparable to the packet transmission time, and the new MAC scheme improves performance significantly. For instance, if the maximum tolerable packet delay is 1s and packets have a size of 1kbyte, then the new MAC protocol improves the maximum arrival rate by 30 percent compared to a conventional MAC protocol. Note that in this case, the maximum arrival rate is the arrival rate for which the delay is 1s. In the case that the packet size is 0.5kbytes this improvement is about 50 percent. On the other hand, if the packet size is increased, this improvement is reduced. For 2kbyte packets, the improvement is little less than 20 percent and for 4kbyte packets there is only marginal improvement.

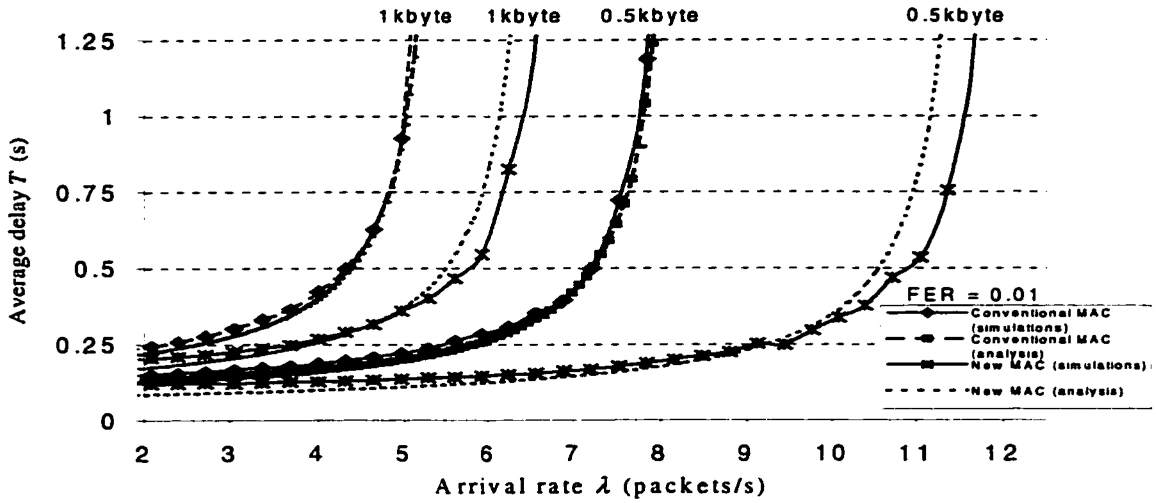

Figure 27. Mean packet delay vs. arrival rate for 0.5kbyte and 1kbyte packets.

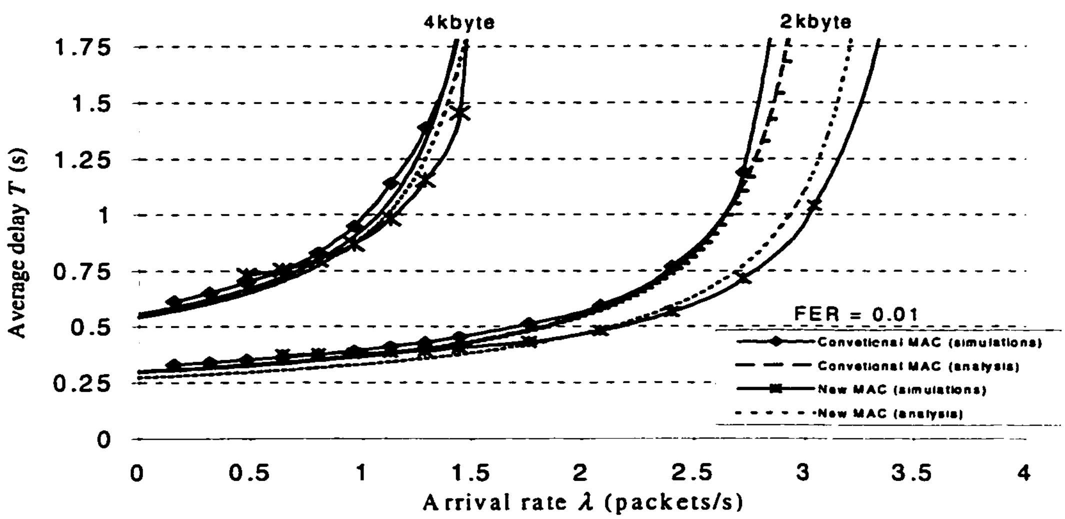

Figure 28. Mean packet delay vs. arrival rate for 2kbyte and 4kbyte packets.

5.2.2. Specific packet size distributions

5.2.2.1. E-mail packet size distribution

In order to see how the new MAC protocol performs with a specific application, it was also simulated with the size of packets following a typical e-mail packet size distribution [29], see Figure 29. The mean packet size for this distribution is 2kbytes.

The results are given in Figure 30 for a FER equal to 0.01. This figure shows an improvement of about 15 percent if a maximum delay of approximately 5s is considered when using the new MAC. The improvement is not very significant because of the mean size of the packets. As shown with Figure 28, the improvement with 2kbyte packets is about 20 percent. Therefore, in the situation where sent packets have variable sizes, it is better to use this new MAC scheme when the transmission time corresponding to the mean packet size is not too big compared to the processing delay at the BS.

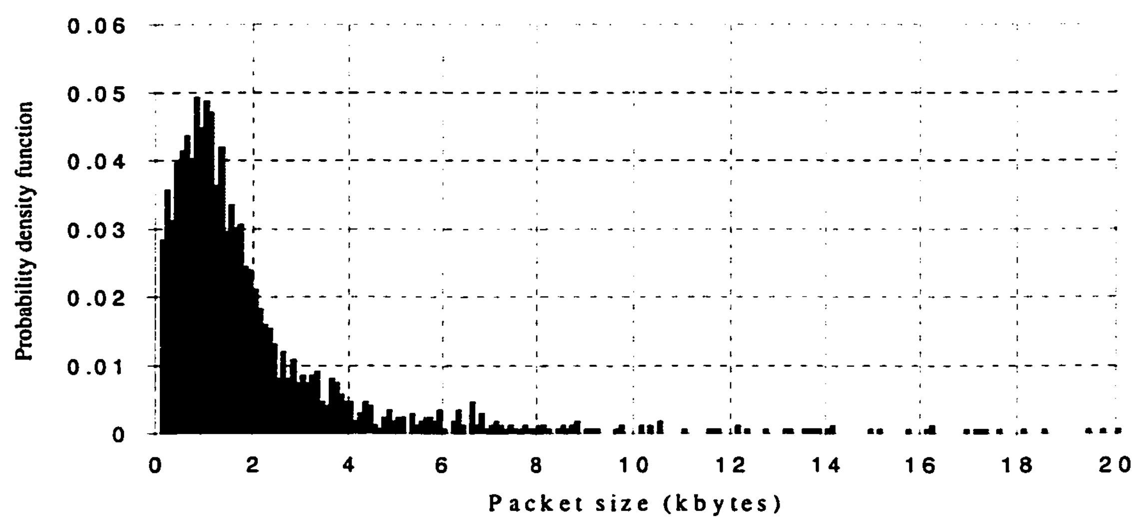

Figure 29. Probability density function of e-mail packet sizes [29].

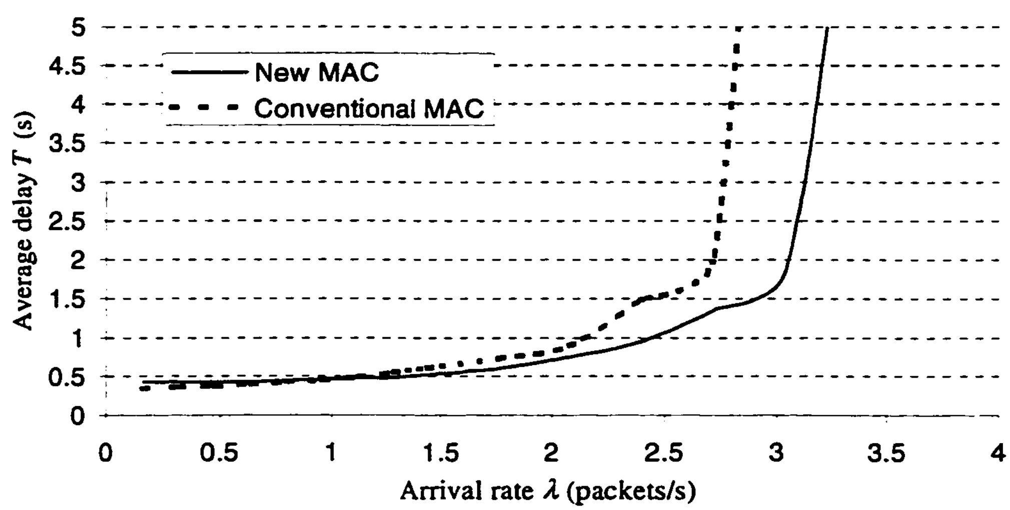

Figure 30. Mean packet delay for the e-mail distribution of Figure 29.

5.2.2.2. Typical Internet packet size distribution

A typical packet size distribution seen in the Internet [40] was also used in the simulations of the two MAC protocols. The distribution is shown on Figure 31. This figure shows that the distribution is characterized by a large probability of having very short packets. These short packets of about 40 bytes correspond to transmission control protocol (TCP) acknowledgements. There is also a large proportion of packets of size 576 bytes or slightly shorter that correspond to the smallest size of packets that must be supported on the Internet [15]. Note that the probability density function given in [40] was truncated for packets larger than 1500 bytes. The reason for this is that these larger packets are Ethernet packets. These packets are not sent by MSs but by local area networks. The mean of the packet size distribution of Figure 31 is 400 bytes

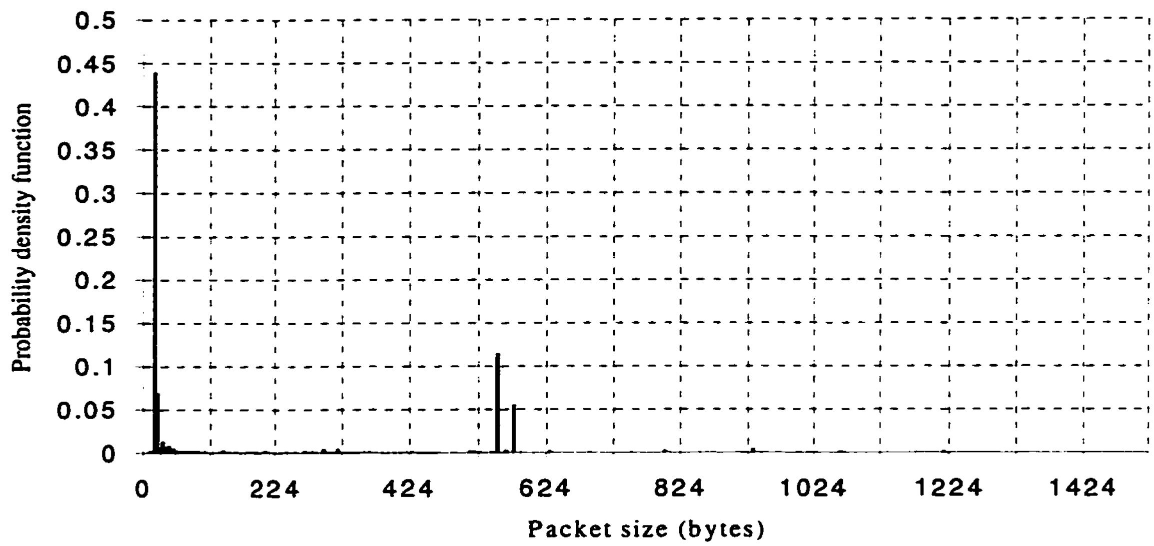

Figure 31. Probability density function of typical Internet packet sizes [40].

The performance of the new MAC protocol and the conventional one is shown on Figure 32. The new MAC protocol improves significantly the maximum arrival rate that the PDTCH/ARCH can handle compared to the conventional protocol. For a maximum delay of 5s, the new MAC protocol improves the arrival rate by about 60 percent. This is due to the vast majority of short packets for which the new MAC protocol is very efficient, as seen previously with 0.5kbyte packets.

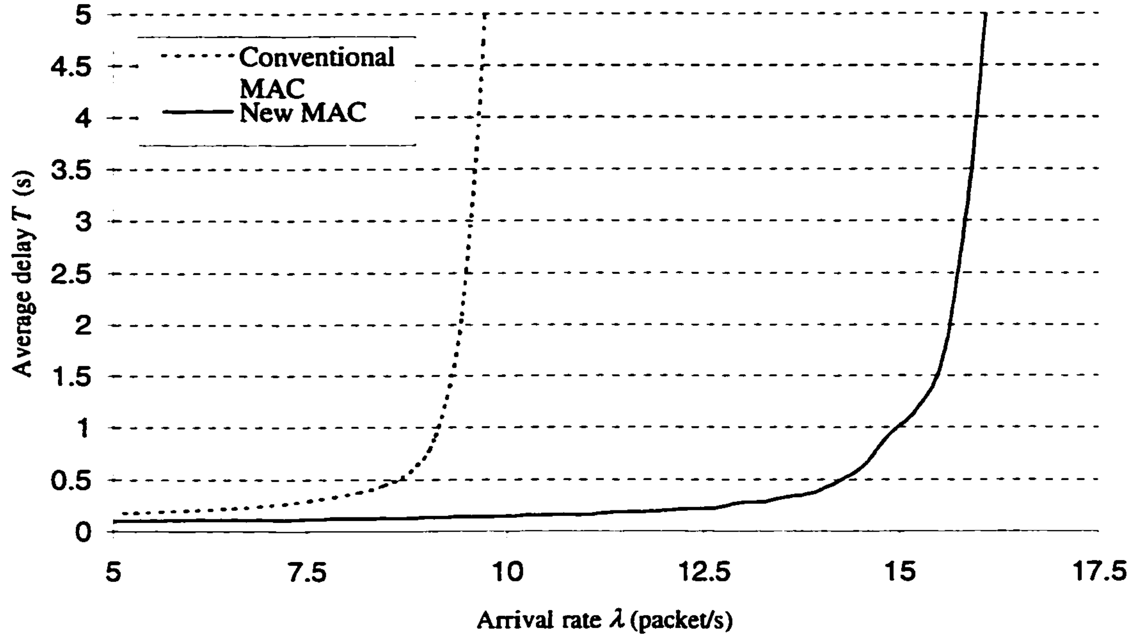

Figure 32. Average packet delay for the typical Internet packet size distribution of Figure 31.

5.2.3. Effect of frame error rate

In order to demonstrate the robustness of the previous discussed MAC protocols, these protocols were also simulated with different frame error rates for 0.5kbyte packets. First note that, for the conventional MAC protocol, the analysis and the simulations give very close results as shown in Figure 33 for a frame error rate (FER) ranging from 0.01 to 0.1.

