Performance Analysis of TETRA and TAPS and Implications for Future Broadband Public Safety Communication Systems

Christian Hoymann, Dirk Kuypers, Peter Sievering,
Peter Stuckmann, Bernhard Walke, Bangnan Xu

Chair of Communication Networks, Aachen University
Kopernikusstr. 16, D–52074 Aachen, Germany
walke@comnets.rwth-aachen.de

Abstract

This paper discusses the traffic performance of TETRA, TAPS, and a new approach for broadband ad hoc communication systems. The results show the performance of TETRA and TAPS regarding multimedia applications for public safety forces and what will be possible with selforganizing broadband networks.
The results were obtained with three simulation tools developed at the chair. The TETRA simulator TETRIS implements the protocols of the TETRA air interface, a channel model and load generators for data and voice applications, and provides simulation results for *Quality of Service* (QoS) measures like random access delay or transmission delay, see Section 1.
Because TAPS is an adaptation of (E)GPRS for TETRA, our GPRS simulator was used. Simulation results for quality of service measures of TAPS for the different applications and GPRS system measures are based on the simulation tool GPRSim that models the application and user behavior, the TCP/IP and WAP protocol architecture, the GPRS protocol architecture and the radio channel, see Section 2.
Finally, design issues of self-organizing broadband wireless networks are examined in Section 3. As centralized solutions suffer from many inherent disadvantages, the responsibilities of organizing and controlling of self-organizing networks should be fully distributed among wireless stations themselves. *Wireless CHannel oriented Ad-Hoc Multihop Broadband* (W-CHAMB) ideas that meet QoS demands for high performance services and realize statistical multiplexing of bursty traffic in a fully distributed and efficient manner are described in detail. The superiority of the performance of the W-CHAMB network can be seen in comparison with that of packet-oriented IEEE 802.11 WLAN. The hidden station problem is completely resolved by means of the energy signal (E-signal) solution. The performance gain of E-signal solution over the RTS/CTS mechanism is evaluated through computer simulation. Finally, the effect of network connectivity on the traffic performance is discussed.

1 Performance Evaluation of TETRA

TETRA Release 1 offers circuit switched speech services and connectionless or connection-oriented data services with data rates at about 4.8 kbit/s. Multimedia applications with higher bandwidth demands are not expected to be served by TETRA Release 1.

This section presents simulation results and provides a performance measure of the TETRA V+D air interface (reference point U_m). Results regarding the maximum number of users which can be served by TETRA systems are provided in [2]. Starting with an outline of the ETSI scenarios, we then introduce the TETRA system simulator TETRIS and conclude with a presentation of the traffic performance results.

1.1 ETSI Scenarios

The TETRA designer's guide [3] describes ten scenarios for the comparison of TETRA systems. For each scenario detailed specifications have been laid down concerning speech activity and offered data traffic of the mobile end user. Furthermore, the channel model to be used, the size of the scenario area, the number and type of the mobile stations, mobile or hand radio terminal, and their maximum velocity have been defined. Due to the fact, that scenario 10 defines the highest amount of offered load per terminal, this scenario has been chosen for the performance analysis in this section. Scenario 10 describes the parameters of a public or private network for airlines ground services, airport security, fire brigades and so on. Table 1

Table 1: Scenario 10 – General parameters

Parameter	Value
Type of area	BU
Covered area	50 km^2
Subscriber density	50 km^{-2}
Subscriber distribution	Gaussian
Class of terminals	80% portable, 20% vehicle
Velocity	3–50 km/h
Grade of Service	5%

Table 2: Scenario 10 – Traffic per radio user

Parameter	Value
Speech activity	$A_s = 20\,\mathrm{mE}$
Call duration	$\bar{\beta}_s = 20\,\mathrm{s}$
Mean waiting time	$\bar{\tau}_w = 4\,\mathrm{s}$
Speech arrival rate	$\lambda_s = 3.6\,\mathrm{h}^{-1}$
Short data (100 byte) arrival rate	$\lambda_{sd} = 20\,\mathrm{h}^{-1}$
Middle data (2 kbyte) arrival rate	$\lambda_{md} = 0.5\,\mathrm{h}^{-1}$

depicts the general parameters defined by scenario 10.

As can be seen from Table 2, the total traffic load per mobile terminal is:

$$\lambda = \lambda_s + \lambda_{sd} + \lambda_{md} = 24.1\,\mathrm{h}^{-1} \qquad (1)$$

The speech arrival rate λ_s has been calculated using $\lambda_s = A_s/\bar{\beta}_s$. The mean waiting time is defined as the duration between the dialling of a subscriber number or pressing the *Push-to-Talk* (PTT) button and the successful completion of the call set-up.

Sixty percent of the voice calls are assumed to be group calls, the mean group size is 20 mobile terminals. We assume that the 2500 radio users are distributed over four TETRA cells providing full indoor coverage. Configurations with 400, 500, 600, 700, and 800 users per radio cell are evaluated, taking into account a non-uniform distribution of the radio users over the four cells.

TETRA systems allow queueing of call set-ups and the Erlang C formula is applicable for trunking capacity estimation under the call queueing strategy. Due to the mean number of 625 radio users per cell the total mean offered speech traffic is $A = 12.5\,\mathrm{E}$. To reach a call delay probability of $p_D = 5\,\%$ at least 20 traffic channels and 1 control channel are required, i.e. at least 6 carrier frequencies, accordingly 150 kHz spectrum per radio cell, are required.

1.2 Simulation Concept

For the traffic performance evaluation of the TETRA protocol stack the protocols of the air interface at reference point U_M have been implemented because they are the key elements.

The structure of the simulator TETRIS is depicted in Figure 1. The protocol stack of the TETRA V+D

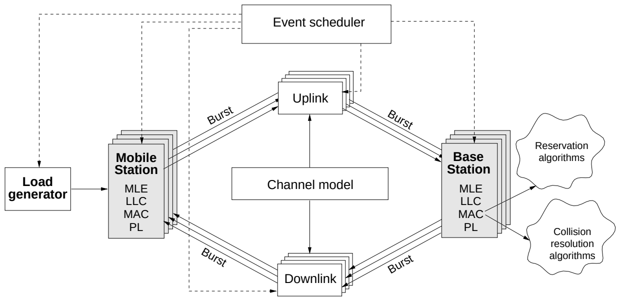


Figure 1: Structure of the simulation environment

system has been specified with the help of *Formal Description Technique* (FDT) to guarantee not only syntactically and semantically unambiguous formal descriptions of the communication protocols but also interoperable and compatible implementations of these protocols independent of their implementation [6]. The *Specification and Description Language* (SDL) is the most widely used FDT in the area of telecommunications [5]. With the help of the C++ code generator `SDL2SPEETCL`, which converts SDL phrase representation to C++ source code, the mobile and base station protocol stacks have been embedded in the C++ simulation environment. The C++ implementations are based on the *SDL Performance Evaluation and Tools Class Library* (SPEETCL)[1]. The SPEETCL provides generic C++ classes as well as a simulation library with strengths in random number generation, statistics evaluation, and event driven simulation control.

The core of the simulator is the simulation control, which creates mobile and base stations and assigns the traffic generators to create specific traffic loads to the individual mobile stations. Depending on the scenario (see Section 1.1) the traffic generators are controlled to offer a certain traffic load. A traffic load is defined by inter-arrival times and the size of the data units.

Mobile and base stations communicate via uplink and downlink channels by exchanging bursts. With the help of error pattern files transmission errors can be introduced on the uplink or downlink dependent on the actual C/I value at the respective receiver. The pattern files have been generated taking into account a channel propagation model, the characteristics of the TETRA physical layer, and the receiver characteristics.

Having a TETRA network comprising of four radio cells, co-channel interference is unlikely. Thus error-free transmission is assumed in this paper. Mobility and the mobility management protocol are not taken into account because of the low terminal speed in comparison to the size of a TETRA cell. For further studies on the TETRIS simulation tool please refer to [4].

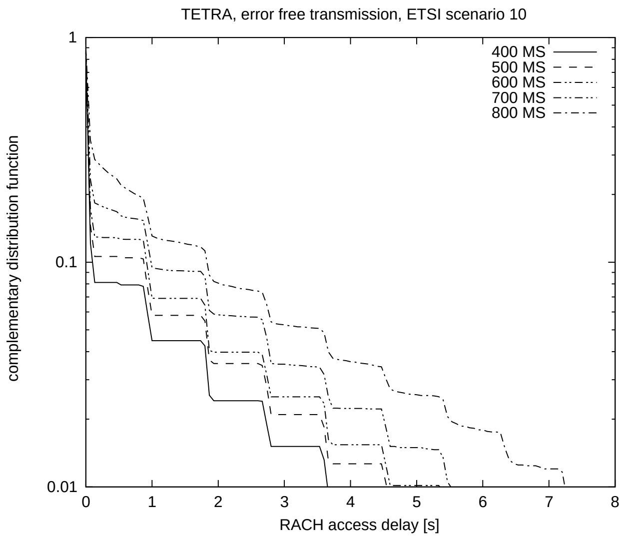


Figure 2: RACH access delay

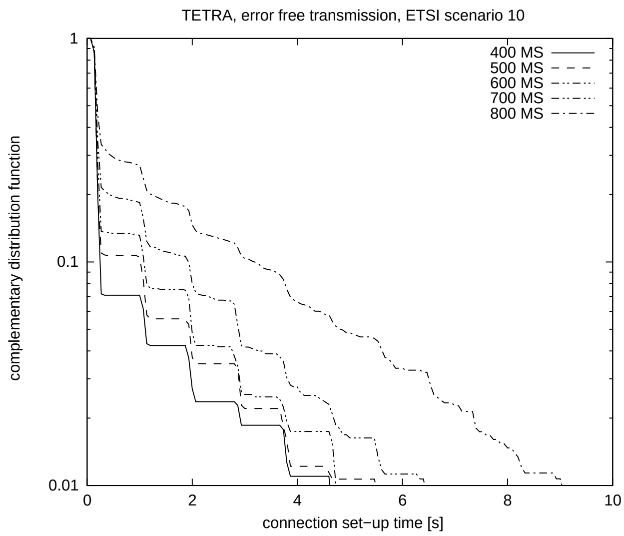


Figure 3: Connection set-up time

1.3 Performance Measurements

Figure 2 shows the complementary distribution function of the RACH access delay for the traffic load parameters of scenario 10. The RACH is used to activate point-to-point connections as well as group communications.

