Traffic Engineering for the Evolution of GPRS/EDGE Networks

Christian Hoymann, Peter Stuckmann

Communication Networks
Aachen University of Technology
{hoy|pst}@comnets.rwth-aachen.de, http://www.comnets.rwth-aachen.de

Abstract

In this paper radio network dimensioning for the evolution of GSM/EDGE networks towards 3rd Generation (3G) mobile radio networks is presented. Taking simulation results for GPRS/EGPRS introduction and evolution scenarios as the basis, the radio resources needed for a given offered traffic and for given Quality of Service (QoS) requirements can be estimated. Both circuit-switched voice traffic and packet-switched Internet traffic are considered sharing the radio resources available in a GSM radio cell. To achieve this, the simulation tool GPRSim is used that comprises a prototypical implementation of the standardized GPRS protocols and stochastic load generators for typical GPRS usage.

Keywords

GPRS, EGPRS, Internet applications, WAP, video streaming, traffic engineering , stochastic simulation.

1. Introduction

In the context of the evolution towards 3rd Generation (3G) mobile radio networks, packet-switched data services like the General Packet Radio Service (GPRS) are already introduced or services like the Enhanced GPRS (EGPRS) are going to be introduced into GSM and IS-136 systems worldwide. For network operators, equipment vendors, and system integrators dimensioning rules have to be developed to plan and estimate the radio capacity that is needed for the predicted amount of user data, when the radio resources are shared between circuit- and packet-switched services. Many of these emerging data services will handle multimedia data. As streaming is strongly associated with multimedia content the providers have to take care that streaming is working well inside their mobile environment.

For circuit-switched networks the Erlang-B-Formula has been successfully applied over decades, while for packet-switched networks such an applicable capacity model is still missing. The analytical description of statistical multiplexing and Internet traffic modeling is more complex than for circuit-switched networks. Several papers concerning GPRS performance analysis were published in the last years [1]. They do not contain results for on-demand channel configurations with coexisting circuit-switched traffic sources, which are the typical configurations in GPRS introduction scenarios. Publications about on-demand channels like [2] do not focus on higher layer traffic performance and do not consider the complete Internet stack with TCP. Dimensioning rules for GPRS focussed on traditional WWW and e-mail traffic were first published in [3].

In Section 2 the problem of GPRS radio network dimensioning is introduced. After the description of the simulation environment GPRSim in Section 3, (E)GPRS dimensioning rules are presented for fixed and on-demand channel configurations in Section 4.

2. Radio Network Dimensioning

A suitable dimensioning approach for the busy hour is based on the number of active users, the corresponding applications and the user behavior. These parameters characterize the offered traffic that has to be served by the network. Second the quality of service, which the operator wants to offer his customers has to be defined. Giving these two dimensioning criteria as the input parameters to an adequate capacity model the needed radio capacity can be determined.

Although analytical and algorithmic models for the performance analysis of packet-switched radio networks are under development, the full detail of the GPRS protocol architecture and the Internet protocols including TCP can currently not be described simply by formulas or equations usable in practice. Producing performance results by measurements in the existing GPRS networks is not easily possible, since scenarios with well-defined traffic load are hard to set-up, the calculation of performance and system measures are limited, and the analysis of different alternative protocol implementations is not possible with the existing network equipment. Therefore computer simulation with prototypical implementation of the standardized GPRS protocols and the Internet protocols, traffic generators for the regarded applications and a simple model for the radio channel is chosen as the methodology to get the needed results rapidly.

3. Simulation Environment

The (E)GPRS Simulator GPRSim, which has been developed at the Chair of Communication Networks models a GSM/GPRS network with its protocol architecture, the radio channel attributes and protocol specific traffic sources. Based on its simulation results it is possible to create dimensioning graphs so that network dimensioning can be performed.

The simulator GPRSim [4] is a pure software solution based on the programming language C++. Up to now models of Mobile Station (MS), Base Station (BS), and Serving GPRS Support Node (SGSN) are implemented (see Figure 1). For the implementation of the simulation model in C++ the Communication Networks Class Library (CNCL) is used that is a predecessor to the SDL Performance Evaluation Tool Class Library (SPEETCL). This allows an object oriented structure of programs and is especially applicable for event driven simulations. The approach of the GPRSim is based on the detailed implementation of the standardized protocols. This enables a realistic study of the behavior of (E)GPRS. The real protocol stacks