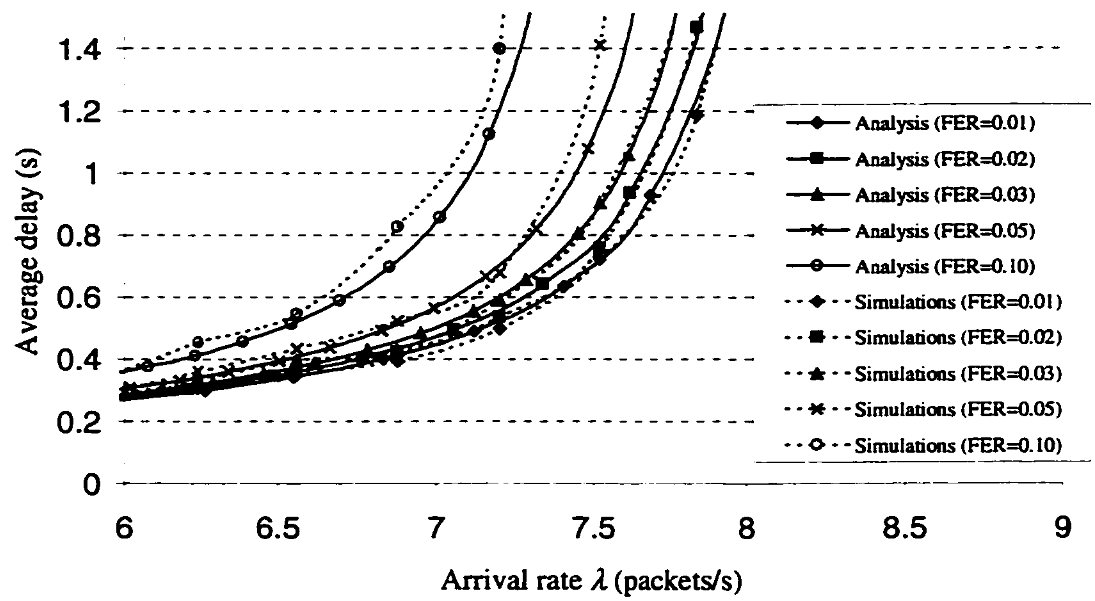

Figure 33. Average delay versus arrival rate as a function of FER for the conventional MAC protocol

Note that as the FER increases, the analysis becomes less accurate with respect to the simulations. This is particularly true for FERs greater or equal to 0.05. The reason for this is that the approximation that was made is Section 4.2.4.1, assuming that packets are sent without retransmissions, is less accurate. The analysis tends to overestimate the arrival rate that the MAC protocol can handle. Nevertheless, this overestimation is small and always smaller than two percent for the entire range of the arrival rates. As for low FERs (smaller than 0.03), the previous approximation is fairly accurate. Thus, if a two-percent error is tolerated, it can be said that the analysis of the conventional MAC protocol gives the same results as the simulations.

For the new MAC protocol, the results are shown in Figure 34. The general trend is that the analysis tends to underestimate the arrival rate. For an FER equal to 0.01 the analysis tends to underestimate the arrival rate by less than 5 percent (at 1s of maximum tolerable delay). For higher FERs, ranging from 0.02 to 0.05, this underestimation becomes bigger and ranges from 8 percent to 25 percent (also at 1s). The explanation for this is that the approximation in Section 4.2.4.2 is very conservative. As a matter of fact, for high FERs (greater than 0.02), assuming that the whole TS is retransmitted when single TS frame is received in error, is very pessimistic. Indeed a high FER leads to a very high probability of TS retransmission, which in turn leads to an increase in the evaluation of the packet transmission time, and therefore poor delay performance (see Section 4.2.4.2). As a result, for FERs greater than 0.02, the analysis of the new MAC (Section 4.2) gives conservative results but they are not very accurate anymore. Thus it is better to use simulations to assess the performance of the new MAC protocol for high FERs. In any case, it can be seen by comparing Figure 33 and Figure 34 that the performance of the new MAC protocol with FER equal to 0.05 is still much better than that of the conventional MAC protocol with FER equal to 0.01. This demonstrate the robustness of the new MAC protocol under high FER.

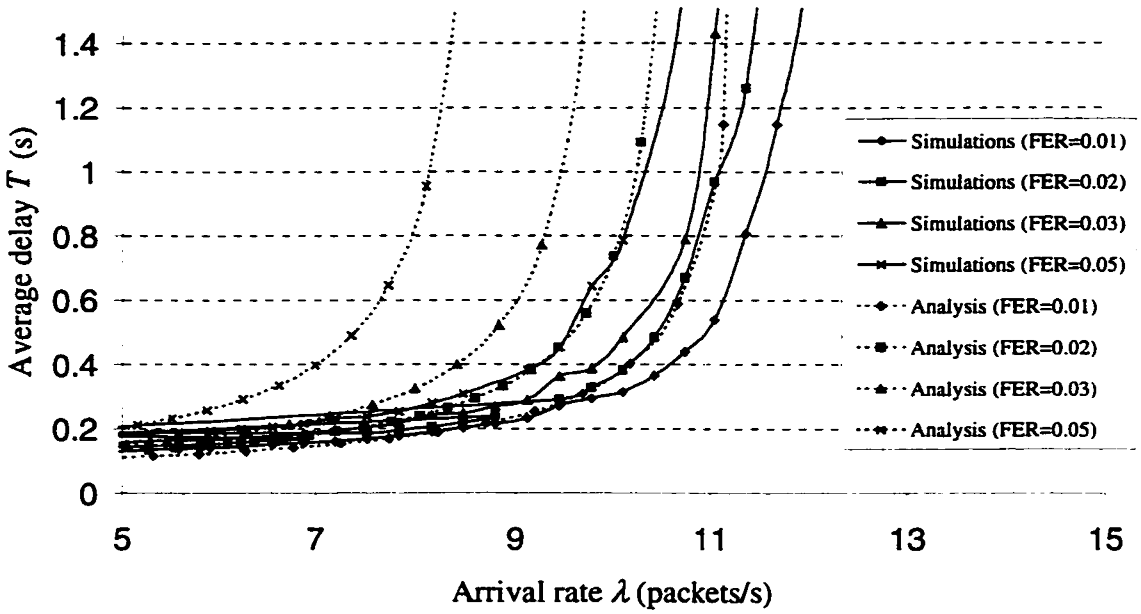

Figure 34. Average delay versus arrival rate as a function of FER for the new MAC protocols

Simulations for the e-mail packet size distribution of Figure 29 were also considered. The results are shown in Figure 35. For a maximum delay of 5s, the new MAC protocol improves the performance by 12 percent for a FER of 0.05, and about 10 percent for a FER of 0.10. Therefore for packet size distributions that have a mean packet size of 2kbytes, the new MAC protocol leads to a slight improvement and this improvement decreases as the FER increases.

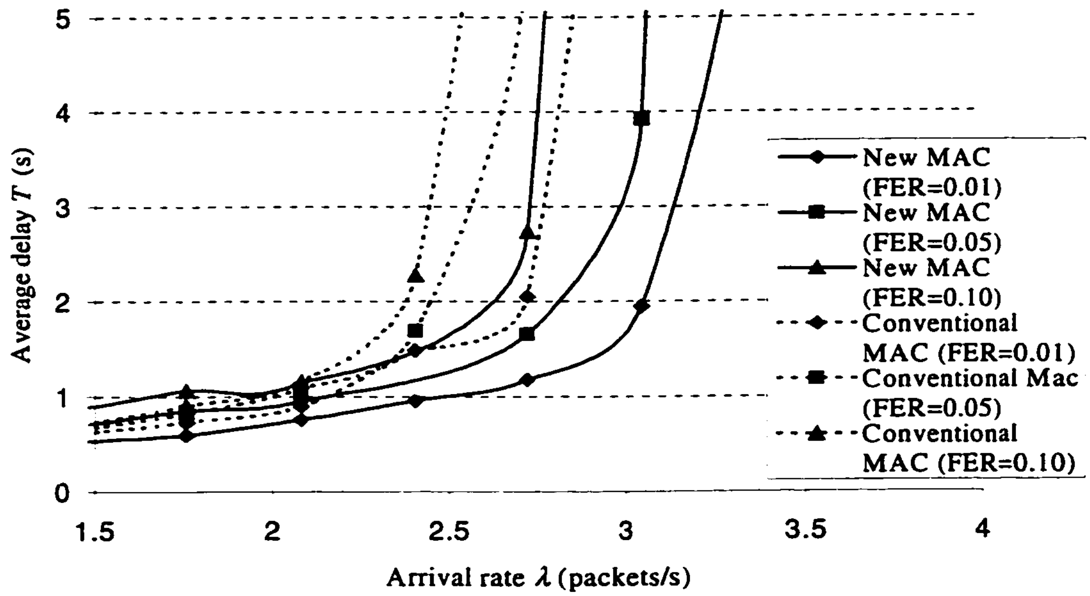

Figure 35. Mean packet delay for the e-mail distribution of Figure 29 as a function of FER.

As for the Internet distribution of Figure 31, the results of the average delay versus the arrival rate as a function of the FER are given in Figure 36. The new MAC protocol performs well even under high FER. In fact, for an FER equal to 0.05, the improvement in the maximum arrival rate is about 50 percent with respect to the conventional MAC. For an FER of 0.10, the new MAC protocol leads to a increase in the maximum arrival rate of about 35 percent. The performance of the new MAC protocol for the Internet packet size distribution is very good despite high FER (e.g. 0.10). This is due to the small average size of the packets.

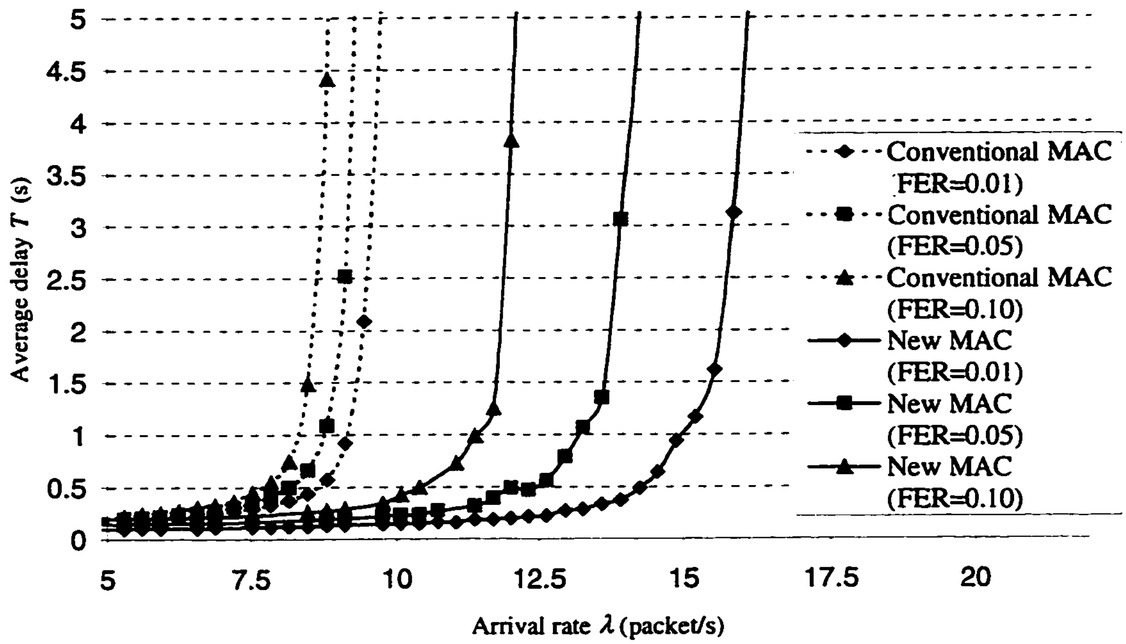

Figure 36. Average delay versus arrival rate as a function of FER for the typical Internet packet size distribution of Figure 31.

5.3. Throughput efficiency

The throughput efficiencies of the conventional and new schemes are compared using both analysis and simulations. Again, the analysis and simulation results are very close as demonstrated by Figures 37 to 40. From these figures, it can be seen that the throughput efficiency of the new MAC scheme is smaller but very close to that of the conventional one. Indeed, the throughput degradation caused by restarting the power control procedure (needed for a trailing segment retransmission) is small. The reason for this small degradation is that the concatenation of RS and convolutional codes makes the probability of frame retransmission low (lower than 0.01 according to the analysis done of Section 4.1), so that trailing segment retransmissions do not appreciably affect throughput. Another point is that throughput degradation is more severe for smaller packets. The new MAC protocol performs about 6 percent worse than the conventional one at an FER equal to 0.1 for 0.5kbyte packets, while it is about 3 percent worse for 1kbyte packets at the same FER, 1.5 percent worse to 2kbyte packets, and 1 percent worse for 4 kbyte packets.

This degradation decreases as the FERs is decreased. For an FER equal to 0.01 the throughput degradation is less than 1 percent for all mentioned packet sizes.

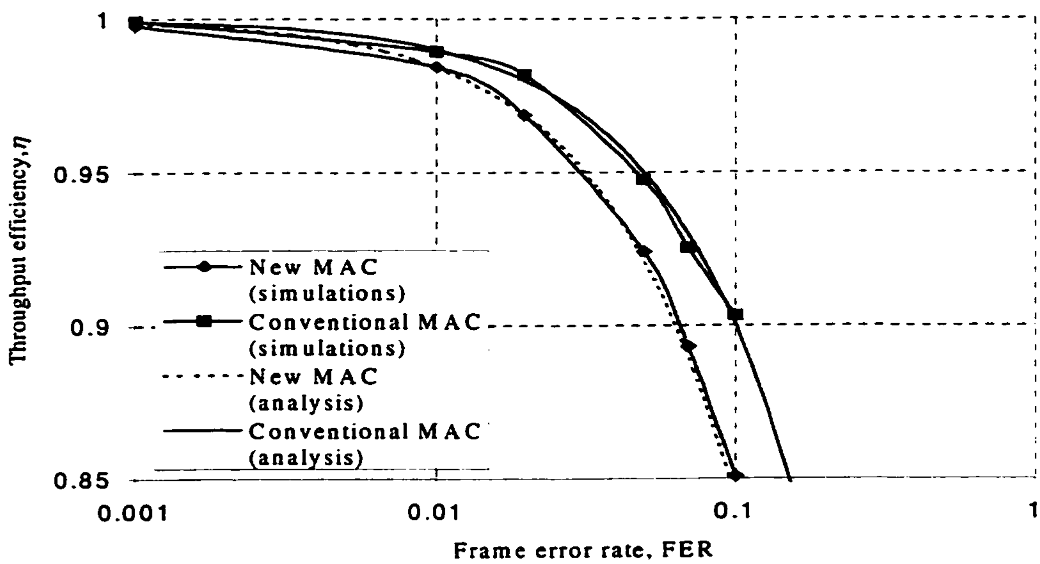

Figure 37. Throughput efficiency for 0.5kbyte packets.