The access delay is defined as the duration between the creation of a connection set-up request and the reception of the acknowledgement of a successful connection set-up sent by the BS. The access delay is influenced by the structure of the logical channels and the collision resolution algorithm.

If the TETRA BS cannot assign capacity to an MS within a preset time after a successful access on the RACH, the MS accesses the RACH again to repeat the capacity request. In the simulation results presented, this time-out is set to 18 TDMA frames. Thus, the access delay, as shown in Figure 2, also takes into account the retransmission on the RACH due to this time-outs.

With a higher traffic load offered, the probability for a repetition of the capacity request is increased because of the limited number of traffic channels. In case of 400 MSs this delay probability is about 8 % and with 800 MSs it becomes more than 20 %.

The connection set-up time, as shown in Figure 3, includes both, the access delay and the waiting time for the assignment of a traffic channel. The measurements include point-to-point as well as group communications.

Because of the low collision probability the RACH access delay for the first access is small and the connection set-up time is mainly influenced by the waiting time for a free traffic channel.

1.4 Conclusions

The results for error-free transmission show that fast call establishment times of about 300 ms are only achievable with low numbers of mobile stations or by adding more control channels. This gets even worse in high traffic load situation or if transmission errors have to be considered.

2 Performance Evaluation of TAPS

The *TETRA Advanced Packet Service* (TAPS) is the new high speed packet data service for TETRA (Release 2). It is heavily based on the (E)GPRS standards for GSM to enable cost effective and efficient use of network resources. The changes introduced are concerned with the matching of the frequency bands. We use our (E)GPRS system simulator GPRSim to get an idea of the performance that can be expected of enhanced data services and multimedia applications with TAPS.

After *High Speed Circuit Switched Data* (HSCSD) has been introduced in some countries in 1999, the first GPRS-based services have been available since 2001 in Europe. Many countries worldwide will introduce GPRS in the next years. With these new services mobile data applications with net bit rates of up to 117 kbit/s will be offered and established on the market. To realize higher data rates the *European Telecommunications Standards Institute* (ETSI) and the *3rd Generation Partnership Project* (3GPP) have developed the *Enhanced Data Rates for GSM Evolution* (EDGE) standard, which offers a net bit rate of up to 384 kbit/s by means of modified modulation, coding and medium access schemes (see [10, 17, 14]).

In parallel to the GSM evolution, the data applications performed by mobile users will evolve. In the first phase of the GSM evolution, where the data services *Circuit Switched Data* (CSD) and GPRS are available, WAP-based applications as defined in [22] running on smart phones and PDAs besides conventional Internet applications running on laptop computers or enhanced PDAs will dominate. Then Video Streaming applications (see [8]) and *Large Data Transfer* (LDT) applications including the *Multimedia Message Service* (MMS) based on WAP version 2.0 as defined in [23] are felt to become more popular with the optimization of GPRS and with the introduction of EDGE and the related packet data service *Enhanced GPRS* (EGPRS).

While for the time period right after the service introduction minimal configurations were chosen supporting only a basic availability of GPRS, with increasing data traffic load in the next years GSM/GPRS cell capacity will have to be extended. For this evolution of GSM/GPRS networks and for the introduction of EGPRS, dimensioning guidelines are needed for operators, equipment manufactures and system integrators. They should describe the relationship between the offered traffic and the radio resources to be allocated to reach a desired quality of service for the different applications (see [20, 19, 15, 14]).

This part of the paper aims at presenting simulation results for two predicted traffic mixes, one for a GPRS evolution scenario and one for an EDGE introduction scenario. The first one is composed of WAP, WWW and e-mail, the second is defined by Streaming, WWW and e-mail.

In Section 2.1 the potential applications and the related traffic models are introduced. After the description of the simulation tool GPRSim in Section 2.2 the traffic performance results are presented and interpreted in Section 2.3.

2.1 Applications and Traffic Models

This section describes the traffic characteristics that are expected in 2.5 and 3G mobile radio networks. After the presentation of traffic models for the conventional Internet applications WWW and e-mail, WAP applications are depicted. Finally an introduction into Streaming applications and the related traffic models are given.

2.1.1 WWW

All applications summarized by *World Wide Web* (WWW) are based on the *Hypertext Transfer Protocol* (HTTP), which uses the TCP/IP protocol stack. HTTP organizes the transfer of *Hypertext Markup Language* (HTML) documents (web pages).

WWW sessions consist of requests for a number of pages. These pages consist of a number of objects with a certain object size. Another characteristic parameter is the delay between two pages depending on the user's behavior to surf around the Web (see [7, 9]). Table 3 gives an overview of the WWW traffic parameters. The small number of objects per page (2.5 objects), and the small object size (3700 byte) were chosen, since Web pages with a large number of objects or large objects are not suitable for thin clients such as PDAs or smart phones served by (E)GPRS. The traffic characteristics of the WWW model can be seen in the distribution functions of the object size.

2.1.2 E-mail

E-mails are transmitted by using the *Simple Mail Transfer Protocol* (SMTP) or the *Post Office Protocol version 3* (POP3) for e-mail download. Since the size of an e-mail download on a mobile device is the crucial parameter for this research, a traffic model defining e-mail sizes is suitable. The introduced e-mail model based on [12] describes the load arising with the transfer of messages performed by an SMTP user. The only parameter is the e-mail size that is characterized by two $\log_2$-normal distributions plus an additional fixed quota of 300 byte (see Table 3). The base quota was assumed to be a fixed overhead. Subtracting the overhead, a bimodal distribution remained. The lower 80 % were said to be text-based mails, while the upper 20 % represent mails with attached files, which can be rather large. The transition between these two distributions is 2 kbyte. The maximum e-mail size is set to 100 kbyte.

2.1.3 WAP

The WAP specifications, which are the basis for the implementation in today's mobile terminals, including the June 2000 Conformance Release, also known as WAP 1.2.1, aim at optimizing the operation in 2G networks. Therefore WAP 1.2.1 defines a distinct technology comprising protocols and content representation. WAP is a suite of specifications that defines an architecture framework containing optimized protocols (e.g., WDP, WTP, WSP), a compact XML-based content representation (WML, WBXML) and other mobile-specific features like *Wireless Telephony Applications* (WTA) as defined in [22].

WAP Release 1.x In addition to the goal of the optimized operation in 2G networks, WAP has been developed because today's graphics-enhanced web services cannot be brought to and displayed on thin clients, e.g., GSM mobile phones, and IP as the network layer may not be applicable in some environments, e.g., WAP over *Short Message Service* (SMS) or *Unstructured Supplementary Service Data* (USSD).

Because of the optimizations and different protocols it is not possible to run WAP end-to-end to a regular Internet site. Instead, a WAP Gateway must be used.

Table 3: Traffic model parameters

WWW Parameter	Distribution	Mean	Variance
Pages per session	geometric	5.0	20.0
Intervals between pages [s]	negative exponential	12.0	144.0
Objects per page	geometric	2.5	3.75
Object size [byte]	$\log_2$-Erlang-k ($k = 17$)	3700	$1.36 \cdot 10^6$
e-mail Parameter	**Distribution**	**Mean**	**Variance**
e-mail size (lower 80 %) [byte]	$\log_2$-normal	1700	$5.5 \cdot 10^6$
e-mail size (upper 20 %) [byte]	$\log_2$-normal	15700	$62.9 \cdot 10^9$
Base quota [byte]	constant	300	0
WAP Parameter	**Distribution**	**Mean**	**Variance**
Decks per session	geometric	20.0	3800
Intervals between decks [s]	negative exponential	14.1	198.8
Size of 'Get Request' packet [byte]	$\log_2$-normal	108.2	$4.1 \cdot 10^3$
Size of 'Content' packet [byte]	$\log_2$-normal	511.0	$3.63 \cdot 10^5$

The main services a WAP Gateway provides is protocol conversion between WAP stack and Internet stack. In addition to this standardized functionality, many gateway vendors provide a variety of value-added services that allow for personalization, for example.

WAP Release 2.0 In the specification WAP 2.0 as defined in [23] some existing WAP protocols have been extended by new capabilities. WAP 2.0 converges with widely used Internet protocols like the *Transmission Control Protocol* (TCP) and the *Hypertext Transfer Protocol* (HTTP). *Internet Engineering Task Force* (IETF) work in the *Performance Implications of Link Characteristics* (PILC) Working Group has been leveraged to develop a mobile profile of TCP for wireless links. This profile is fully interoperable with the common TCP that operates over the Internet today. Further, WAP 2.0 does not require a WAP proxy, since the communication between the client and the server can be conducted using HTTP 1.1. However, deploying a WAP proxy can still optimize the communication process and may offer mobile service enhancements, such as location, privacy, and presence based services. In addition, a WAP proxy remains necessary to offer Push functionality.

In addition to protocol work, the WAP Forum has continued its work on service-enabling features for the mobile environment, like the Push service or synchronization issues. Although WAP 2.0 has been finished in 2001, WAP 1.x protocol stacks will still be used in the mobile terminals in the next years. In this paper, only WAP 1.x is regarded.

WAP Traffic Model A WAP traffic model has been developed and applied in [16, 18].

A WAP session consists of several requests for a deck performed by the user. The maximum amount of data that can be transferred by one request defaults to 1400 bytes. The parameters are summarized in Table 3. The main characteristic is a very small mean packet size (511 byte) modelled by a $\log_2$-normal distribution with a limited maximum packet size of 1400 byte (see Table 3).

2.1.4 Video Streaming

Many Internet portal sites are offering video services for accessing news and entertainment content from a *Personal Computer* (PC). Beside *Motion Picture Expert Group* (MPEG), H.263 is the currently most accepted video coding standard for Video Streaming applications. In the near future, mobile communication systems are expected to extend the scope of today's Internet Streaming solutions by introducing standardized Streaming services as described in [8].

In the scope of modelling video sources, a lot of attention has been paid to long range dependent or self-similar models of traffic streams in telecommunication networks (see [21]). Many of such models have been used to investigate *Variable Bit Rate* (VBR) video sources with a statistical analysis of empirical sequences and estimation of the grade of self-similarity (see [13]). Since MPEG and H.263 video traffic consists of a highly correlated sequence of images due to its encoding, the correct modelling of the correlation structure of the video streams is essential (see [24]).

In this work no stochastical models of video streams with self-similar or high-correlated traffic characteristics are applied. Real video sequences coded by an H.263 coder are used to generate the Streaming traffic.

The Video Streaming traffic model used within the scope of this work is based on three video sequences in the format *Quarter Common Intermediate Format* (QCIF) with the resolution of 176×144 pixels. The sequences are proposed by the *Video Quality Expert*

Sequences	offered IP traffic	
	Q20	80-10-10 Mix
`Claire`	10.9 kbit/s	
`Carphone`	26.7 kbit/s	**14.39 kbit/s**
`Foreman`	31.7 kbit/s	

Table 4: Offered IP traffic of video sequences

Group (VQEG) and are for this reason commonly used. Each sequence is representing a particular group of videos with different intensities of motion.

- `Claire` stands for a very low motion intensity and can be seen as a characteristic video conferencing sequence or inactive visual telephony.

- `Carphone` includes both, periods with rather high motion and periods of low motion intensity. It represents many kinds of vivid or active video-conferences or even visual telephony.

- The third video, `Foreman`, is a sequence with permanently high motion intensity of both, the actor and the background. This permanent motion is characteristic for sport events or movie trailers.

The H.263 coder was used with a skip factor of 2, which means that every second frame of the original sequence was skipped so that the frame rate of the coded sequences was reduced from 25 to only 12.5 frames/s.

The quantization level 20 (Q20) was adjusted for *Intra* (I)- and *Predictive* (P)-frames. The resulting video quality is marginal. But it is acceptable for mobile devices with its limited visual output capacities.

A conservative mix of sequences including 80 % `Claire`, 10 % `Carphone` and 10 % `Foreman` has been selected for the simulations performed. The mix shall represent video streams with low motion and only a few streams with higher motion intensity.

Due to the negligible size of *Real-Time Streaming Protocol* (RTSP) and *Real-Time Protocol* (RTP) control messages in comparison to the size of real-time data, they have been neglected. The resulting average IP traffic offered by this particular mix is 14.39 kbit/s (see Table 4).

Beside visual telephony all of the new emerging applications are relatively short in duration. So called *heavy users*, generating long streams with huge amounts of data, have not been taken into account. The duration of video sessions is modelled by a negative-exponential distribution with an average value of 60 s. This is an assumption with regards to the prognosis for 3G networks in [9] where the duration of real-time calls is proposed to be modelled by a negative-exponential distribution.

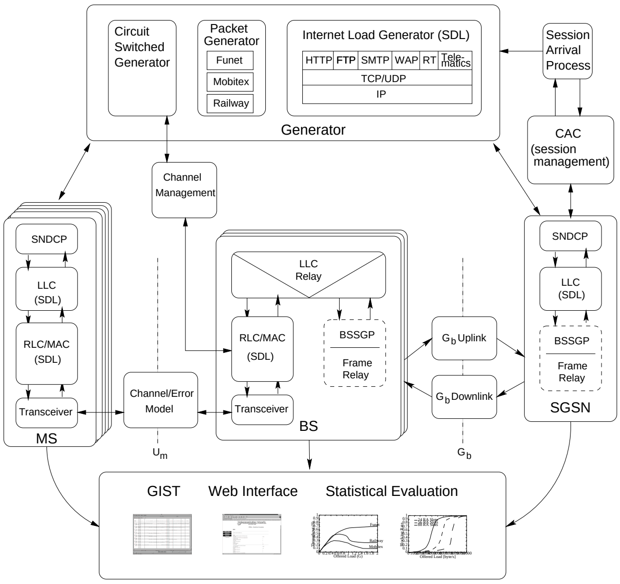


Figure 4: The (E)GPRS Simulator GPRSim

2.2 Simulation Concept

The full details of the GPRS protocol stacks of the radio and the fixed network and of the Internet protocols including the characteristics of TCP currently cannot be described by formulas usable in practice. Since GPRS networks are presently introduced in the field, traffic engineering and related performance results are needed soon, so that capacity and performance estimations become possible for GPRS/EDGE introduction and evolution scenarios.

Measuring the traffic performance in an existing GPRS network is not possible, since a scenario with a well-defined traffic load is hard to set-up, the evaluation of the performance by measurement is very difficult, and the analysis of different protocol options is not possible in an existing radio network.

Therefore computer simulation based on the prototypical implementation (called emulation) of the standardized GPRS protocols and the Internet protocols in combination with traffic generators for the regarded applications and models for the radio channel are chosen as the methodology to get the needed results rapidly.

The (E)GPRS Simulator GPRSim is a pure software solution based on the programming language C++. Up to now models of *Mobile Station* (MS), *Base Station* (BS), *Serving GPRS Support Node* (SGSN), and *Gateway GPRS Support Node* (GGSN) have been implemented. The simulator offers interfaces to be upgraded by additional modules (see Figure 4).

For the implementation of the simulation model in C++ the *Communication Networks Class Library* (CNCL) (see [11]) is used, a predecessor to the *SDL Performance Evaluation Tool Class Library* (SPEETCL) presented in [6]. This enforces an object oriented structure of programs and is especially suited for event driven simulation.

Different from usual approaches to establish a simulator, where abstractions of functions and protocols are being implemented, the approach of the GPRSim is based on the detailed implementation of the standardized GSM and (E)GPRS protocols. This enables a realistic study of the behavior of EGPRS and GPRS. The real protocol stacks of (E)GPRS are used during system simulation and are statistically analyzed under a well-defined and reproducible traffic load.

The complex layers of the protocol stacks like SNDCP, LLC, RLC/MAC based on (E)GPRS Release 99, the Internet traffic load generators and TCP/IP itself are specified formally with the *Specification and Description Language* (SDL), translated to C++ code by means of the Code Generator `SDL2CNCL` (see [6]) and finally integrated into the simulator.

2.3 Performance Measurements

2.3.1 Simulation Scenario Parameter Settings

The cell configuration is given by the number of *Packet Data Channels* (PDCHs) permanently available for GPRS. In this paper 1, 4, 6 and 8 fixed PDCHs have been regarded. For the GPRS simulation series a C/I of 12 dB (13.5% BLEP) has been regarded and Coding Scheme 2 (CS-2) has been used. For EGPRS the channel conditions are determined by the cell and cluster size that are the basis for the C/I calculation as described in [17]. Cluster size 7, a cell size with a radius of 3000 meters and a velocity of 6 km/h has been regarded. Both *Link Adaptation* (LA) and *Incremental Redundancy* (IR) are applied.

LLC and RLC/MAC are operating in acknowledged mode for WWW, e-mail and WAP and in unacknowledged mode for Streaming. The multislot capability is 1 uplink and 4 downlink slots. The MAC protocol instances in the simulations are operating with 3 random access subchannels per 52-multiframe. All conventional MAC requests have the radio priority level 1 and are scheduled with a FIFO strategy. LLC has a window size of 16 frames. TCP/IP header compression in SNDCP is performed. The maximum IP datagram size is set to 1500 byte for WAP and 552 byte for the TCP-based applications. In the Internet stack for WWW and e-mail TCP is operating with a maximum congestion window size of 8 kbyte. The transmission delay in the core network and external networks, i. e., the public Internet is neglected, since it is assumed that the servers are located in the operator's domain and the core network is well dimensioned. Since the high round-trip time in GPRS networks is mainly caused by *Temporary Block Flow* (TBF) establishment procedures at the air interface, the delay in well dimensioned IP subnetworks does not have a great effect on the end-to-end performance.

2.3.2 Performance and System Measures

To characterize the traffic performance of GPRS several performance and system measures are defined in the following. For the different applications different critical performance measures will be regarded, since different performance characteristics are required for transaction-oriented applications and real-time applications, respectively.

Mean IP throughput per user is the downlink IP throughput measured during transmission periods, e. g., the download period of a single object of a web page. This is an important QoS parameter from a user's point of view. The statistical evaluation of this measure is done by counting the amount of IP bytes transmitted in each TDMA frame period for each user, if a packet train is running. Thus, the throughput is not averaged over inactive periods. The number of IP bytes transmitted divided by the TDMA frame duration represents a simulation sample value in the evaluation sequence. At the end of the simulation the mean throughput is calculated from this evaluation sequence.

Mean application response time is the difference between the time when a user is requesting a web page, a WAP deck or an e-mail and the time when it is completely received.

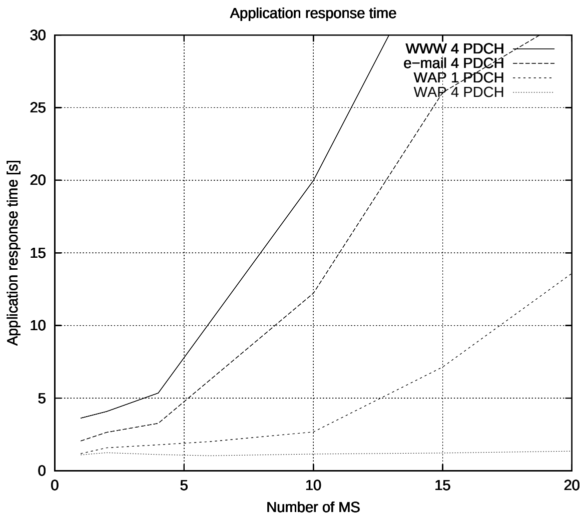
Mean IP datagram delay is the end-to-end delay of IP datagrams evaluated by means of time stamps given to the datagrams, when the IP layer performs an SNDCP data request for transmission. When the datagram arrives at the receiver, the difference of the actual time and the time stamp value is calculated as a sample of the respective evaluation sequence.

Mean throughput per cell is also called system throughput and is calculated from the total IP data transmitted on all PDCHs of the regarded radio cell and for all users during the whole simulation duration, divided by the simulation duration. Since a loss of IP datagrams over fixed subnetworks is not modelled, this parameter equals the offered IP traffic in the radio cell.