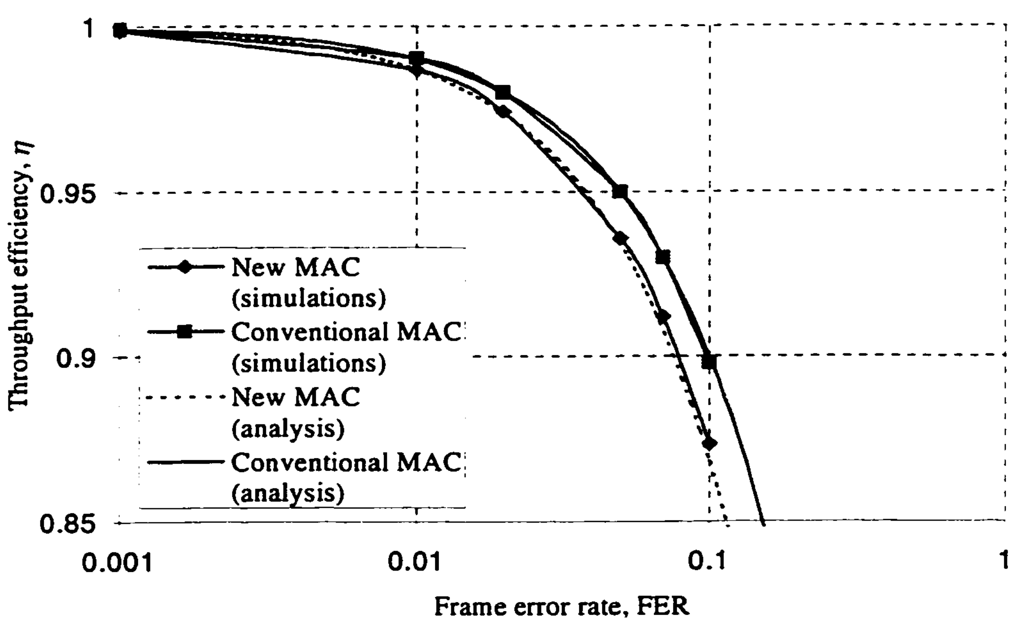

Figure 38. Throughput efficiency for 1kbyte packets.

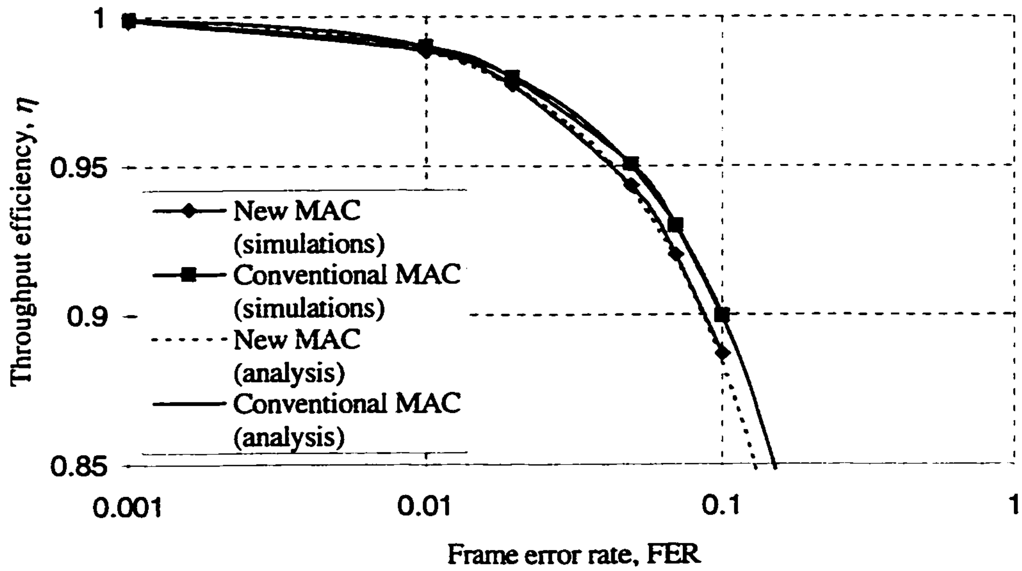

Figure 39. Throughput efficiency for 2kbyte packets.

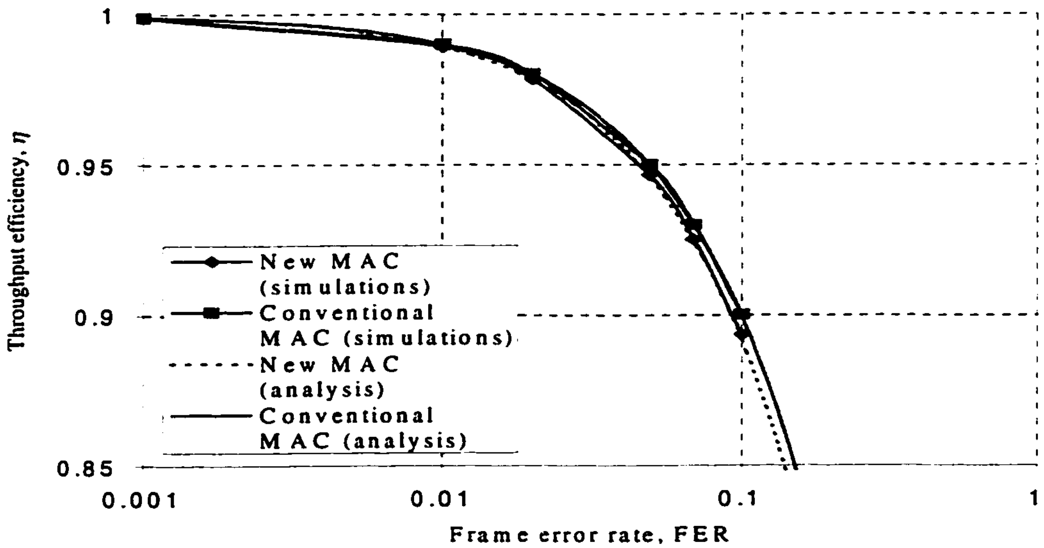

Figure 40. Throughput efficiency for 4kbyte packets.

5.4. Data activity factor

Both analysis and simulations have been performed for the data activity factor. Figure 41 shows that analysis and simulations give very close results for packets of size

1kbyte. For the new MAC protocol, the analysis gives results which are off by less that 5 percent compared to the simulations. This agreement is also true for packets of size 0.5, 2, and 4 bkytes (the corresponding figures are not shown). Note that the data activity factor of the conventional scheme reaches a maximum roughly equal to the ratio of packet transmission delay to the packet transmission plus processing delay, while that of the new scheme closely approaches unity. This demonstrates that the pair PDTCH/ARCH can be active all the time, with the new MAC scheme, if the arrival rate is high enough.

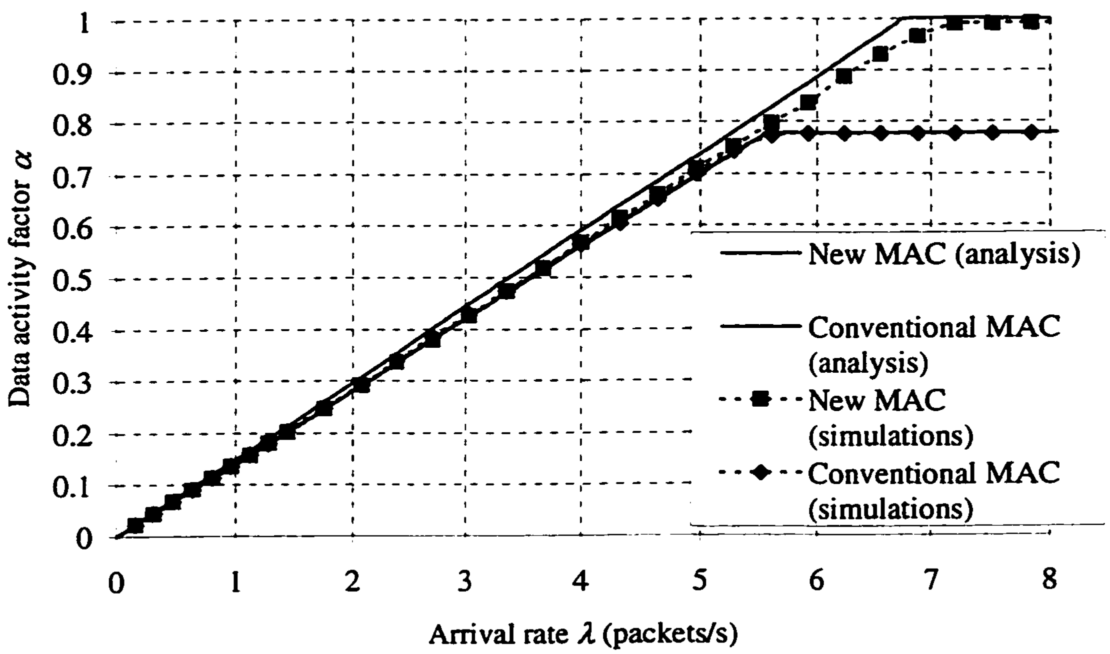

Figure 41. Data activity factor versus arrival rate for 1kbyte packets

Data activity factors for the two protocols were also obtained for packet sizes varying from 0.5 to 4kbytes. Figure 42 and Figure 43 show data activity factors for 0.5, 1, 2, and 4kbyte packets obtained through simulations. For all the cases, the new MAC scheme allows the data activity factor to reach a value close to unity, while the conventional scheme does not allow the data activity factor to reach unity because of the processing delay at the receiver. A near-optimal utilization of the PDTCH is achieved with the new MAC. However, in interference-limited systems, such as DS-CDMA systems, the less interference a channel causes, the higher the overall capacity, as mentioned in Section 2.1.1. Therefore there is a trade-off between arrival rate the PDTCH can handle and the capacity of DS-CDMA systems. Hence, the capacity of the system should be derived taking into account trade-offs between different parameters (maximum tolerable mean packet delay, MS arrival rate, and number of voice channels). This trade-off is discussed in Chapter 6.

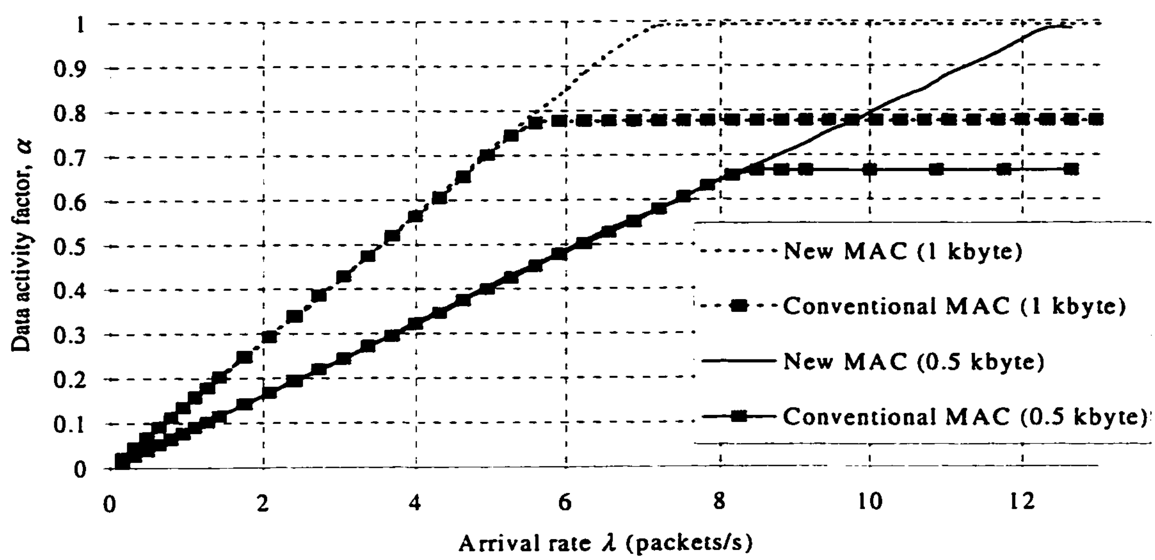

Figure 42. Data activity factor versus arrival rate for 0.5kbyte and 1kbyte packets.

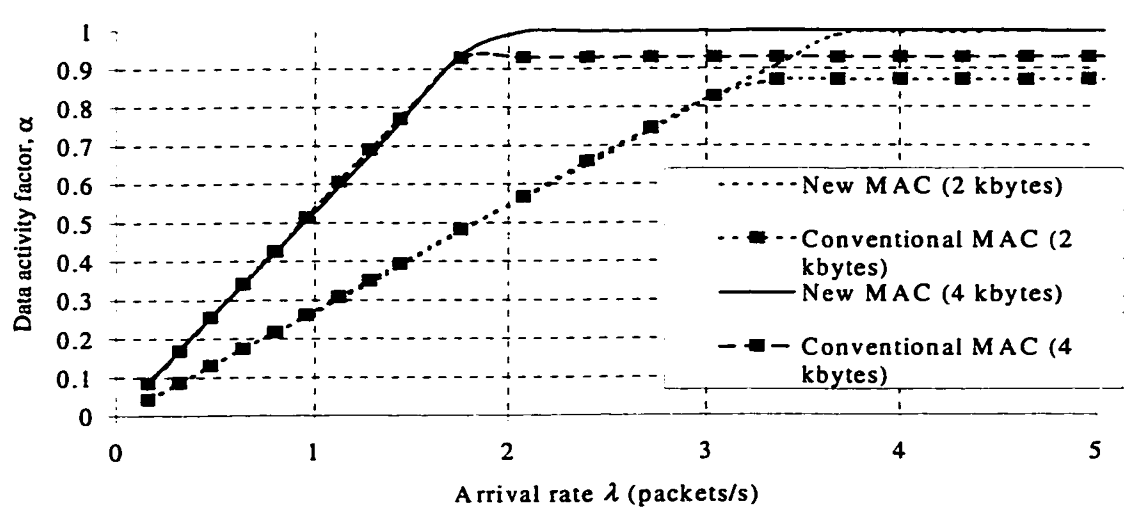

Figure 43. Data activity factor versus arrival rate for 2kbyte and 4byte packets.

Figure 44 illustrates the robustness of the new MAC protocol as a function of the FER. The data activity factor variation for the conventional scheme is negligible as the FER increases. For the new MAC scheme, as the FER increases the data activity increases for fixed λ. In fact, the overhead caused by the power control procedure and the PN code re-synchronization becomes non-negligible and contributes to increasing the activity of the PDTCH as the FER increases. This degradation in data activity factor, α, is less than 8 percent for an FER equal to 0.05 for the entire range of simulated λ compared to the case where FER is 0.01. This degradation is about 13 percent for the case where FER is equal to 0.10.