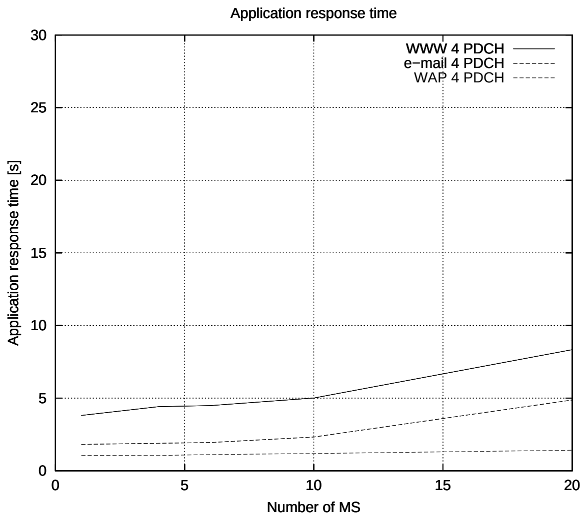
PDCH utilization: is the number of MAC blocks utilized for MAC data and control blocks normalized to the sum of data, control and idle blocks. Thus existing capacity reserves in the scenario under consideration can be seen from this measure.

2.3.3 WAP in Comparison to Internet Applications over GPRS

To be able to compare the user-perceived performance of WAP in comparison to conventional Internet ap-



(a) Pure WWW/e-mail and WAP traffic



(b) Traffic Mix

Figure 5: Mean application response time

plications, the application response time is shown in Figure 5(a) for pure WWW, e-mail and WAP traffic.

In situations with low traffic load the response time for a WAP deck is below 2 s, while the response time for a web page is around 4 s. The reason is that a web page has a larger content size and is transmitted over TCP.

In load situations with higher traffic load the response time for a WAP deck remains nearly constant for up to 20 MS. If only 1 PDCH is available, the WAP response time increases to more than 10 s for 20 MS in the radio cell. Because of the larger content size the response time for web pages passes 20 s already with 10 active MS in the radio cell even if 4 PDCHs are available. The reason for the strong increase in response time for WWW and e-mail can be seen in other evaluated measures like the downlink PDCH utilization in Figure 7(a). 100 % PDCH utilization is reached for WWW/e-mail traffic with 15 MS, while 15 WAP users are only utilizing the PDCHs with 30 % for the same PDCH configuration.

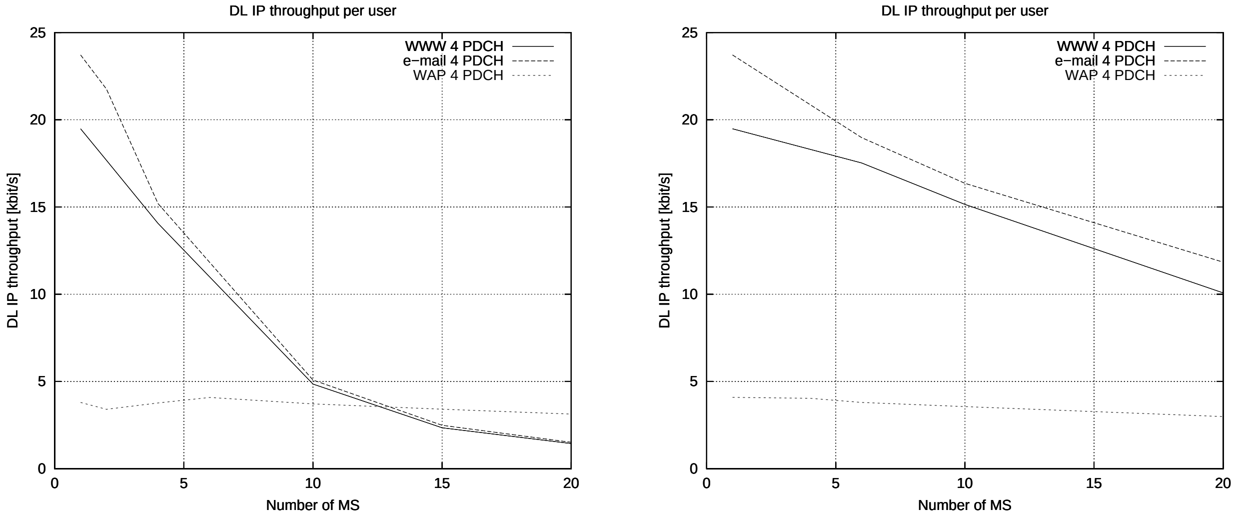
Figure 6(a) shows the mean downlink IP throughput per user during transmission periods. While the throughput performance for pure WAP traffic remains relatively constant with an increasing number of mobile stations and 4 PDCHs, it decreases dramatically for pure WWW/e-mail traffic because of the higher offered traffic and the higher utilization. The poor throughput performance for WAP traffic can be explained by the low WAP deck size. Such transaction-oriented applications are more influenced by the high round-trip-time, which is mainly caused by the high delay over the air interface, than by the available bit rate. Since the response time for a WAP deck is less than 1.5 s, which should be acceptable for a wireless application, the user is not aware of this low throughput performance.

Since WWW and e-mail applications comprise larger file sizes to download than WAP-based applications do, the throughput performance perceived by a user in situations with low traffic load ranges from 14 to 24 kbit/s. These performance values are mainly influenced by the characteristics of the offered traffic. Since the e-mail traffic model has larger file sizes than WWW, the throughput performance is better. With an increasing number of mobile stations up to 15 the saturation is reached and the performance for WWW and e-mail users gets unacceptable and even gets worse than the low throughput for pure WAP traffic. In this situation with high traffic load the WWW and e-mail traffic performance is less influenced by the characteristics of the traffic model like the file size, but by the load on the air interface.

2.3.4 Traffic Mix with WAP and WWW/e-mail over GPRS

Since the predicted traffic mix for GPRS networks will be composed of WAP traffic and conventional Internet applications like WWW and e-mail, the GPRS traffic performance for a traffic mix of 60 % WAP, 28 % e-mail and 12 % WWW sessions will be regarded, here.

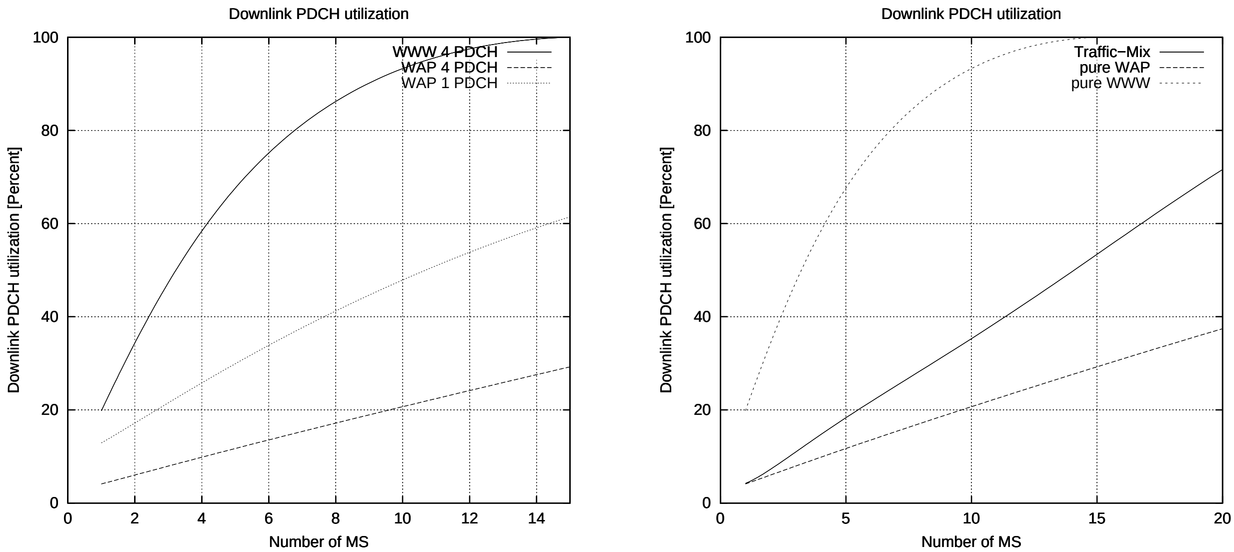
Figure 5(b) shows the application response time for WAP decks, e-mails and WWW pages, respectively. Compared to the graphs in the previous section, the WWW and e-mail performance is not strongly affected by WAP traffic, since small WAP packets can be multiplexed seamlessly with the TCP-based



(a) Pure WWW/e-mail and WAP traffic

(b) Traffic Mix

Figure 6: Mean downlink IP throughput per user



(a) Pure WWW/e-mail and WAP traffic

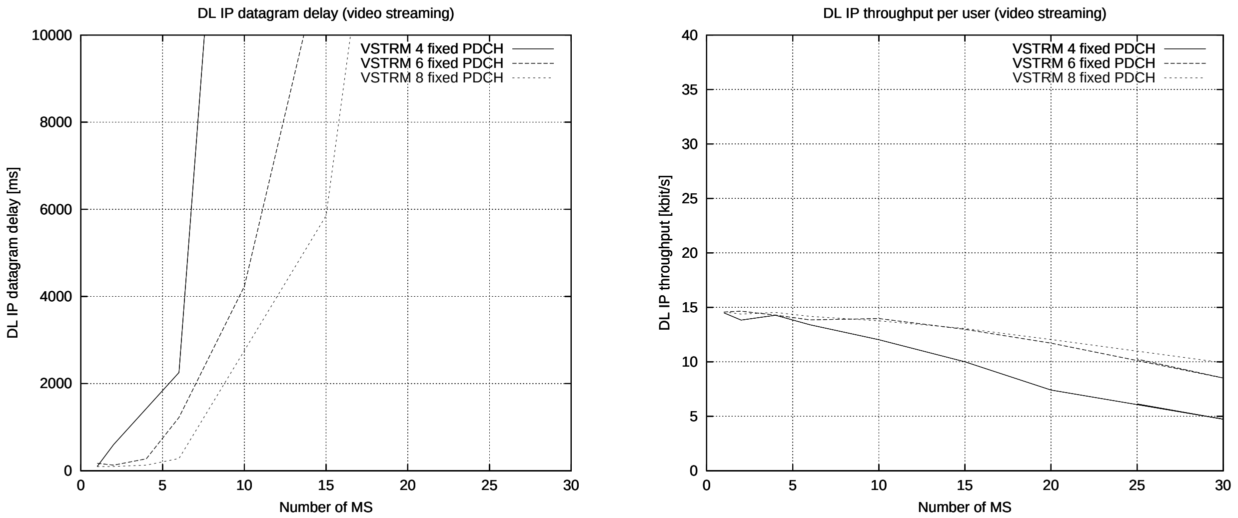
(b) Traffic Mix compared to pure WWW and pure WAP traffic

Figure 7: Mean downlink PDCH utilization

WWW and e-mail traffic. The throughput (see Figure 6(b)) decreases slower with an increasing number of mobile stations than in Figure 6(a) with pure WWW, e-mail and WAP traffic regarded separately, since here WAP represents the main part of a traffic mix and the total offered traffic per radio cell is increasing much slower.

The same applies for the response time. In the scenario with traffic mix WWW pages have a response time of 5 s with 10 active stations generating a traffic mix, while 10 stations generating pure WWW traffic have to wait for more than 20 s.

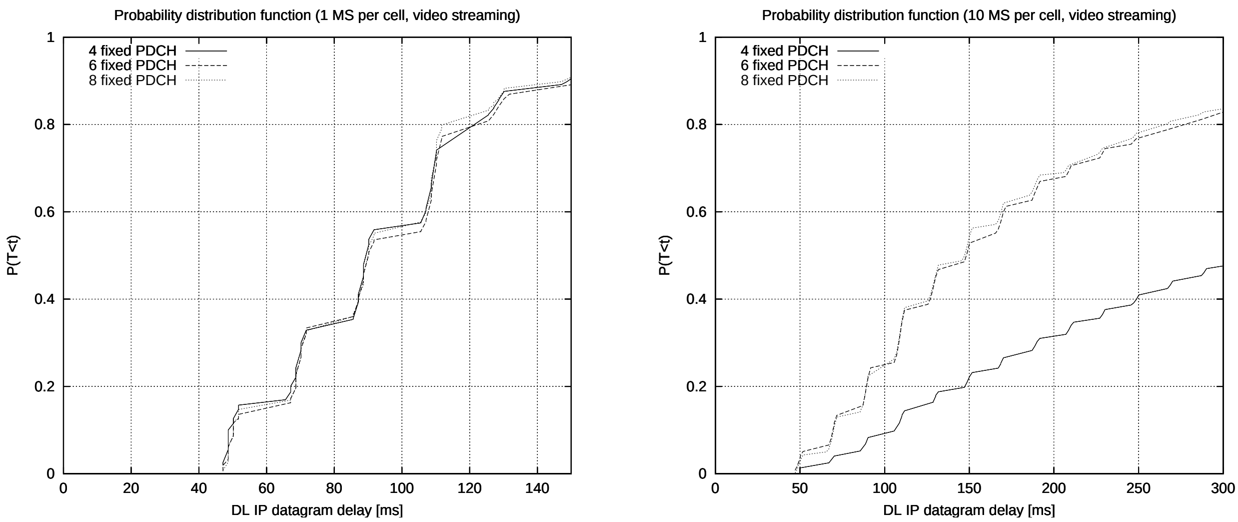
The WAP response time increases slightly from 1.2 s for pure WAP traffic to 2.1 s for the traffic mix scenario. The reason is that WWW and e-mail sessions are composed of larger application packets that leave less resources open for WAP users (see Figure 7(b)). Nevertheless a response time for WAP decks of 2.1 s



(a) Mean downlink IP datagram delay

(b) Mean downlink IP throughput per user

Figure 8: Performance of Video Streaming applications (traffic mix)



(a) 1 mobile station

(b) 10 mobile stations

Figure 9: Distribution of the downlink IP datagram delay for Video Streaming applications (traffic mix)

still should be acceptable.

2.3.5 Traffic Mix with Video Streaming Applications over EGPRS

As a typical EGPRS introduction scenario, Video Streaming applications over EGPRS are examined in coexistence with WWW and e-mail applications. Due to the conservative predictions concerning the future usage of Streaming applications the mix only contains 10 % Streaming sessions. The remaining part is assumed to 63 % e-mail and 27 % WWW session. As the critical performance measure the downlink IP datagram delay is regarded (see Figure 8(a)). Additionally the downlink IP throughput per user in Figure 8(b) indicates, if the Streaming data rate of 14.39 kbit/s can be maintained under a regarded number of mobile stations offering the Multimedia traffic mix. In addition the distribution of the downlink IP datagram delay for 1 and 10 mobile stations is shown in Figure 9(a) and 9(b).

With 4 available PDCHs in the regarded radio cell the IP datagram delay for Streaming increases dramatically to more than 10 s with more than 6 active stations that generate a traffic mix. With more than 6 stations the throughput of 14.39 kbit/s can not be maintained, which shows that the performance for Streaming users becomes unacceptable. With 6 and 8 PDCHs available in the cell 15-20 users generating the traffic with 10 % Streaming can be satisfied. The delay starts increasing dramatically with 15 and 20 users, respectively.

Regarding the downlink IP throughput per user in Figure 8(b) there is no significant difference in the performance between 6 and 8 available PDCHs. The throughput for 4, 6 and 8 PDCHs start at the same level of 14.39 kbit/s. This is exactly the data rate needed for the chosen video sequence. The downlink IP throughput per user is remaining constant as long as the necessary data rate for Streaming is provided. Depending on the number of fixed PDCHs the real time data rate is decreasing below the required rate of 14.39 kbit/s. At this point the IP datagram delay is increasing dramatically. With 15 users the required data rate can not be maintained any more.

The distribution functions in Figure 9(a) and 9(b) confirm these interpretation. For one mobile station 90 percent of the IP packets for the Streaming applications are delivered within 150 ms. The performance is not depending on the number of PDCHs available, since the regarded mobile stations can use only maximum 4 slots on the downlink. For 10 mobile stations and 4 PDCHs available more than 50 percent of the IP packets are delayed more than 300 ms and the slow increase of the distribution function indicates a high delay variance, which makes the delay performance for Streaming applications unacceptable. With 6 and 8 PDCHs 85 % of the IP packets or Streaming applications are delivered within less than 300 ms, which makes the Streaming performance just acceptable for 10 users generating the Multimedia traffic mix. The steps in the distribution functions are affected by the segmentation of IP packets into radio blocks and the number of radio blocks transmitted within a GPRS radio block period of 20 ms.

The different requirements of the applications can be supported by *Quality of Service* (QoS) management functions in the RLC/MAC layer (see [15]). The transmission of Streaming data may be privileged on the expense of background traffic. While the application response times for WWW and e-mail would increase, the Streaming application would be able to proceed, although high background traffic load occurred in the cell.

2.4 Conclusions

In this section the performance of different Multimedia applications in packet-switched cellular radio networks based on GPRS and EGPRS is presented. As TAPS is based on GPRS the performance of TAPS can be deduced. For GPRS introduction and evolution scenarios WAP applications and a traffic mix of WAP and conventional Internet applications over GPRS are examined. After the performance characteristics of WAP and Internet applications have been regarded separately, the effects of coexisting Internet traffic on WAP traffic and vice versa are outlined. It has been shown that WAP traffic can be multiplexed seamlessly with the Internet traffic because of the small and limited WAP deck size, while Internet traffic slightly slows down WAP traffic in situations with high traffic load. Regarding Video Streaming applications in coexistence with TCP-based applications over EGPRS it has been shown that only a small number of Streaming users can be served by EGPRS, even if the percentage of Streaming in the traffic mix is low. At least more than 4 fixed PDCHs should be available to support Streaming applications together with background TCP traffic. Privileged transmission of real-time data, realized by QoS management, is one approach to provide the required bit rate for video streaming in situations with high traffic load.

3 Implications for Broadband Public Safety Communication Systems

Since a self-organizing network can work without any preexistent infrastructure, it can be rapidly deployed. This feature is especially beneficial for temporary application scenarios, such as short term events, extension of the radio coverage of fixed infrastructure radio networks, disaster relief and military applications. Meanwhile, such a network is very reliable as failure or departure of some wireless stations will not cause the failure of the whole network. Due to the simplicity, flexibility and low cost to deploy a self-organizing network, the interest in such kind of networks will be ever increasing.

3.1 Centralized vs. Decentralized

It seems that self-organizing means that no central control will be used in the network. So a self-organizing network should naturally be decentralized. This is fully true as long as high performance multimedia must not be supported. The controversy whether a self-organizing broadband wireless network should be controlled in a centralized or decentralized way arises due to the consideration that provision of QoS requirements may be easily realized by a central controller. Based on this consideration, many researchers have developed algorithms to select central controllers in a self-organizing environment. With a simple network topology, in which a central controller can be optimally selected and the other wireless stations can successfully receive the control information from the central controller, the centralized self-organizing network may function well. But in most cases, no matter how good the selection algorithm is, a centralized solution suffers from the following inherent problems:

1. To realize self-organization, most wireless stations should have an in-built central controller function. As a central controller in a broadband wireless network needs very high computing capacity, the hardware requirements on the wireless stations increase dramatically.

2. The network will be complicated and vulnerable. The failure or departure of the selected central controller will cause temporary chaos in the whole network.

3. Direct mode and multihop communication cannot be realized efficiently as communication is possible only under control of a central controller.

4. The scarce frequency spectrum cannot be used efficiently. Neighboring central controllers must use different frequencies. Dynamic channel allocation which is inherent in decentralized networks is not easy to perform.

5. A wireless station may not be able to associate because it may not be able to receive the information from the central controller or the central controller cannot hear this wireless station.

As the centralized solution for self-organizing broadband wireless networks suffers from the limitations described above, we are sure that the best way for self-organizing broadband wireless networks is still the decentralized solution. The responsibilities of the organizing and controlling of self-organizing networks should be fully distributed among the wireless stations themselves. Every wireless station decides by itself when and how to send its information according to predetermined algorithms and protocols.