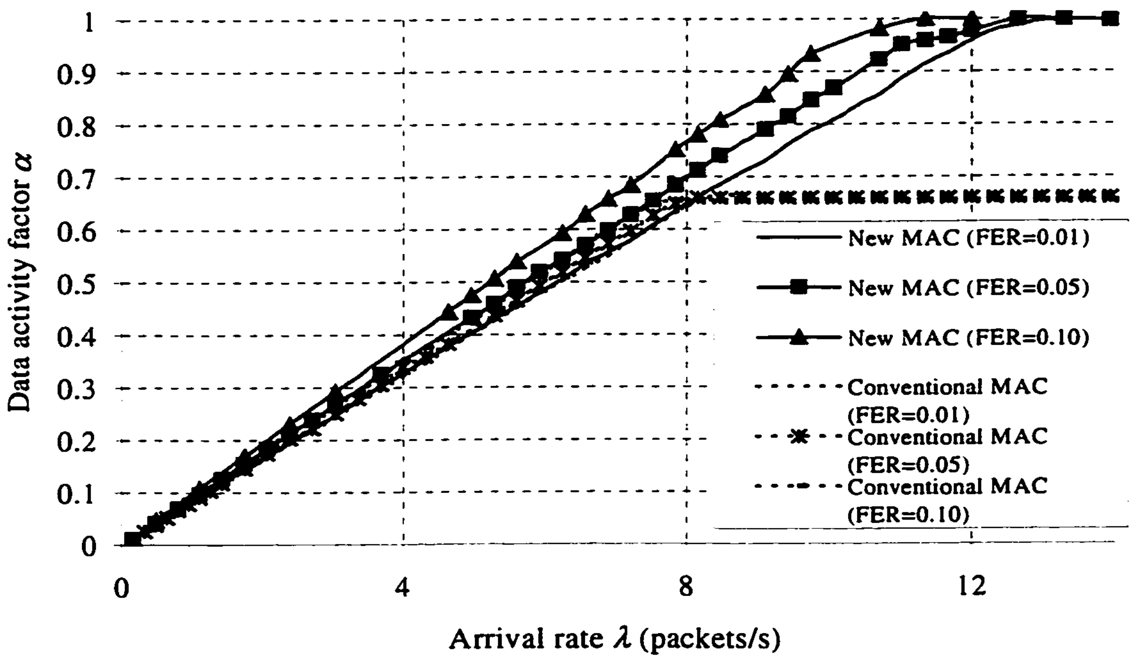

Figure 44. Data activity factor versus arrival rate as a function of FER (0.5kbyte packets).

The data activity factors of the e-mail packet size distribution of Figure 29 and the typical Internet packet size distribution from Figure 31 were also obtained. Figure 45 shows the simulation results for the e-mail distribution described in Section 5.2.2.1. It can be seen that the new MAC protocol data activity factor, α_N, reaches unity while the conventional scheme data activity factor, α_c, is limited to about 0.87. We note that this limit is roughly the same as for the case for which packets have a mean size of 2kbytes.

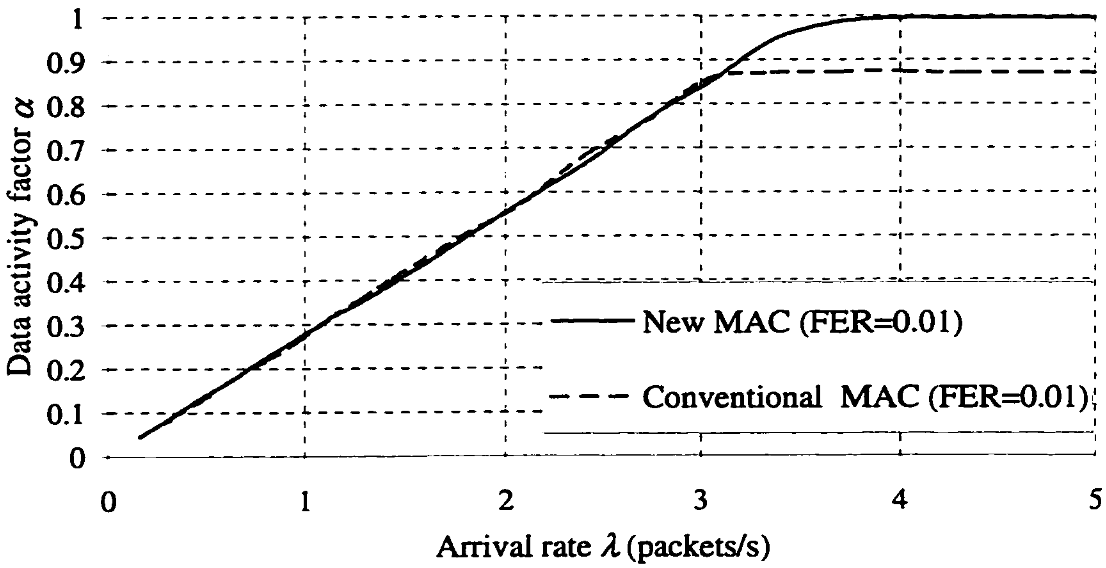

Figure 45. Data activity factor versus arrival rate for the e-mail packet size distribution of Figure 29.

As far as the typical Internet packet size distribution of Section 5.2.2.2 is concerned, the simulation results for the data activity factor are given in Figure 46. The behaviour of the activity factor is similar to the previous cases. The new MAC protocol allows the PDTCH/ARCH channel pair to operate at a near-optimal utilization. This channel pair, in the case of the conventional scheme, is limited by the processing delay and the data activity factor is limited to a value less than 0.6.

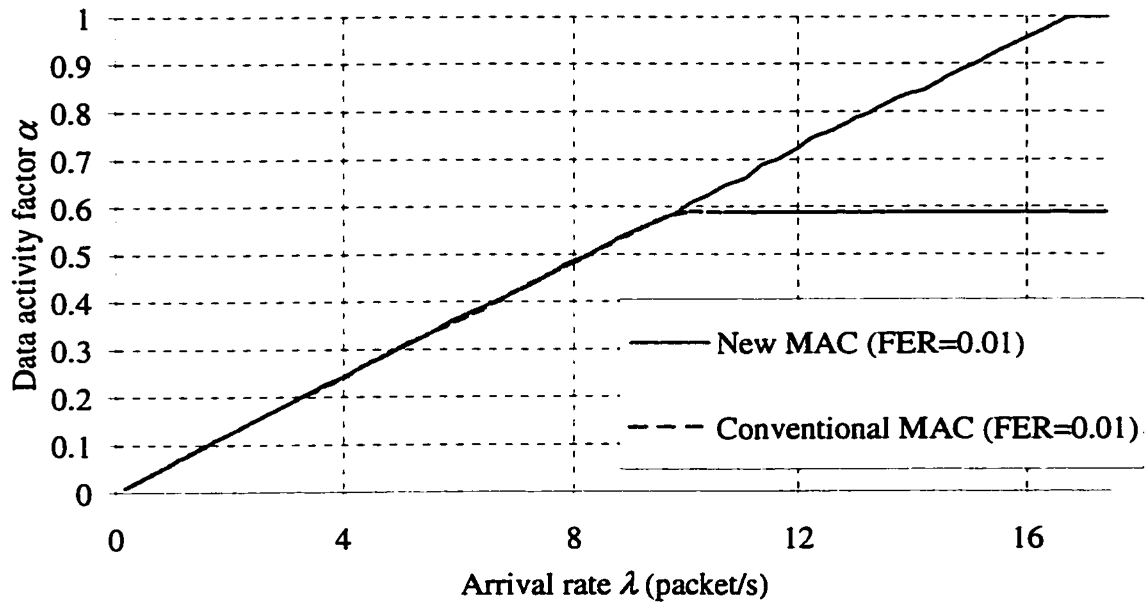

Figure 46. Data activity factor versus arrival rate for the typical Internet distribution of Figure 31.

5.5. Fairness of the algorithm

In order to check the capability of the MAC scheme to serve a pool of MSs attached to the PDTCH/ARCH in a fair manner, the long-term probability that an access request of each of the MSs is served has been determined by simulations and analysis. This is a measure of the fairness of the algorithm. Figure 47 shows the probability of an access request being served versus the MS number. Recall that for the simulations, a pool of 16 MSs was considered.

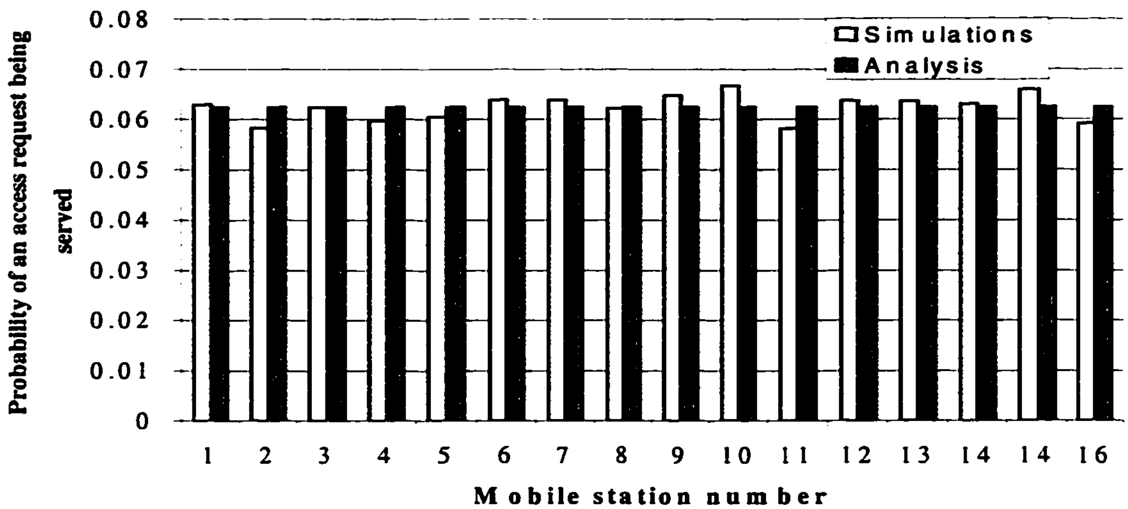

Figure 47. Fairness of the new MAC protocol.

A fair algorithm gives the same probability of an access request being served for all the MSs registered to the packet data service, that is, access is equiprobable for all the MSs (as shown with the grey bars on Figure 47). The new scheme achieves a fair treatment of all the MSs as shown by the white bars on that figure.

5.6. Conclusion

This chapter showed simulation results and numerical results of the analysis discussed in Chapter 4 for the new and conventional MAC schemes. First the performance of the synchronizer was shown for a Rayleigh fading channel. The results obtained for the synchronizer were then used to obtain the delay performance of the new MAC scheme under different conditions and scenarios. At 64kbps and for small packets (0.5 and 1kbyte), the new MAC protocol shows a significant increase in the arrival rate that the PDTCH can handle. For a typical e-mail packet size distribution, the new MAC protocol improves the maximum arrival rate by a fair amount. For a typical Internet packet size distribution, this improvement is quite significant. Under a high frame error rate, the new MAC scheme still outperforms the conventional one, but as the frame error rate increases, this improvement decreases. The throughput degradation due to the new MAC scheme compared to a conventional one is marginal. Finally, the data activity factor that the new MAC scheme can reach is close to unity, which means that the new

MAC scheme can achieve a near-optimal utilization of the PDTCH. The data activity factor of the conventional MAC scheme is limited by the processing delay of the receiver. The limit is all the more restricting as the packet size decreases. Finally, simulations show that the new MAC scheme allows a fair treatment of MS access requests.

Chapter 6. System capacity

In this section, a capacity analysis will be presented for a DS-CDMA cellular system using the new MAC protocol. Capacity analyses of the conventional MAC protocol and the CODIT approach discussed in Section 2.4.1 will also be presented for the purposes of comparison. In order to compare these MAC protocols, a capacity analysis involving two classes of users will be presented as in [35]. These two classes are voice users and packet-switched data users (or circuit-switched data user in the case of the CODIT approach).

6.1. Capacity analysis

6.1.1. CODIT MAC protocol for data transmission

In this section a capacity analysis for the first MAC protocol of Section 2.4.1 will be presented. This MAC protocol provides a data service in a very straightforward manner. The DPCCH is continuously transmitted to allow synchronization and power control to be continuously maintained. The DPDCH can therefore be discontinuous. Note that this method of providing data services is a circuit-switched approach as stated in Section 2.4.1.1. The analysis is performed for comparison purposes.

Let $I_{0,d}$ be the power spectral density of the interference seen by the DPDCH of a data user, N_d-the number of data users in the system, α_d the data activity factor of the DPDCH, $E_{c,d}$ the energy per chip of the DPDCH, $E'_{c,d}$ the energy per chip of the DPCCH, and I_v the power spectral density of the interference created by voice users. Then, using the same argument as in [35] and neglecting the effect of thermal noise, $I_{0,d}$ is given by

$$I_{0,d} = \alpha_d(N_d - 1)E_{c,d} + E'_{c,d}N_d + I_v \tag{63}$$

Note that the second term of (63) does not contain the data activity factor since the DPCCH is continuously transmitting in the CODIT MAC protocol. Moreover, there are N_d interfering sources because of the $N_d - 1$ DPCCHs of the other users and also the user's own DPCCH that is also treated like an additional interfering source. Taking into account the characteristics of the pulse-shaping filter and the self-interference created by multipath propagation given by $\alpha_d \delta E_{c,d}$ [24], (63) becomes

$$I_{0,d} = \alpha_d \delta E_{c,d} + \alpha_d \beta (N_d - 1) E_{c,d} + E'_{c,d} \beta N_d + I_v \tag{64}$$

Where β is given by the shape of the pulse-shaping filter used. For a square root raised cosine filter with ρ as the roll-off factor, $\beta = 1 - \rho / 4$. Also, $\delta = 1 - 1/L'$ for equal-strength paths and normalized energy equal to one. L' is the number of resolvable paths of the mobile channel and L' will be taken equal to four as in [24].