As we decided to adopt the decentralized solution, we face another challenge, that is how to realize QoS guarantee in a fully distributed broadband wireless network.

3.2 Packet-oriented vs. Channel-oriented

3.2.1 The Packet-oriented solution: IEEE 802.11

Due to the worldwide great success of Ethernet, the packet-oriented *Carrier Sense Multiple Access* (CSMA) protocol is pervasive in the area of LAN. So the WLAN standard IEEE 802.11, viewed as a wireless extension of Ethernet, has also adopted the packet-oriented *Carrier Sense Multiple Access with Collision Avoidance* (CSMA/CA) access scheme [28]. A high speed physical layer in the 5 GHz Band, called IEEE 802.11a, has also been specified as a supplement to IEEE 802.11. The data rates of IEEE 802.11a will be up to 54 Mbit/s [26].

Although the *Distributed Coordination Function* (DCF) of IEEE 802.11 is fully decentralized and self-organizing, it is not, however, suitable for self-organizing broadband wireless networks because of its inefficiency and no means to guarantee QoS.

The success of the Ethernet is due to its simplicity. The increasing bandwidth requirements have been met by the high speed Ethernet of 100 Mbit/s and 1 Gbit/s. IEEE 802.11 WLAN, however, has a different transmission medium. It is impossible for IEEE 802.11 to apply CSMA/CD because the sending station can not detect any ongoing collision. So IEEE 802.11 can only use a CSMA/CA protocol. To achieve *Collision Avoidance* (CA), a large protocol overhead, such as backoff, is necessary. In addition, the existence of hidden stations in wireless environments makes the CSMA-like protocol very inefficient [32]. To overcome this problem, IEEE 802.11 has specified the RTS/CTS mechanism that increases the protocol overhead. It is worth mentioning that RTS/CTS can not solve the hidden station problem completely (see Section 3.3). Moreover, one of the main advantages of the packet-oriented CSMA protocol namely that the transmission of large packets can achieve high efficiency, is no longer valid in the wireless environment because the packet error probability will be higher for larger packets. Assume that the *Packet Error Rate* (PER) of a short packet is 3%, then the PER of a large packet of 10 short packets length will be $1 - (1 - 0.03)^{10} = 26\%$, which is no longer acceptable. So large packets have to be fragmented to short packets to achieve sufficient transmit reliability. Such a fragmentation will increase the overhead.

In our research, we have found that provision of QoS requirements of high performance multimedia applications in a packet-oriented self-organizing wireless network appears to be impossible.

3.2.2 The Channel-oriented solution: W-CHAMB

Inspired from GPRS and DECT concepts, we developed a channel-oriented solution — *Wireless CHannel-oriented Ad-hoc Multihop Broadband* (W-CHAMB) networks — for a self-organizing broadband wireless network. W-CHAMB adopts the key idea of GPRS, that is statistical multiplexing of bursty traffic through packet reservation, and the most advanced feature of DECT, that is dynamic channel selection according to the measured signal level *Radio Signal Strength Indicator* (RSSI) [20]. It differs from the GPRS and DECT with its ability to operate in a multihop environment and in a fully distributed manner. The most significant feature of W-CHAMB is that it meets QoS demands for different services and realizes statistical multiplexing of bursty traffic in a fully distributed and efficient manner.

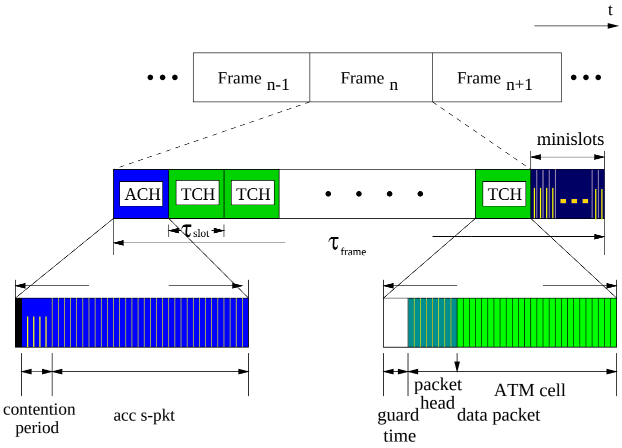


Figure 10: W-CHAMB Access structure

W-CHAMB channel access structure Transmission of packets in W-CHAMB networks is channel-oriented. The transmission time scale in W-CHAMB networks is organized in periodic frames, each containing a fixed number of time slots, see Figure 10. All *Wireless Stations* (WSs) of the network are synchronized on a frame and slot basis.

To use channel resources more flexibly for heterogeneous applications, periodic slots are used as physical channels to provide transmit capacity for several *Logical Channels* (LCH), e.g. LCH1/2 uses one slot every two frames, see Figure 11. The first slot of the frame is used as *Access Channel* (ACH), where a number of energy signals and an access signaling packet (*acc s-pkt*) can be transmitted. The other slots are used as physical *Traffic Channels* (TCHs), each for one data packet transmission per frame. Each data packet has a user data unit, such as an ATM cell, and a packet header containing information, such as packet identifier, sequence number, time information, etc., see [29].

At the end of the frame a number of minislots follow. Each minislot is associated with a TCH. If the TCH shall stay reserved, the receiving WS sends an energy signal (E-signal) on the corresponding minislot. The introduction of minislots may reduce the time available for message transmission. However, each minislot carries only a single on-off pulse of the unmodulated carrier, so that the related overhead is generally quite small. Moreover, as we use the same kind of E-signals to realize the distributed transmit priority, this solution does not increase hardware complexity. The E-signal is transmitted on the carrier frequency, so that the detection delay is negligible. The performance gain of this solution is evaluated in Section 3.3.

Dynamic Channel Reservation As an example, we demonstrate the procedure of *Dynamic Channel Reservation* (DCR) for S_1 to send an information burst to S_2, see Figure 12 and 11. At the beginning of

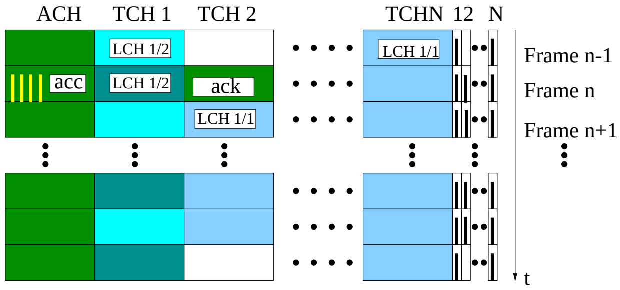


Figure 11: Dynamic channel reservation

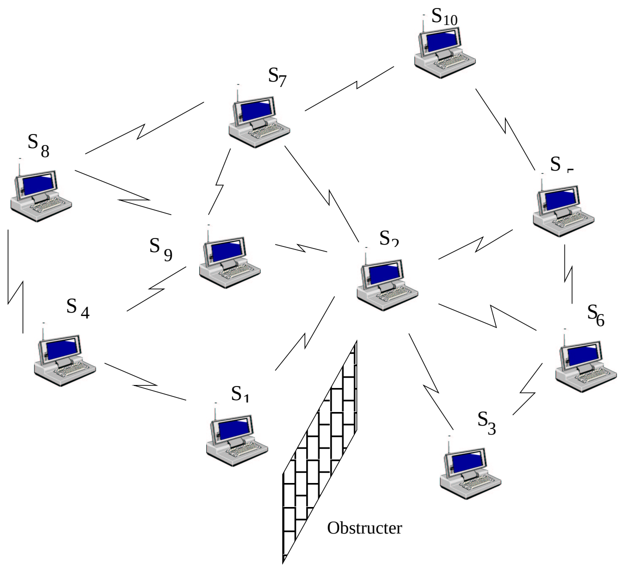


Figure 12: Self-organizing broadband wireless network

the channel reservation, S_1 contends for transmitting an *acc s-pkt* via the ACH using the distributed access priority. The *acc s-pkt* contains a set of free LCHs that have a low noise level and could be used in the view of S_1. In the event that S_1 did send the *acc s-pkt*, and the addressed station, e.g. S_2, could successfully receive this *acc s-pkt* and could find at least one of the LCHs proposed by S_1, to be also free in the view of S_2, it responds to S_1 with an acknowledgment (*ack*) *s-pkt* via the selected LCH, e.g. LCH1/1 on TCH 2 in Figure 11, and starts sending an E-signal on the corresponding minislot 2. By this procedure, a LCH is reserved between S_1 and S_2. All other WSs in the detection range of S_1 and/or S_2 will mark this LCH as reserved. The hidden station problem is resolved by the E-signal sent by S_2. At the end of its information burst, S_1 stops transmitting on the reserved LCH and S_2 stops sending the E-signal on the corresponding minislot. WSs in the range of S_1 and/or S_2 that detect then the TCH 2 unused will mark it in their local channel occupancy list free again.

Distributed Access Priority To be able to prioritize a real time VBR service, we define the QoS of real time VBR traffic services in terms of a maximum tolerable packet delay D_{max} and a packet dropping probability, P_{drop}. To give the more urgent information burst a higher access priority, we use a distributed access priority algorithm [30, 31].

Adaptive Back-off In spite of using the distributed access priority, a collision on the ACH may happen if: (1) two WSs use the same priority; (2) contending WSs are hidden to each other. To avoid repeated collisions, we use an adaptive back-off algorithm [30, 31].

ABR reservation interrupted on demand ABR traffic is multiplexed with rt-VBR traffic in the air interface. To use bandwidth efficiently, ABR services will use bandwidth resources which are temporarily not used by VBR services. As WSs reserve LCHs in an uncoordinated manner, an algorithm is necessary to ensure that a rt-VBR burst can always find a free LCH before its access deadline and at the same time the free bandwidth can be efficiently used by the ABR service. For that purpose any station receiving a request to open a LCH for a rt-VBR service will interrupt an ongoing ABR transmission and allocate the LCH freed by this to the rt-VBR service requested.

For the ABR traffic, the available bandwidth resources should be shared fairly among the WSs and should be used efficiently. On the one hand, it should be avoided that a WS reserves a LCH for a long time while other WSs may have no opportunity to send any ABR traffic. On the other hand, it should also be avoided that a WS is not allowed to transmit ABR traffic continuously even though there are enough free LCHs. To deal with this problem, a WS must check the channel occupancy situation at the end of each ABR burst transmission to evaluate the system traffic load. If the spectrum is highly loaded, the WS must back off for some time before it is allowed to apply for a LCH for ABR traffic again.

3.2.3 Performance comparison

Compared to the packet-oriented IEEE 802.11, the channel-oriented W-CHAMB LAN has the following advantages: (1) It achieves high network efficiency. (2) It supports rt-traffic in a fully distributed manner. With realistic Ethernet packet sizes with a mean of 434 bytes for an IEEE 802.11 WLAN at the transmission rate of 24 Mbit/s, the superiority of the performance of W-CHAMB can be seen from the simulation results in [34].

3.3 The hidden station problem

Hidden stations may result in extreme inefficiency in self-organizing wireless networks. A hidden station is a station that cannot sense the transmission of the sending WS, but will cause interference to the receiving WS if it transmits. Hidden stations can be caused by obstruction, see Figure 12. Assume that S_2 is receiving data from S_1. But S_3 cannot sense the transmission of S_1 because of the obstructer. If S_3 transmits at the same channel as used by S_1, S_2 will be interfered. S_3 is a hidden station in this case. Hidden stations may be caused by the multihop environment. Even if there is no obstructer, hidden stations can result from the different distances among wireless stations. Assume that S_2 is receiving from S_1 again. As S_5 is out of the detection range of S_1, it cannot sense the transmission of S_1. So S_5 is a hidden station which may cause interferences to S_2. Hidden stations may degrade the network performance substantially. There are two basic approaches to solve the hidden station problem:

1. The busy tone solution was firstly proposed in [32] to combat hidden stations in CSMA systems. A busy tone signal is sent by the receiving station on a narrow band channel to make a hidden station aware of an ongoing transmission and prevent it from transmitting and interfering. The limitation of the busy tone solution is the need of a separate channel, the need of additional hardware for the receiver to transmit the busy tone while receiving on the data channel and the large delay of detecting the busy tone in a narrow band channel. The last limitation makes it impossible to use the busy tone solution in broadband wireless networks.

2. The RTS/CTS (request to send/clear to send) mechanism as specified in IEEE 802.11 is able to combat hidden stations. A station that intends to send a data packet sends a RTS packet to the receiver. After reception, in turn, the receiver sends a CTS packet to indicate it is ready to receive data. Other stations that receive the RTS and/or CTS packet will defer their access for a period of time according to the transmission duration information contained in the RTS/CTS packets. The goal of the RTS/CTS mechanism is that hidden stations should receive the CTS packet and would cooperate then. But the RTS/CTS mechanism does not solve the hidden station problem completely. Some cases remain where hidden stations cannot receive the CTS packet. For example, see Figure 12, S_{10} cannot receive the CTS packet from S_2 because it is out of the decode range of S_2. But S_{10} can cause interference to S_2 as S_{10} is still in the interference range of S_2. The interference range is usually much larger than the

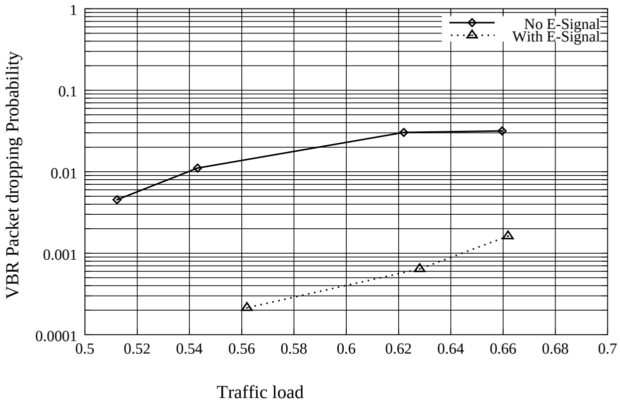


Figure 13: VBR packet dropping probability, $c = 0.58$

> decode range. In some cases, even though a station is in the decode range of the receiver, it may not be able to receive the CTS packet because of the interference of other stations. For example, S_9 is in the decode range of S_2. But if S_8 is sending while S_2 transmits the CTS packet, S_9 cannot receive the CTS of the S_2. So S_9 may access the channel after S_8 ended the transmission, which causes interference to S_2.

In the W-CHAMB network, the hidden stations problem is solved completely through E-signals transmitted on minislots, see Section 3.2.2. Although the introduction of minislots does increase the system overhead, the performance gain of using E-signal can be derived from Figure 13. The network throughput is defined as the number of successfully transmitted data packets divided by the simulated time (counted in slots). The simulated network consists of 20 WSs with a connectivity of 0.58, see Section 3.4. The length of a minislot is assumed 10% of a normal slot. ABR traffic and rt-VBR traffic are mixed, each 50%. The packet error rate is assumed to be 3%. If the E-signal is not used, all WSs that receive an *ack s-pkt* keep silent for the duration of the transmission so that the *ack s-pkt* has a function similar to CTS packets [34]. So the performance resulting in Figure 13 indicates that the RTS/CTS mechanism is not enough to solve the hidden station problem.

3.4 Impact of network connectivity

To study the impact of the network connectivity on the traffic performance, we consider a 5×5 square grid network with 25 wireless stations. A desired network connectivity is achieved by adjusting the fixed transmit power of the wireless stations accordingly. The connectivity is defined as the mean number of neighbors to a WS, normalized by the number of the maximum possible number of neighbors. $c = \frac{1}{N(N-1)} \times \sum_{i=1}^{N} n_i$, where n_i is the number of neighbors to station i, N is the number of stations in the network. This means a fully connected network has a connectivity of 1.

The packet error rate depends on the *Carrier-to-Interference Ratio* (C/I). We assume the same physical layer as defined for H/2 and use the results of [33] concerning the relation between the C/I and the packet error rate. We assume that the power at the distance γ from the transmitter is $W = k\gamma^{-\alpha}$, where k is a constant for all stations. A typical value for WLAN environments is $\alpha = 4$.

Each wireless station produces a Poisson traffic stream and randomly selects another station as its traffic sink. The burst length of traffic is geometrically distributed with mean of 30 packets, each sent in one time slot. The packet size is 54 bytes. The interarrival time of bursts is negative exponentially distributed. The mean value of the interarrival time is adjusted to meet the traffic load to the whole network. In our simulation, we decided to drop packets that have exceeded a delay of 15000 slots, e.g. 300 ms.

Figure 14 and 15 show the impact of the network connectivity on the network throughput and the mean end-to-end packet delay. With a connectivity of 0.93, the network has the highest traffic performance. There, the throughput is increased linearly with the traffic load until a traffic load of about 0.71, where the network becomes saturated. Many packets are dropped then due to large delay. At the traffic load of 0.84, the network throughput declines by about 8% owing to changes of lengths of the connections counted in hops. It can be seen that the traffic performance is reduced with a smaller connectivity. With a connectivity of 0.24, the network is saturated at a traffic load of 0.31. The strong decline of the throughput at the load of 0.38 is due to a increased mean number of hops of each connection still served. The mean delay versus traffic load at different connectivities is shown in Figure 15. It can be seen that if the network is lightly loaded, the mean delay increases slowly with increased traffic load. The mean delay increases significantly if the network approaches saturation. With $c = 0.93$, the mean delay indicates saturation at a traffic load of 0.71. Under higher traffic load, the mean delay increases further, but with a reduced slope since packets with large delay are dropped there and are not counted in the mean delay. The significant change of the packet delay when the network approaches saturation can also be seen from the *Complementary Distribution Functions* (CDF) of end-to-end packet delay with the different traffic loads at the connectivity of 0.93, see Figure 16 . It can be seen that there are two groups of delay distribution curves. One comprises distributions with the traffic loads under 0.71. The other those with traffic loads above 0.71. As the network is lightly loaded under traffic loads of 0.64, data packets experience smaller delay. Above traffic loads of 0.71, the network is saturated. So data packets experience larger delay. All the curves decline at

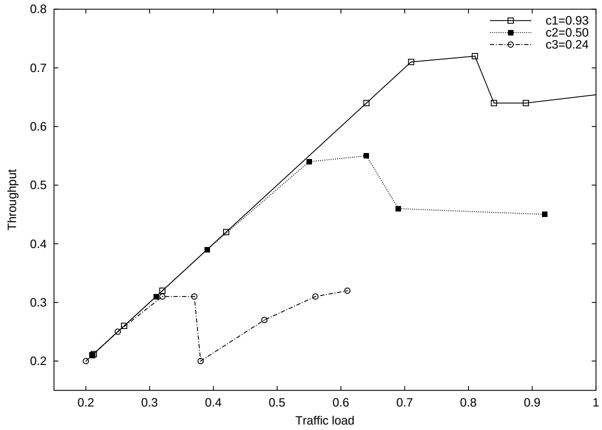


Figure 14: Throughput vs. traffic load

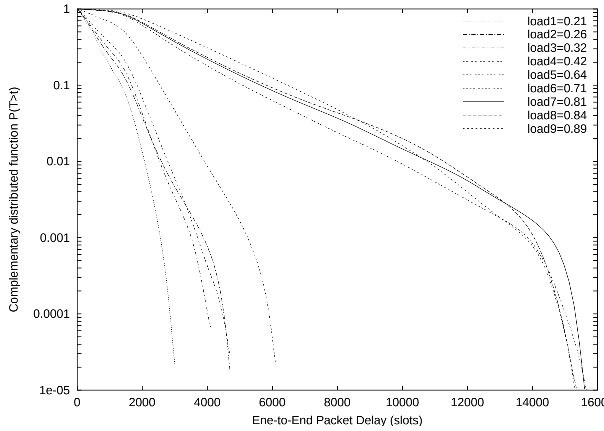


Figure 16: End-to-end packet delay CDF, $c = 0.93$

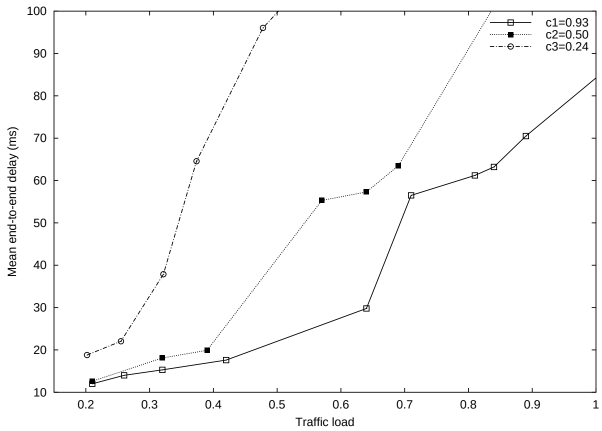


Figure 15: End-to-end packet mean delay vs. traffic load

the delay of about 15000 slots (300 ms) dramatically since packets that exceed that delay are dropped at the source stations. The delay of dropped packets is not considered in the curves. Some packets may experience a delay larger than 15000 slots as packets that have left the source station are not dropped in the relay stations.