Let N_v be the number of voice users, $E_{c,v}$ the chip energy of the DPDCH from the voice users, $E'_{c,v}$ the chip energy of the PDCCH from the voice users, and α_v the voice activity factor. Then the interference caused by voice users is given by

$$I_v = \alpha_v \beta E_{c,v} N_v + \beta E'_{c,v} N_v \tag{65}$$

Note that in (65), as in (64), the second term does not depend on α_v because the DPCCH is continuously transmitted. This expression can be simplified by introducing the ratio of the DPCCH chip energy to the DPDCH chip energy, $\rho_v = \frac{E'_{c,v}}{E_{c,v}}$. (65) then becomes

$$I_v = \alpha_v \beta N_v (1 + \tfrac{\rho_v}{\alpha_v}) E_{c,v} \tag{66}$$

Now if sectorization, power control inaccuracy, and interference coming from other cells are taken into account in (66) and (64), the expression for $I_{0,d}$ becomes

$$I_{0,d} = \frac{1}{n\mu} \left\{ \alpha_d \delta E_{b,d} + \alpha_d \beta [(1+\theta) N_d - 1] E_{b,d} + \beta (1+\theta) N_d E'_{c,d} + \alpha_v \beta (1+\theta) N_v (1 + \tfrac{\rho_v}{\alpha_v}) E_{c,v} \right\} \tag{67}$$

where n is the number of sectors, which refers to the subdividing of a cell into n parts where all the parts are isolated. $n = 3$ will be taken for the comparison. μ is the degradation factor due to imperfect power control, which is equal to one if perfect power control is achieved. $\mu = 0.85$ will be taken just as in [35]. Finally θ is a factor that takes into account interference coming from users in other cells, $\theta = 0.5$ is usually taken.

It is then possible to compute the ratio of the energy per bit to the interference power spectral density for the packet data users of the CODIT MAC protocol, $\left(E_b/I_0\right)_d$. This ratio is defined as

$$\left(E_b/I_0\right)_d = \frac{E_{b,d}}{I_{0,d}} = \frac{G_d E_{c,d}}{I_{0,d}} \tag{68}$$

With G_d, the processing gain is defined as the ratio of the chip rate to the data user DPDCH bit rate. The bit rate of the data is supposed to be 64kbps and the chip rate is supposed to be 4.096Mchips/s. Using (67) in (68) and making algebraic simplifications leads to

$$\left(\frac{E_b}{I_0}\right)_d = \frac{n\mu G_d}{\alpha_v\delta + \alpha_d\beta[(1+\theta)N_d - 1] + \beta(1+\theta)N_d\rho_d + \alpha_v\beta(1+\theta)N_v(1+\frac{\rho_v}{\alpha_v})\frac{E_{c,v}}{E_{c,d}}} \tag{69}$$

where $\rho_d = \frac{E'_{c,d}}{E_{c,d}}$.

In order to satisfy the quality of service of the data transmission, $\left(E_b/I_0\right)_d$ must be greater than a certain target signal to interference ratio, γ_d. This γ_d is chosen in order to have a BER smaller than 10^{-9}. According to Section 4.1, γ_d must be equal to 3.25dB. Solving for N_d, the inequality $\left(E_b/I_o\right)_d \geq \gamma_d$ gives

$$N_d \leq \tilde{N}_d - aN_v \tag{70}$$

where:

$$\tilde{N}_d = \frac{1}{\beta(1+\theta)(1+\frac{\rho_d}{\alpha_d})}\left[\frac{n\mu G_d}{\alpha_d}\frac{1}{\gamma_d} + \beta - \delta\right] \tag{71}$$

$$a = \frac{\alpha_v}{\alpha_d}\frac{(1+\frac{\rho_v}{\alpha_v})}{(1+\frac{\rho_d}{\alpha_d})}\frac{E_{c,v}}{E_{c,d}} \tag{72}$$

Now the same steps can be repeated in order to determine an upper bound on the number of voice users, N_v. Using the same notations as before, the power spectral density of the interference a voice user sees from other data users and other voice users, $I_{o,v}$ is given by

$$I_{0,v} = \frac{1}{n\mu}\{\alpha_v \delta E_{c,v} + \alpha_v \beta[(1+\theta)N_v - 1]E_{c,v} + \beta(1+\theta)N_v E_{c,v}^{'} + \alpha_d \beta(1+\theta)N_d E_{c,d} + \beta(1+\theta)N_d E_{c,d}^{'}\} \quad (73)$$

The first term of (73) represents the self-interference term due to multipath propagation, the second term represents interference created by other voice user DPDCHs, the third term represents interference created by all the voice DPCCHs, the fourth term represents interference created by all the data users DPDCHs, and the last term represents interference from all the data users DPCCHs.

Using (73) it is possible to determine the ratio of the energy per bit to the interference power spectral density a voice user sees in CODIT, $\left(E_b/I_0\right)_v$. Once this $\left(E_b/I_0\right)_v$ has been found, the same argument as for data users can be applied again. The quality of voice communications requires that BER is smaller than 10^{-3}. This quality requirement is satisfied only if $\left(E_b/I_0\right)_v$ is greater than a certain threshold γ_v. According to [11], a value for γ_v that gives this quality of service is 6.3dB. Solving the inequality $\left(E_b/I_0\right)_v \geq \gamma_v$ for N_v leads to

$$N_v \leq \tilde{N}_v - \tfrac{1}{a} N_d \quad (74)$$

where:

$$\tilde{N}_v = \frac{1}{\beta(1+\theta)(1+\frac{\beta_v}{\alpha_v})}\left[\frac{n\mu G_v}{\alpha_v}\frac{1}{\gamma_v} + \beta - \delta\right] \quad (75)$$

Here G_v is the processing gain of the voice channel defined as the ratio of the chip rate over the nominal voice coder rate. The voice coder rate is assumed to be 8kbps. The parameter a is given by (72).

Following the same argument as in [35], (74) and (70) must give the same result for N_d when N_v is equal to zero. This leads to

$$\tilde{N}_d = a\tilde{N}_v \quad (76)$$

Combining (74), (70), and (76) gives

$$\frac{N_d}{\tilde{N}_d}+\frac{N_v}{\tilde{N}_v}\leq 1 \tag{77}$$

(77) clearly shows the trade-off between the number of data users and the number of voice users. (77) can be rewritten as

$$N_v \leq \tilde{N}_v - \frac{\tilde{N}_v}{\tilde{N}_d} N_d \tag{78}$$

In (78), the term $\frac{\tilde{N}_v}{\tilde{N}_d} N_d$ represents the cost in terms of the number of voice users having N_d data users. This term will be called, C_{N_d}

$$C_{N_d} = \frac{\tilde{N}_v}{\tilde{N}_d} N_d \tag{79}$$

This is the parameter that is going to be used to compare different MAC protocols. (78) also shows that $\tilde{N}_v$ is the maximum number of voice users that the cellular system can handle. When $N_d = 0$, the number of voice users, N_v, is equal to the maximum number of voice users, $\tilde{N}_v$.

6.1.2. Conventional and new MAC protocols

In this section, the system capacity analysis of the conventional (described in Section 2.4.1.3) and the new MAC scheme is presented. The analysis for the two MAC protocols is similar, but the difference between them is the data activity factor shown in Section 5.4.

The approach used for the analysis in this section is the same as the one in the previous section. The interference seen by a given user (data or voice) will be derived first, and then the signal-to-interference ratio will be computed. The quality of service required leads to an inequality that is used to determine the maximum number of users (data or voice).

The maximum number of data users will be computed first. In the case of both the conventional and new MAC protocols, the interference power spectral density seen by a data user, denoted the same way as in Section 6.1.1, $I_{0,d}$ is given by

$$I_{0,d} = \frac{1}{n\mu}\left\{\alpha_d \delta E_{c,d} + \alpha_d \beta[(1+\theta)N_d - 1]E_{c,d} + \alpha_d \beta(1+\theta)N_d E_{c,d} + \alpha_d \beta \tfrac{n_s}{n_p} E_{c,ARM} + \alpha_v \beta(1+\theta)(1+\tfrac{\beta_v}{\alpha_v})N_v E_{c,v}\right\} \quad (80)$$

In (80), the first three terms are the self-interference, the interference from other data user DPDCHs, and the interference from all the data user DPCCHs. Note that the interference coming from all the DPCCHs depends on the data activity factor, α_d, of these users. In this present case, the DPCCH of a data user is torn down when it is not transmitting in order to reduce interference as discussed in Section 2.4.1. The fourth term of (80) represents the interference coming from the ARCH. This term depends on the average number of ARMs needed to acquire synchronization, n_s (defined in Section 3.1.1.2), the average number of frames per packet when no retransmission is needed, n_p (defined in Section 4.2.4.2), the data activity factor, α_d, and the energy per chip used for the ARMs, $E_{c,ARM}$. The fifth term is the interference caused by voice users (both DPDCHs and DPCCHs included).

The following relation relates the energy per chip used by the ARMs, $E_{c,ARM}$, to the energy per PN code period, $E_{c,period}$

$$E_{c,ARM} = \frac{E_{c,period}}{N_{code}} \quad (81)$$

where N_{code} is the period of the short PN code used to spread the ARMs. In [20], N_{code} is set equal to 128. The ratio of $E_{c,ARM}$ to $E_{c,d}$ is k (assumed to be constant). Some straightforward manipulations give

$$E_{c,period} = k\frac{N_{code}}{G_d}E_{b,d} \quad (82)$$

If both sides of (82) are divided by the power spectral density of the interference, I_0, then

$$\frac{E_{c,period}}{I_0} = \bar{\mu} = k\frac{N_{code}}{G_d}\left(\frac{E_{b,d}}{I_0}\right) \quad (83)$$

Figure 25 shows that $\bar{\mu}$ must be at least 12.4dB to allow the synchronizer to detect the ARMs with a probability of 0.99. In order to allow communications with a BER smaller than 10^{-9}, $E_{b,d}/I_0$ must be at least 3.25dB. Therefore solving for k in (82) gives k equal to 16.5.

Using (80), it is now possible to determine the energy per bit to interference ratio for data users, $(E_b/I_0)_d$. Once this ratio is determined, solving the following inequality $(E_b/I_0)_d \geq \gamma_d$ for N_d gives

$$N_d \leq \tilde{N}_d' - a' N_v \tag{84}$$

where:

$$\tilde{N}_d' = \frac{1}{\beta(1+\theta)(1+\rho_d)}\left[\frac{n\mu G_d}{\alpha_d}\frac{1}{\gamma_d} + \beta(1+\tfrac{n_s}{n_p}k) - \delta\right] \tag{85}$$

$$a' = \frac{\alpha_v}{\alpha_d}\frac{1+\frac{\rho_v}{\alpha_v}}{1+\rho_d}\frac{E_{c,v}}{E_{c,d}} \tag{86}$$

The differences between (71) and (85) are: the term $1+\frac{\rho_d}{\alpha_d}$ is now $1+\rho_d$ because the data user DPCCHs are not continuous; the term β is now $\beta(1+\frac{n_s}{n_p}k)$ because of the interference introduced by ARMs used for each transmitted packet. The difference between (72) and (86) is that the term $1+\frac{\rho_d}{\alpha_d}$ is now $1+\rho_d$ because of the discontinuous data user DPCCHs. Note that (70) and (84) are very similar.

The same steps can then be followed to derive the same type of bound for voice users. The first quantity to determine is the interference power spectral density that a given voice user sees. This is given by

$$I_{0,v} = \frac{1}{n\mu}\left\{\alpha_v \delta E_{c,v} + \alpha_v \beta[(1+\theta)N_v - 1]E_{b,v} + \beta(1+\theta)N_v \dot{E}_{c,v} + \alpha_d \beta(1+\theta)N_d(1+\rho_d)E_{c,d} + \alpha_d \beta \tfrac{n_s}{n_p}(1+\theta)N_d E_{c,ARM}\right\} \tag{87}$$

Using (87), it is possible to compute the ratio of the energy per bit to interference for a voice user, $(E_b/I_0)_v$. This $(E_b/I_0)_v$ must be greater than the threshold γ_v to keep the BER smaller than 10^{-3} for the voice service. Solving for this inequality leads to a bound on the number of users the voice service

$$N_v \le \tilde{N}_v^{'} - dN_d \tag{88}$$

where:

$$\tilde{N}_v^{'} = \frac{1}{\beta(1+\theta)(1+\frac{\rho_v}{\alpha_v})}\left[\frac{n\mu G_v}{\alpha_v}\frac{1}{\gamma_v} + \beta - \delta\right] \tag{89}$$

$$d = \frac{\alpha_d}{\alpha_v}\frac{1+\rho_d+\frac{n_t}{n_p}k}{1+\frac{\rho_v}{\alpha_v}}\frac{E_{c.d}}{E_{c.v}} \tag{90}$$

Note that (89) and (71) are exactly the same equations. When the number of PDTCHs is zero, $N_d = 0$, there are only voice users left in the system. In that special case, the only interference a voice user sees is interference coming from other voice users. Therefore, when $N_d = 0$, it does not matter which MAC protocol (the original CODIT approach or the new MAC protocol) is used to provide data transmission; hence $\tilde{N}_v^{'}$ is equal to $\tilde{N}_v$. The difference between the inverse of (72) and (90) is that $1+\frac{\rho_d}{\alpha_d}$ has become $1+\rho_d+\frac{n_t}{n_p}k$ because of the discontinuous transmission of the data user DPCCHs and the ARMs used for each packet.

Once again, by using (84) and (88) and using the same argument as in [35], the number of data users, N_d, and the number of voice users, N_v, can be shown to be related by

$$\frac{N_d}{\tilde{N}_d^{'}} + \frac{N_v}{\tilde{N}_v^{'}} \le 1 \tag{91}$$

It is more meaningful to rewrite this equation to express a bound on the number of voice users (as was done in (78) with the CODIT approach)

$$N_v \le \tilde{N}_v^{'} - \frac{\tilde{N}_v^{'}}{\tilde{N}_d^{'}}N_d \tag{92}$$

This last inequality shows the cost for voice user channels, $\frac{\tilde{N}_v^{'}}{\tilde{N}_d^{'}}N_d$, if the system has N_d ARCH/PDTCHs. The cost for N_d ARCH/PDTCHs, $C_{N_d}^{'}$, is therefore

$$C'_{N_d} = \frac{\tilde{N}'_v}{\tilde{N}'_d} N_d \tag{93}$$

C'_{N_d} is basically the reduction of the number of voice users caused by the N_d ARCH/PDTCHs interferers. An important comment must be made here. The number, N_d, of data channels is not equal to the number of data users, as it is the case with CODIT approach. In fact, because of statistical multiplexing, in this present case, a PDTCH represents several MSs attached to it. The number of MSs that can be attached to a single PDTCH depends on the arrival rate of the packets at the MSs, as will be explained in the next section.

6.2. Framework of the numerical results

In this section, numerical results corresponding to the previous analysis will be given. Before comparison results are given, it is necessary to describe the framework used in these comparisons.

6.2.1. CODIT MAC protocol

This MAC protocol is a circuit-switched approach and, therefore, each MS is given a separate DPDCH/DPCCH channel pair. MSs are assumed to have the same arrival rate, λ_{MS}. The data activity factor of the DPDCH is obtained by using (60), which is $\alpha = \lambda_{MS} \cdot \bar{A}$, where $\bar{A}$ is the mean packet duration. This data activity factor is then used as input for the equations derived in 6.1.1.