Although frequency spatial reuse is possible in networks with a small connectivity and has been considered in our simulations, the system capacity is used up rapidly owing to the multihop transmissions needed for end-to-end connections. The benefits of frequency spatial reuse in multihop networks is adverse to that of cellular networks that only use one hop per connection. It can be derived from the results presented that in multihop networks, frequency spatial reuse does not remarkably increase the network capacity. A comparison of the saturation throughputs of 0.71 and 0.31 under connectivities of 0.93 and 0.24 indicates that a factor of about 4 in connectivity does correspond to a factor of about 2.3 in saturation throughput. This is due to the frequency reuse.

3.5 Conclusion

We have presented some ideas and solutions for the design of self-organizing broadband wireless networks. We have shown that QoS cannot be guaranteed to high performance multimedia applications in a packet-oriented self-organizing wireless network.

Channel oriented packet transmission has proven to be appropriate to control QoS in a self-organizing wireless network. The maximum number of hops of a connection should be limited to achieve a reasonable traffic performance. Another lesson we learned is that a network with decentralized control is best suited for the operation of a self-organizing broadband wireless network. It can be expected that the centralized design of H/2 phase 1 will reach its limitations soon if the number of users in the 5–6 GHz unlicensed frequency spectrum will increase. So an extension of H/2 to a fully decentralized self-organizing network appears to be necessary in the future.

The traffic performance of existing TETRA and (E)GPRS systems give lower bounds for achievable delays and throughput in broadband communication systems.

References

[1] AixCom GmbH. http://www.aixcom.com. 1.2

[2] Martin Steppler. *Maximum Number of Users Which Can Be Served by TETRA Systems.* In Bernhard Walke, editor, *Tagungsband der European Wireless '99*, Vol. 157 of *Reihe ITG-Fachbericht*, pp. 315–320, München, VDE-Verlag, 6th–8th October 1999. 1

[3] ETSI. (RES06.1). *Scenarios for Comparison of Technical Proposals for MDTRS.* Working document, European Telecommunications Standards Institute, Sophia Antipolis, France, June 1991. 1.1

[4] Martin Steppler. *TETRIS—A simulation tool for TETRA systems.* In *Mobile Kommunikation*, Vol. 135 of *ITG-Fachbericht*, pp. 403–410, Berlin, Offenbach, Informationstechnische Gesellschaft (ITG) im Verein Deutscher Elektrotechniker (VDE), VDE-Verlag GmbH, September 1995. 1.2

[5] Martin Steppler, Matthias Lott. *SPEET – SDL Performance Evaluation Tool.* In Ana Cavalli, Amardeo Sarma, editors, *Proceedings of the 8th SDL Forum '97*, pp. 53–67, Evry, France, Elsevier Science Publishers, 23rd–26th September 1997. ISBN 0-444-82816-8. 1.2

[6] Martin Steppler. *Performance Analysis of Communication Systems Formally Specified in SDL.* In *Proceedings of The First International Workshop on Simulation and Performance '98 (WOSP '98)*, pp. 49–62, Santa Fe, New Mexico, USA, 12th–16th October 1998. 1.2, 2.2

[7] Arlitt, M. and Williamson, C. (1995). A Synthetic Workload Model for Internet Mosaic Traffic. In *Proc. of the 1995 Summer Computer Simulation Conference*, pages 24–26, Ottawa, Canada. 2.1.1

[8] Elsen, I., Hartung, F., Horn, U., Kampmann, M., and Peters, L. (2001). Streaming Technology in 3G Mobile Communication Systems. *IEEE Computer*, 34(9):46–52. 2, 2.1.4

[9] ETSI 3GPP (1998). Selection Procedures for the Choice of Radio Transmission Technologies of the Universal Mobile Telecommunication System UMTS (UMTS 30.03, 3G TR 101 112). Technical report, European Telecommunications Standards Institute, Sophia Antipolis, France. 2.1.1, 2.1.4

[10] Furuskär, A., Mazur, S., Müller, F., and Olofsson, H. (1999). EDGE: Enhanced Data Rates for GSM and TDMA/136 Evolution. *IEEE Personal Communications*, pages 56–65. 2

[11] M. Junius et al. (1993). CNCL: a C++ library for event driven simulation, statistical evaluation and random number generators and distributions. Technical Report, Communication Networks, Aachen University of Technology. 2.2

[12] Paxson, V. (1994). Empirically-derived analytic models of wide-area tcp connections. *IEEE/ ACM Transactions on Networking*, 2(4):316–336. 2.1.2

[13] Rose, O. and Frater, M. (1994). A comparison of models for VBR video traffic sources in B-ISDN. *IFIP Transactions C-24:Broadband communications II*, pages 275–287. 2.1.4

[14] Stuckmann, P. (2002). *The GSM Evolution - Mobile Packet Data Services.* John Wiley & Sons, Chichester. 2

[15] Stuckmann, P. (2002). Quality of Service Management in GPRS-based Radio Access Networks. *Telecommunication Systems 19:3, Kluwer Academic Publishers*, pages 515–546. 2, 2.3.5

[16] Stuckmann, P., Finck, H., and Bahls, T. (2001). A WAP Traffic Model and its Appliance for the Performance Analysis of WAP over GPRS. In *Proc. of the IEEE International Conference on Third Generation Wireless and Beyond (3Gwireless '01)*, San Francisco, USA. 2.1.3

[17] Stuckmann, P. and Franke, J. (2001). The Capacity and Performance Gain Reachable with Link Quality Control in EGPRS Networks. In *Proc. of the IEEE International Conference on Third Generation Wireless and Beyond (3Gwireless '01)*. 2, 2.3.1

[18] Stuckmann, P. and Hoymann, C. (2002). Performance Evaluation of WAP-based Applications over GPRS. In *Proceedings of the IEEE International Conference on Communications (ICC '02)*, New York, USA. 2.1.3

[19] Stuckmann, P. and Paul, O. (2001). Dimensioning GSM/GPRS Networks for Circuit- and Packet-Switched Services. In *Proceedings of the 10th Symposium on Wireless Personal Multimedia Communications, ISBN 87-988568-0-4*, pages 597–602, Aalborg, Denmark. 2

[20] Walke, B. (2001). *Mobile Radio Networks - Networking, Protocols and Traffic Performance, 2nd Edition.* John Wiley & Sons, Chichester. 2, 3.2.2

[21] Willinger, W., Taqqu, M., and Sherman, R. (1997). Self-similarity through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level. *IEEE/ACM Transactions on Networking*, 5:71–86. 2.1.4

[22] Wireless Application Protocol Forum (1999). Wireless Application Protocol - Wireless Application Environment Overview. Technical report. 2, 2.1.3

[23] Wireless Application Protocol Forum (2001). Wireless Application Protocol 2.0 - Wireless Application Protocol Architecture Specification. Technical report. 2, 2.1.3

[24] Zaddach, M. and Heidtmann, K. (2001). Measurement and Traffic Characterization of H.26x-coded Video Streams. In *Proc. of the 11th GI/ITG Conference on Measuring, Modelling and Evaluation of Computer and Communication Systems.* VDE-Verlag, Berlin. 2.1.4

[25] ETSI BRAN, High Performance Radio Local Area Network - Type 2, System Overview, TR 101 683, Sophia Antipolis, France, Oct. 1999.

[26] IEEE, High Speed Physical Layer in the 5 GHz band, Draft Supplement to Standard IEEE 802.11, New York, Jan. 1999. 3.2.1

[27] ETSI-RES10, High Performance Radio Local Area Network (HIPERLAN), Type 1, Functional specification, Standard EN 300 652, Sophia Antipolis, France, Jul. 1998.

[28] IEEE, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Standard, IEEE 802.11, New York, Nov. 1997. 3.2.1

[29] B. Walke, S. Böhmer, and M. Lott, Protocols for a Wireless ATM Multihop Network, in *Proc. 1998 International Zurich Seminar on Broadband Comm.* pp. 75-82, ETH Zürich, Switzerland, Feb. 1998. 3.2.2

[30] B. Xu, B. Walke, Protocols and Algorithms supporting QoS in an Ad-hoc Wireless ATM Multihop Network, in *Proc. EPMCC'99*, pp. 79-84, Paris, France, Mar. 1999. 3.2.2, 3.2.2

[31] B. Xu, B. Walke, A Multiple Access Protocol for Ad-hoc Wireless ATM Multihop Networks, in *Proc. IEEE VTC'99*, pp. 1141-1145, Houston, U.S.A., May. 1999. 3.2.2, 3.2.2

[32] F. A. Tobagi and L. Kleinrock, Packet Switching in Radio Channels: Part II - The Hidden Terminal Problem in Carrier-Sense Multiple Access and the Busy-Tone Solution, *IEEE Trans. on Comm.*, vol.23, no.12, pp. 1417-1433, Dec. 1975. 3.2.1, 1

[33] J. Khun-Jush, P. Schramm, U. Wachsmann, and F. Wenger, Structure and Performance of the HiperLAN/2 Physical Layer. In *Proc. VTC'99 Spring*, pp. 1094-1100, Holland, April 1999. 3.4

[34] B. Xu, B. Walke, W-CHAMB: A Wireless Channel-oriented Ad-hoc Multihop Broadband Network - Comparison with IEEE 802.11. In *Proc. European Wireless'99*, Munich, Germany, October 1999. pp. 79-84 3.2.3, 3.3