6.2.2. Conventional and new MAC protocol

Again, it is assumed that all the MSs have the same arrival rate, λ_{MS}. According to one of the properties of Poisson processes (given in Section 2.5.4), the aggregate process is also Poisson. If λ is the arrival rate of the aggregate process, then it is equal to $m \cdot \lambda_{MS}$, where m is the actual number of MSs attached to the channel pair ARCH/PDTCH. The actual number of MSs attached to the PDTCH is always smaller or equal to N_{slot}, the number of slots provided in the ARCH (see Section 3.1.1.2).

For the new MAC scheme and the conventional scheme, the actual number of MSs that can be attached to a PDTCH is determined by the maximum arrival rate, $\lambda_{\max}$, a PDTCH can handle. The maximum arrival rate a PDTCH can handle is in turn determined by the maximum tolerable delay, $T_{\max}$, assumed for the data service. In the case of packets with constant packet size, $\lambda_{\max}$ is found using the data of Figure 27 and Figure 28 given $T_{\max}$. Analysis could also be used to determine $\lambda_{\max}$ for $T = T_{\max}$ in (45) with $\overline{X}$ and $\overline{X^2}$ given by (46) and (47) for the conventional MAC scheme, and (52) and (53) for the new MAC scheme. In order to have the maximum number of users registered to the ARCH/PDTCH (this number is also called the capacity of the channel pair ARCH/PDTCH), m must be maximized under the following constraints

$$m \leq N_{slot} \tag{94}$$

$$m\lambda_{MS} \leq \lambda_{\max} \tag{95}$$

(94) and (95) show that the capacity of the channel pair ARCH/PDTCH is either limited by the number of available slots on the ARCH or the maximum arrival rate that the pair can handle.

Once m is found, it is possible to obtain the data activity factor, α_d, using the analysis with (61) and (62) for the conventional and the new MAC protocol, respectively. This data activity factor can also be obtained using the simulation results shown in Section 5.4. α_d will then be used as an input to the cost function of (93) obtained in Section 6.1.2.

An overload of the PDTCH/ARCH can occur when the aggregate arrival rate is greater than the maximum arrival rate, $\lambda_{\max}$, that the PDTCH can handle. This happens when the number of MSs attached to the ARCH/PDTCH, m, is lower than the number of available ARCH time-slots, N_{slot}. In this situation, the average delay of packets exceeds the maximum tolerable value and, in this case, the quality of service is not satisfied anymore. Therefore, in order to maximize m without inducing an overload of the PDTCH/ARCH, a call admission control (CAC) scheme must be used. This CAC scheme is responsible for a normal operation of the ARCH/PDTCH. The CAC scheme allows MSs to register only if the additional arrival rate they add to the aggregate arrival rate does not lead to a saturation of the ARCH/PDTCH. The CAC scheme is beyond the scope of the thesis and will not be covered.

6.3. Capacity analysis numerical results

This section presents the numerical results of the analysis performed in Section 6.1. First the improvement from the CODIT circuit-switched to a conventional MAC protocol for packet-switched data will be shown. Then the following trade-off will be shown: the trade-off between the increase in the data user capacity and a decrease in the voice user capacity that the use of the new MAC protocol implies.

The following parameter values are used to obtain the numerical results:

- A roll-off factor, ρ, of 0.35 as in [24] for the square root raised cosine pulse shaping filter. Therefore, β is equal to 0.9125.
- A channel with four equal-strength paths and normalized channel energy equal to one as in [24]. Therefore, $\delta = 1 - \frac{1}{4}$.
- Three sectors per cell, $n = 3$, a degradation factor due to imperfect power control, $\mu = 0.85$, and intercell interference factor, $\theta = 0.5$, as used in [35].
- A maximum number of slots per ARCH, N_{slot}, equal to 16.
- An average number of ARMs needed for synchronization, n_s, equal to 1.01.
- An number of frames per packet, n_p, equal to 7, 13, 26 and 52 for packets of 0.5, 1, 2, 4 kbytes, respectively, at a data rate of 64kbps.
- A ratio of the energy per chip on the ARCH channel to the energy per chip of the DPDCH channel, k, equal to 16.5.
- A processing gain of the PDTCH, G_d, equal to 64 (chip rate of 4.096Mchips/s over 64kbps).
- A processing gain the voice channel, G_v, equal to 512 (8kbps voice coder).
- A voice activity factor, α_v, of 3/8.
- A ratio of DPCCH chip energy to DPDCH chip energy of the PDTCH, ρ_d, equal to 0.25 as given in [26].

- A ratio of DPCCH chip energy to DPDCH chip energy of voice channels, ρ_v, equal to 0.18 as given in [26].
- The maximum tolerable delay is assumed to be 1s for fixed size packets and the typical Internet packet size of Figure 31 while it is taken to be 5s for the e-mail packet size distribution of Figure 29.

6.3.1. Conventional MAC protocol increase of data user number over circuit-switched approach

For the comparison between the CODIT approach and a conventional MAC scheme for packet data transmission, the number of channels used in CODIT circuit-switched approach is set equal to m, the number of users that a single channel pair ARCH/PDTCH can handle. In this manner, these approaches can be compared on the basis of the same number of data users being served. In order to ease the comparison, the reduction in the number of CODIT voice users, C_{N_d}, given in (79), will be given relative to the voice reduction of number of voice users of the conventional MAC protocol, C'_{N_d}, given in (93).

Figure 48 shows the voice number decrease of CODIT circuit-switched approach relative to the conventional MAC scheme for packet-switched tra ısmission for packets of constant size. As shown by this figure, the decrease is quite significant for the entire range of MS arrival rates, λ_{MS}. It is all the more significant as the packet size decreases. These results justify the statements made in Section 2.4.1.1. That is, the interference created by the continuous data user DPCCHs is very detrimental to the voice users. The gain in voice user capacity obtained by using a packet-switched approach for data transmission is quite significant. This gain clearly justifies the use of such an approach in DS-CDMA even if DTX requires extra investment (fast synchronizer and more efficient power control algorithm). This figure also shows that as the MS arrival rate increases the decrease is less severe. The reason for this behaviour is that a higher arrival rate implies a more frequent access to the PDTCH. If a data source requires very frequent access to the shared channel, the network is better off giving it a fully dedicated traffic channel.

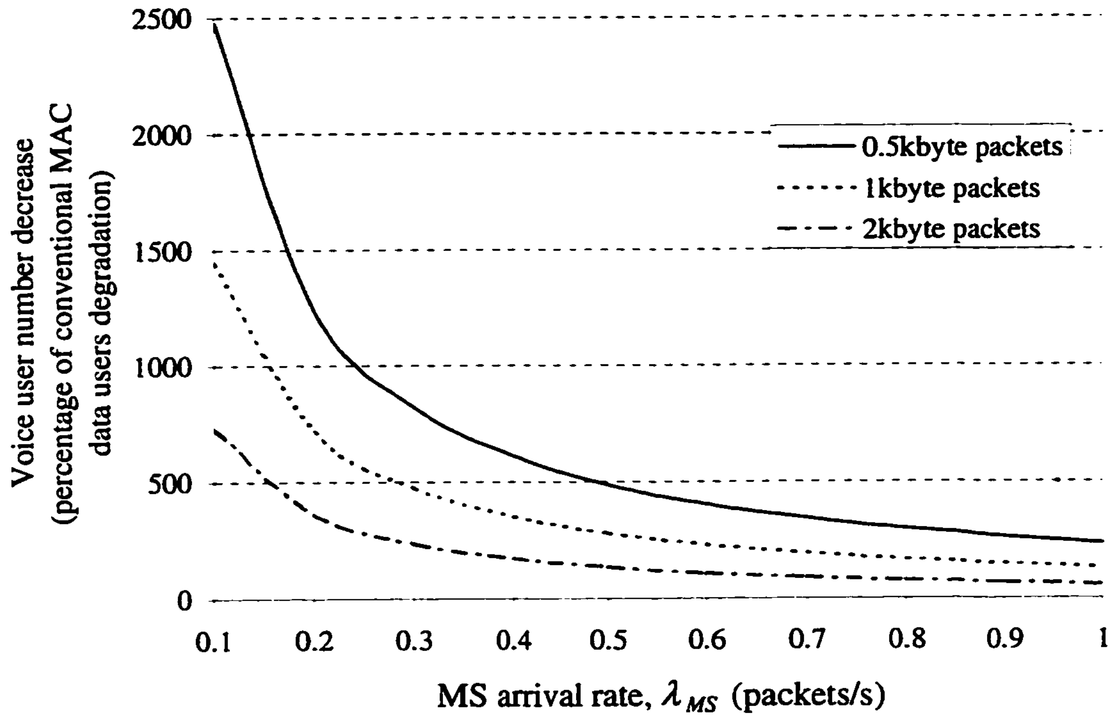

Figure 48. Voice user number decrease of the CODIT circuit-switched approach relative to the voice user number decrease of the conventional MAC protocol.

Figure 49 shows this voice user decrease when packets follow the typical e-mail packet size distribution of Figure 29. This figure shows that the decrease follows the same behaviour as in the case where packets have a constant size. This decrease combined with a high arrival rate is not smaller than 100 percent for λ_{MS} greater than 0.6. The reason for this lower decrease is that packets are longer and the packet-switched approach is becoming less and less attractive as the MS arrival rate increases compared to a circuit-switched approach.

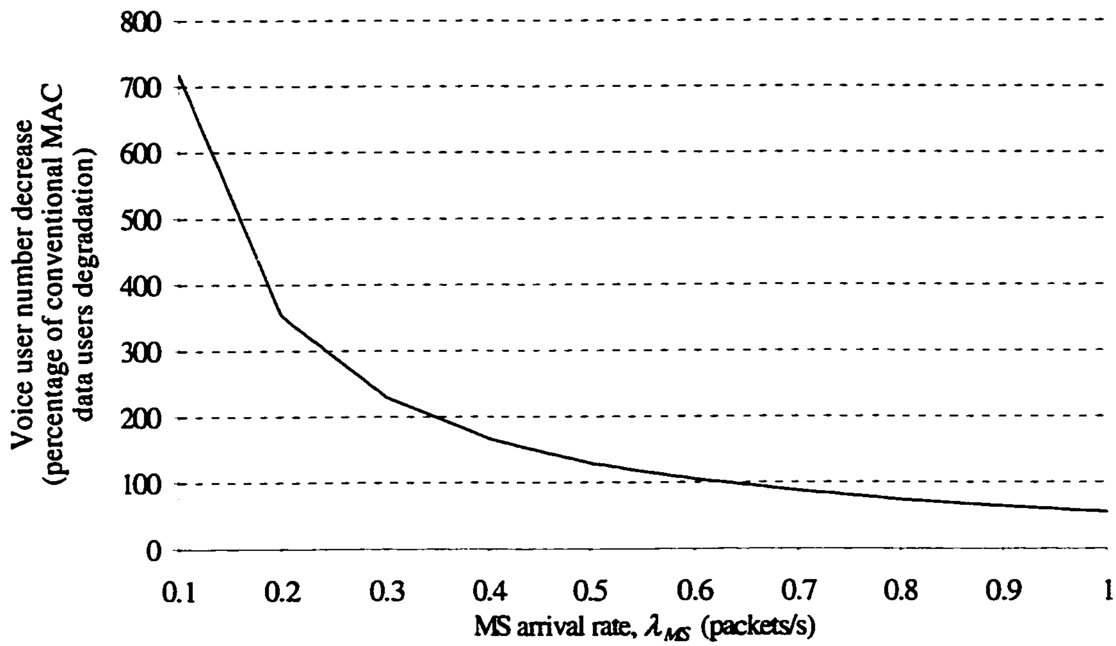

Figure 49. Voice user number decrease of the CODIT circuit-switched approach relative to the conventional MAC scheme with the typical e-mail packet size distribution of Figure 29.

6.3.2. New MAC and conventional MAC protocol trade-off

In this section, the trade-off mentioned in Section 5.4 is going to be discussed. Once again the benchmark for the comparison is going to be the conventional MAC protocol. The reason for this choice is that the voice number decrease caused by the new MAC protocol is higher due to a greater interference caused by the ARCH/PDTCH.

Figure 50 shows the voice user decrease of the new MAC scheme compared to a conventional one. As expected with the new MAC scheme, having a PDTCH with a data activity factor close to unity (see Section 5.4) leads to greater interference than with the conventional scheme. This greater interference leads to a reduction in the number of voice user as shown in Figure 50. For 0.5kbyte packets, there is no decrease for λ_{MS} smaller than 0.4 packets/s. For this range of λ_{MS} values, the number of users per ARCH/PDTCH is limited by the number of slots of the ARCH, N_{slot}. That is, the maximization of m is limited by (94) (see Section 6.2.2). In this case, the number of users of the PDTCH is the same for the conventional scheme and the new MAC scheme. Therefore, the decrease that they cause to voice users is the same, that is, the relative degradation equals zero. When λ_{MS} is greater than 0.4 packets/s, the maximization of m is limited by (95) because the activity of every single MS becomes such that the limiting factor becomes λ_{max}, the maximum arrival rate the pair ARCH/PDTCH can handle. For 1kbyte packets, there is no relative decrease for λ_{MS} smaller than 0.3 packets/s. The maximization of the integer m under the constraints specified in (94) and (95) leads to a relative decrease curve that goes up and down as shown in Figure 50. The behaviour of the curves will be explained in the following.

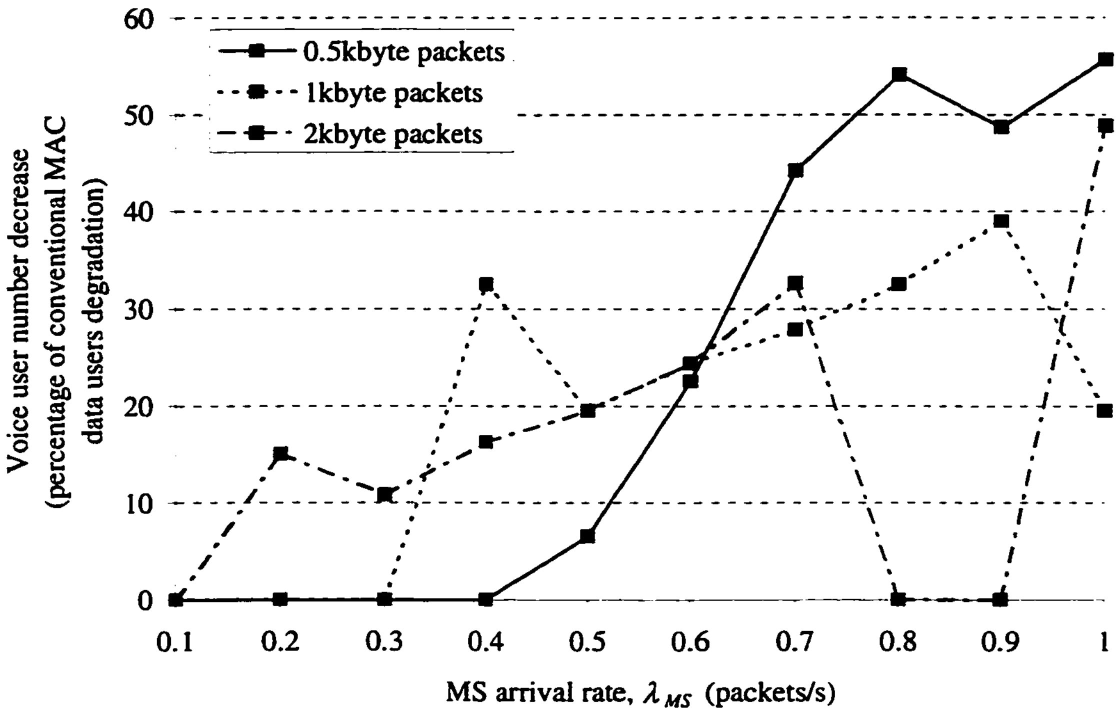

Figure 50. Voice user number decrease with the new MAC scheme relative to the conventional MAC scheme.

Now the improvement due to the new MAC scheme can be shown. Figure 51 shows the increase of the number of data users that the new MAC scheme can achieve over a conventional one. The first thing to notice is that Figure 50 and Figure 51 are very similar. This is actually not surprising and shows the trade-off between capacity and interference, typical to DS-CDMA systems. Interference is the main limiting factor in these systems. What the new MAC scheme can improve with respect to data users necessarily leads to higher interference. This higher amount of interference must bring about a decrease of the number of voice users. Figure 51 shows that the new MAC scheme can lead to an increase in the number of users attached to the pair ARCH/PDTCH compared to a conventional one. This increase in data users is almost exactly equivalent to the decrease in voice users shown previously. Therefore the same comments as in the previous paragraph can be made. The improvement of the new MAC scheme is seen when λ_{MS} is greater than a threshold value that depends on the packet size. This clearly shows the trade-off between increasing data users and decreasing voice users. The increase in the number of data users ranges from 0 to about 60 percent depending on the data rate of the data users. For 0.5kbyte packets, the increase starts at a value of λ_{MS} equal to 0.4 packets/s and reaches about 60 percent for λ_{MS} equal to 1 packets/s. For 1kbyte packets, the increase starts at 0.3 packets/s and can reach 40 percent. For 2kbyte packets, the increase starts at 0.1 packets/s and can reach 50 percent. Again the maximization of the integer *m* discussed in Section 6.2.2 leads to a degradation curve that goes up and down as shown in Figure 51.

In this paragraph, explanations are given for the somewhat complex behaviour of the curves of Figure 51 and in fact the behaviour of all figures in this section. For 2kbyte packets, a maximum tolerable delay chosen to be 1s (see Section 6.1) leads to a maximum arrival rate, $\lambda_{\max}$, of 3.16 and 2.75 packets/s for the new MAC protocol and the conventional MAC protocol, respectively (see Figure 28). If MS arrival rates, λ_{MS}, greater than 0.7 packets/s are considered then the maximum number of MSs attached to the ARCH/PDTCH, *m*, is limited by the maximum arrival rate the PDTCH can handle, $\lambda_{\max}$, and not by the number of available slots, N_{slot}, on the ARCH. In other words, the maximization of *m* is limited by (95) and not (94). Given λ_{MS} equal to 0.7 packets/s, dividing $\lambda_{\max}$ by λ_{MS} gives 4.51 and 3.92, for the new MAC scheme and the conventional one, respectively. Since *m* is an integer, these values must be rounded down in order to obtain *m*, which is then equal to 4 and 3, respectively. These values show that the new MAC scheme can increase the number of data users by about 33 percent for λ_{MS} equal to 0.7 packets/s (see Figure 51). Now, if λ_{MS} is equal to 0.8, dividing λ_{max} by λ_{MS} gives 3.95 and 3.43 which leads to values of m equal to 3 and 3 for the new MAC scheme and the conventional scheme, respectively. Therefore, the improvement is zero for λ_{MS} equal to 0.8 packet/s. This improvement is also zero foı λ_{MS} equal to 0.9 packets/s as shown by Figure 51. For λ_{MS} equal to 1 packet/s, λ_{max} divided by λ_{MS} leads to 3.16 and 2.75 for the new MAC scheme and the conventional one, respectively. In that case, m is equal to 3 and 2, respectively. This gives rises to a data user number increase of 50 percent for the new MAC scheme over the conventional one for λ_{MS} equal to 1 packet/s.

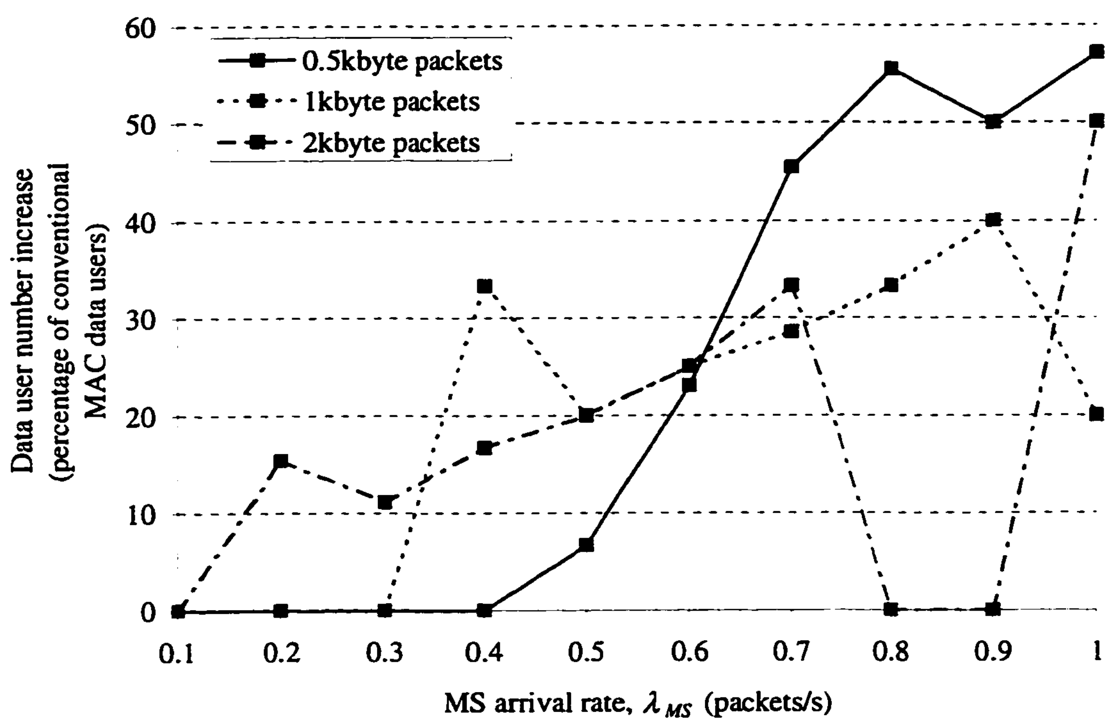

Figure 51. Increase in the number of data users with the new MAC scheme relative to the conventional MAC scheme.

The same trade-offs can be shown but with different magnitudes for the typical e-mail packet size distribution of Figure 29. Figure 52 and Figure 53 show the voice user number degradation and the data user number improvement respectively for this distribution. An increase in the number of data users (or decrease in voice user number) is noticeable over the whole range of MS arrival rate simulated and can reach 50 percent. Again the maximization of the integer *m* discussed in Section 6.2.2 leads to a decrease curve that goes up and down as shown in Figure 51.

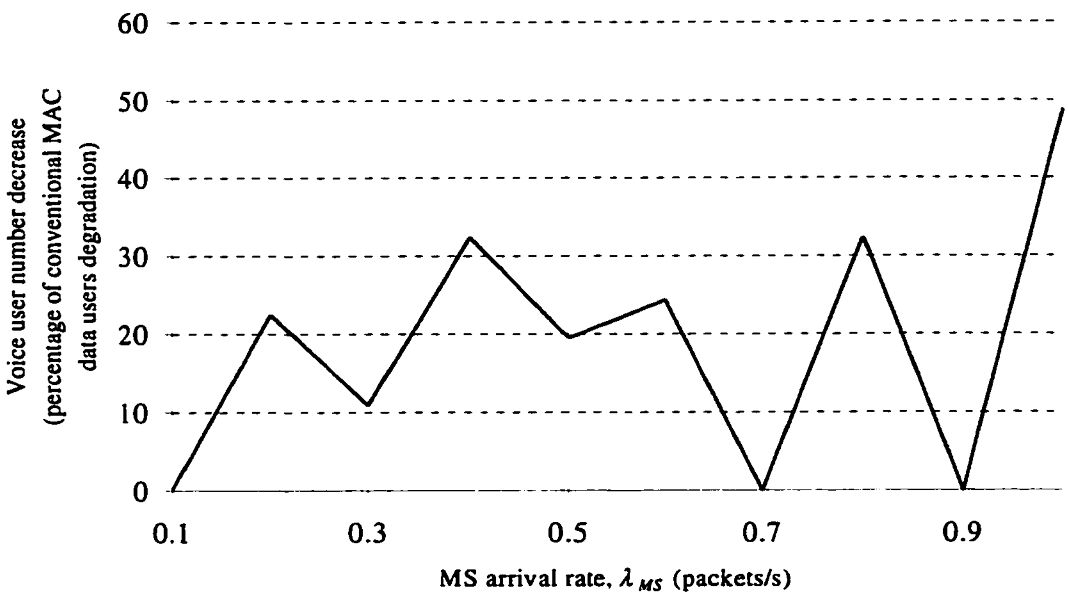

Figure 52. Voice user number decrease of the new MAC scheme relative to the conventional MAC scheme for the typical e-mail packet size distribution of Figure 29.

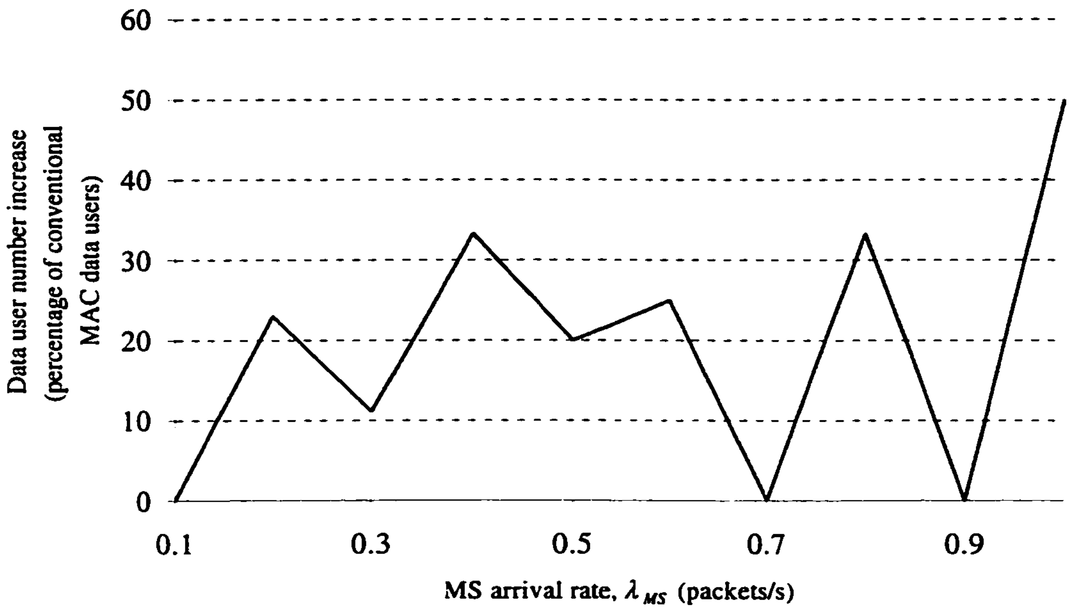

Figure 53. Data user number increase of the new MAC scheme relative to the conventional MAC scheme for the typical e-mail packet size distribution of Figure 29.

For the typical Internet packet size distribution of Figure 31, Figure 54 and Figure 55 show the voice user number degradation and the data user number improvement of the new MAC protocol respectively. The improvement in data users (or degradation in voice users) only starts for a MS arrival rate of 0.5 packets/s and is steadily increasing to about 65 percent. This slow take off of the curve is explained by the fact that for an arrival rate less than 0.5 packets/s, the number of ARCH time slots limits the capacity of the PDTCH. In that case, the ARCH time-slots are all used up before the PDTCH is fully exploited with both the new MAC and the conventional MAC scheme. This leads to the same performance for both MAC protocols.

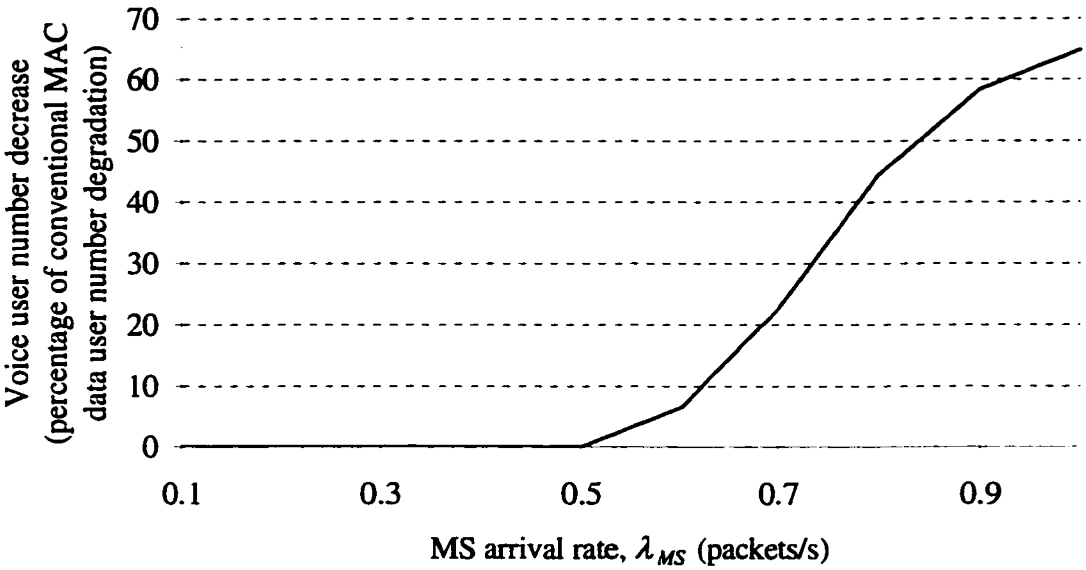

Figure 54. Voice user number decrease of the new MAC scheme relative to the conventional MAC scheme for the typical Internet packet size distribution of Figure 31.

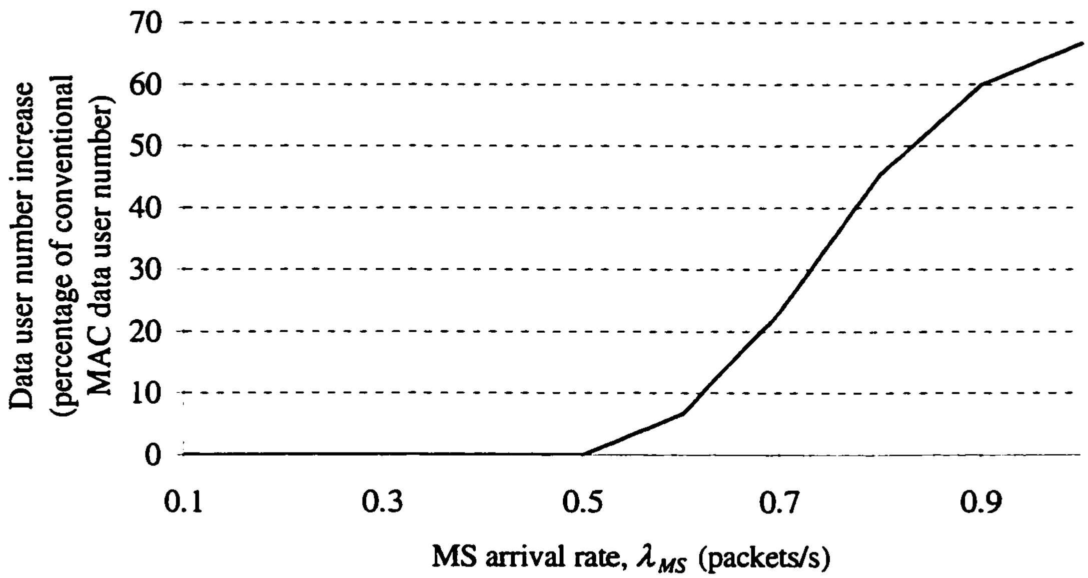

Figure 55. Data user number increase of the new MAC scheme relative to the conventional MAC scheme for the typical Internet packet size distribution of Figure 31.

6.4. Flexibility of new MAC protocol

The new MAC protocol can achieve a higher data user capacity than that of a conventional MAC scheme at the expense of a reduction in voice user capacity. This new MAC protocol can also achieve the same performance of a conventional MAC if the aggregate arrival rate is kept to the same maximum value of a conventional MAC scheme. The aggregate arrival rate is controlled by a call admission control scheme as described in Section 6.2.2. If the cellular operator wants an increase in data user capacity, the advantages of the new MAC scheme can be used to boost performance. The CAC scheme must then admit more users while keeping them to a value that still allows the quality of service to be satisfied. The price to pay for this extra performance is a decrease in the number of voice users. This shows the flexibility introduced by using the new MAC scheme. This MAC scheme allows a very flexible use of radio resource for cellular operators. They can use the extra data user capacity the new MAC protocol can achieve while switching back to the lower data user capacity of the conventional MAC scheme if the voice user degradation is not desirable.

6.5. Conclusion

This chapter showed that the use of packet-switched data transmission is much more efficient than circuit-switched data transmission. It was shown that the packet-switched approach saves more voice user circuits than the circuit-switched approach. By a proper use of a CAC scheme, the new MAC scheme can achieve a higher data user capacity than the conventional one at the expense of a lower voice user capacity. If a reduction in the number of the voice is not desirable, then the new MAC scheme can be operated with the same data and voice user capacity as a conventional one by control admission of users. This feature shows the flexibility that the new MAC protocol can bring to cellular systems. The flexibility introduced by the new MAC scheme is an asset for cellular network operators.

Chapter 7. Conclusion

The focus of this thesis has been a proposed new way to achieve more efficient uplink packet-switched data transmission in third generation DS-CDMA cellular systems. The 3G system considered was the European RTT proposal, called UTRA.

7.1. Thesis summary

This thesis describes a new MAC scheme using a time-shared channel, the PDTCH, applicable to third generation wideband DS-CDMA systems. The novelty of this new MAC protocol is that it admits new users while the previous ones are still being processed at the receiver, and this therefore increases data user capacity. Chapter 2 gives background information related to the thesis. This chapter introduces background information such as: DS-CDMA systems, forward error correction and automatic repeat request, the radio interface protocol architecture of cellular systems, issues related to packet-switched transmission in DS-CDMA, and relevant queueing models. Chapter 3 presents the new MAC protocol and its advantages. This new protocol comprises an access procedure as well as a retransmission procedure. Chapter 4 shows the performance analysis of the new scheme. For comparison purposes, the performance of the conventional MAC scheme, where the processing delay at the receiver is not used, is also shown. Chapter 5 presents simulation results as well as the numerical analysis. This chapter shows that analysis and simulations give close results. It is shown that, for short packets, the new scheme can handle a greater packet arrival rate (for a given mean packet delay) than a conventional scheme. The throughput efficiency of the new MAC scheme is only slightly lower than that of a conventional scheme. This new scheme permits a better utilization of the PDTCH than the conventional scheme. It also treats MSs in a fair manner. Chapter 6 uses the previous results to perform a capacity analysis of the new MAC scheme and the conventional one in a cellular context. The cellular system assumed is a system with packet-switched users and voice users. First, this chapter shows that packet-switched approaches, provided by either the new MAC scheme of the conventional one, definitely lead to better spectrum efficiency than a circuit-switched approach. Then this chapter shows the trade-off between the increased performance the new scheme can provide and the interference it creates to voice users. Finally, the flexibility of the new MAC protocol is discussed, and it is shown that this flexibility should be an asset for cellular systems.

The new MAC scheme may be applicable to packet data transmission in third generation DS-CDMA systems, such as CODIT or FRAMES mode 2.

7.2. Future work

The main focus of the thesis is to show that the proposed new MAC protocol for packet-switched data transmission can achieve a higher data user capacity than current proposed MAC protocols. This increase is, however, accompanied by an increase in the interference seen by voice users. A good question to answer now is how to use the flexibility provided by this new scheme. What kind of call admission control algorithm must be used to exploit this flexibility? As seen in the thesis, the new MAC protocol is mainly geared towards the transmission procedure since it uses processing delay at the receiver. In order to improve its performance, it would be interesting to pursue further research on ways to improve the access procedure. As shown by the results, the number of slots of the ARCH sometimes limits the performance of the MAC scheme. Therefore, ways to increase this number are desirable. Finally, the new MAC scheme proposed has been studied in the context of third generation systems that all use DS-CDMA as an access technique. It can be noted, however, that using the processing delay to admit more users is not specific to DS-CDMA. Therefore, it would be interesting to see how this new MAC scheme can be used in commercial second-generation TDMA systems.

References

[1] "UMTS/IMT-2000, Experimental WCDMA System Description," Ericsson Radio Systems AB, 1997.

[2] "Statistics on Cellular Subscriber in EU-15," European Commission, DG XIII-B.4, May 1998.

[3] "Cellular Subscriber Penetration in EU-15," European Commission, DG XIII-B.4, May 1998.

[4] "Information on Subscribers of Cellular Telephone, Pager, and Personal Handy-Phone (PHS) in Japan," Policy Report, Ministry of Post and Telecommunications, 1998.

[5] G. Brasche and B. Walke, "Concepts, Services, and Protocols of the New GSM Phase 2+ General Packet Radio Service", *IEEE Communications Magazine*, August 1997, pp. 94-104.

[6] W. C. Y. Lee, *Mobile Cellular Telecommunications Systems*, New York: McGraw Hill Book, 1989.

[7] T. Rappaport, *Wireless Communications Principles and Practice*, New Jersey: Prentice Hall 1996.

[8] J. S. DaSilva and B. E. Fernandes, "The European Reseach Program for Advanced Mobile Systems", *IEEE Communications Magazine*, February 1995, pp. 14-19.

[9] W. C. Y. Lee, "Overview of Cellular CDMA", *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 2, May 1991, pp. 291-302.

[10] A. Baier, U.-C. Fiebig, W. Granzow, W. Koch, P. Teder, and J. Thielecke, "Design Study for a CDMA-Based Third-Generation Mobile Radio System," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 4, May 1994, pp. 733-743.

[11] P.G. Andermo, "CODIT Final Report Review," R2020/ERA/PM/DS/P/050/b1, September 1995.

[12] A. Sasaki, "A Perspective of Third Generation Mobile Systems in Japan", Proceedings of the ACTS Mobile Telecommuniations Summit, Granada, Spain, Nov. 27-29, extra booklet.

[13] T. Nakamura, J. Hagiwara, E. Nakano, K. Ohno, and S. Onoe, "Configuration and Characteristic Estimation of W-CDMA Experimental System for Third Generation Mobile Communications", Proceedings of the *Vehicular Technology Conference*, May 18-21, Ottawa, Ontario, Canada, pp. 973-977.

[14] J. Shwarz da Silva, B. Arroyo-Fernandez, B. Barani, J. Pereira, D. Ikonomou, "Evolution Towards UMTS," European Commission, DG XIII-B.4.

[15] E. G. Tiedemman, Y.-C. Jou, and J. P. Odenwalder,"The Evolution of IS-95 to a Third Generation System and to the IMT-2000 Era", Proceedings of the *ACTS Mobile Communications Summit '97*, 7-10 Oct. 1997, Aalborg, Denmark, pp. 924-929.

[16] Radio Transmission Technology Submissions, International Telecommunication Union, June 30, 1998.

[17] D. Bertsekas and R. Gallager, *Data Networks*, 2nd edition. Englewood Cliffs, New Jersey: Prentice Hall, 1992.

[18] R. L. Peterson, R. E. Siemer, D. E. Borth, *Introduction to Spread Spectrum Communications*, Englewoods Cliffs, New Jersey: Prentice-Hall, 1995.

[19] K. S. Gilhousen, I. J. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, and C. E. Wheatley III, "On the Capacity of a Cellular CDMA System", *IEEE Transactions on Vehicular Technology*, vol. 40, no. 2, May 1991, pp. 303-312.

[20] S. Sun and W. Krzymien, "Matched Filter Based Synchronizer Design for Rapid Synchronization of Packet Data Transmissions in DS-CDMA Personal Communications Systems", Proceedings of the *PACRIM '97*, 20-27 Aug. 1997, Victoria, British Columbia, Canada, pp. 5-8.

[21] "An Overview of the Application of Code Division Multiple Access (CDMA) to Digital Cellular Systems and Personal Cellular Networks", Qualcomm Inc., 1992.

[22] S. Lin, D. J. Costello Jr., *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, New Jersey: Prentice Hall, 1983.

[23] J. G. Proakis, Digital Communications, 3rd edition, New York: McGraw Hill, 1995.

[24] R. D. Cideciyan, E. Eleftheriou, and M. Rupf, "Concatenated Reed-Solomon/Convolutional Coding for Data Transmission in CDMA-Based Cellular Systems", *IEEE Transactions on Communications*, vol. 45, no. 10, Oct. 1997, pp. 1291-1303.

[25] S. Sun, W. Krzymien, and A. Jalali, "Reliable Hybrid ARQ Data Transmission with Optimal Signal-to Interference Ratio Assignment in CDMA PCS", accepted for publication in the *ICUPC '98*, 5-9 October, Florence, Italy.

[26] "Submission of Proposed Radio Transmission Technologies", ETSI-SMG2, June 30 1998.

[27] EIA/TIA/IS-99, "Data Services Option Standard for Wideband Spread/Spectrum Digital Cellular System", 1995.

[28] "The cdma2000 ITU-R RTT Candidate Submission (0.18)", TIA, 1998.

[29] P. R. Chevillat, R. D. Cideciyan, M. Rupf, and W. Schott, "A Packet Data Service for CODIT UMTS," Proceedings of the *RACE Mobile Telecommunications Summit*, Cascais, Portugal, Nov. 1995.

[30] S. B. Wicker, *Error Control Systems for Digital Communications and Storage*, Englewoods Cliffs, New Jersey: Prentice Hall, 1995.

[31] H. Holma, F. Ovesjö, E. Dahlman, M. Latva-aho, A. Toskala, "Physical Layer of FRAMES Mode 2 - Wideband CDMA", Proceddings of the *Vehicular Technology Conference '98*, 18-21 May 1998, Ottawa, Ontario, Canada, pp. 978-982.

[32] S. Sun, W. Krzymien, A. Jalali, "Performance Comparison of Two Medium Access Control Schemes for Packet Data Traffic in DS-CDMA Personal Communications Systems", in Proc. of the *IEEE GlobeCom'97/CTMC'97*, Nov. 1997, Phoenix, USA, pp. 185-189.

[33] F. Ovesjö, E. Dahlman, T. Ojanperä, A. Toskala, A. Klein, "FRAMES Multiple Access Mode 2 - Wideband CDMA," Proceedings of the *Personal Indoor and Mobile Rardio Communications Conference `97*, 1-4 September 1997, Helsinki, Finland, pp. 42-46.

[34] I. M. Onyszchuk, "Finding the Complete Path and Weight Enumerator of Convolutional Codes," TDA Progress Report, Jet Propulsion Laboratory, pp. 203-213, February 1990.

[35] M. Mangal, J. A. Roberts, and D. L. Machamer, "Performance Analysis of a DS-CDMA Packet Cellular/PCS Network in a Multi-Rate Multi-User Environment", ", Proceedings of the *Vehicular Technology Cconference '98*, 18-21 May 1998, Ottawa, Ontario, Canada, pp. 1461-1465.

[36] S. Aridhi, C. Despins, D. Skellern, "Error Control Performance for Data Transmission in DS-CDMA Correlated Rayleigh Fading Channels," Proceedings of the *ACTS Mobile Communications Summit '97*, 7-10 Oct. 1997, Aalborg, Denmark, pp. 623-627.

[37] S. B. Wicker, "High-Reliability Data Transfer Over the Land Mobile Radio Channel Using Interleaved Hybrid-ARQ Error Control", *IEEE Transactions on Vehicular Technology*, vol. 39 , no. 1, Feb. 1990, pp. 48-55.

[38] R. J. McEliece and L. Swanson, "On the Decoder Error Probability for Reed-Solomon Codes", *IEEE Transactions on Information Theory*, vol. IT-32, no. 5, Sept. 1986, pp. 701-703.

[39] R. Hang, W. A. Krzymien, and C. Despins, "A MAC Protocol for Efficient Uplink Data Transmission of Short Packets in Wideband DS-CDMA", Proceedings of the *ACTS Mobile Communication Summit '98*, Rhodes, Greece, June 1998.

[40] "WAN Packet Size Distribution", National Library for Applied Network Research, June 1997.

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6″

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989