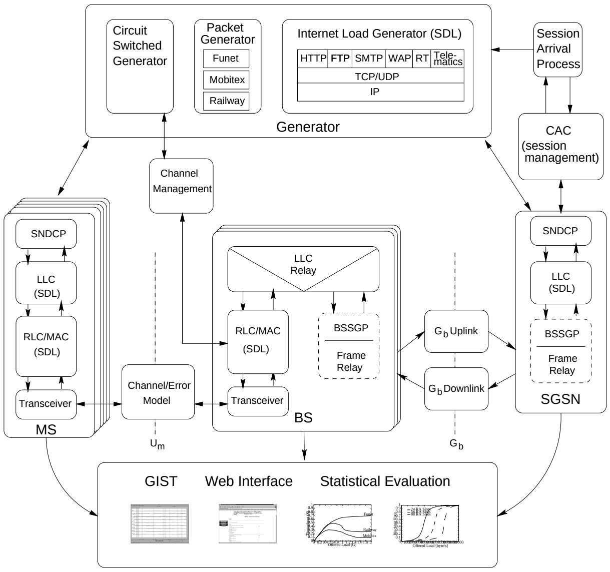

Figure 1: The GPRS/EGPRS Simulator GPRSim

are used during system simulation and statistically analyzed under a well-defined traffic load. The complex protocols like LLC, RLC/MAC based on GPRS/EGPRS release 99, the Internet traffic load generators and TCP/IP are specified formally with the Specification and Description Language (SDL) and are translated to C++ by means of the Code Generator SDL2CNCL and are finally integrated into the simulator.

3.1. Packet Traffic Generators

Internet sessions consist of the applications World Wide Web (WWW) and e-mail running the TCP/IP protocol stack and on the UDP/IP based applications WAP and video streaming. Therefore Internet traffic models are necessary for simulative examinations of the performance of data services of mobile radio networks. In the following, model parameters of these applications and their distributions for generating protocol specific traffic are presented. Related documentation can be found in [5] and [6]. The parameters of these models are updated by parameters given by ETSI/3GPP suppositions for the behavior of mobile Internet users (see Table 1).

3.1.1. WWW Model

WWW sessions consist of requests for a number of *pages*. These pages consist of a number of *objects* with a dedicated *object size*. Another characteristic parameter is the delay between two pages depending on the user's behavior to surf around the Web [5]. Table 1 gives an overview of the WWW 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.

3.1.2. E-mail Model

The e-mail model describes the traffic arising with the transfer of a message downloaded from a mail server by an electronic mail user. The only parameter is the amount of data per e-mail. A constant base quota of 300 byte is added to this size [6]. The parameters of this distribution are shown in Table 1. An overall mean value of 10000 byte for the e-mail size is chosen, since it is assumed that no e-mails with large attachments will be downloaded on mobile devices.

3.1.3. Wireless Application Protocol (WAP) Model

A WAP traffic model was developed and applied in [7]. The main characteristics of the model are very small packet sizes (511 byte) approximately following a log2-normal distribution and a limited value of the packet size (1400 byte).

3.1.4. Video Streaming Model

The video streaming traffic model is based on three video sequences proposed by the Video Quality Expert Group (VQEG). Each sequence is representing a particular group of videos with different intensities of motion. To model evolution and migration scenarios a conventional mixture of sequences including 80 % `Claire` (low motion intensity), 10 % `Carphone` (periods with rather high motion and periods of low motion intensity) and 10 % `Foreman` (permanently high motion intensity) has been selected. They are randomly concatenated to one single stream with low motion infrequently interrupted by phases of higher motion intensity. The resulting average IP traffic of this particular stream is $14.39\,\mathrm{kbit/s}$. 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. Thus, the duration of video sessions is modeled by a negative exponential distribution with an average value of 60 s.

3.2. Circuit-switched Traffic Generator

The circuit-switched traffic generator generates events with an inter-arrival time determined by a negative-exponential distribution. These events correspond to calls initiated in a cell. The traffic value in Erlang is given by the two configurable mean values of the call inter-arrival and the call duration times.

4. Dimensioning Rules for GPRS/EGPRS

In GSM networks a radio cell may allocate resources on one or several physical channels in order to support (E)GPRS data traffic. Those channels shared by the (E)GPRS mobile stations are taken from the common pool of GSM physical channels available in the radio cell. The operator can decide to dedicate permanently or temporarily physical channels for (E)GPRS data traffic. In this context GSM physical channels allocated permanently for data services are called *fixed Packet Data Channels* (PDCHs), channels allocated temporarily are called *on-demand* PDCHs. The dynamic allocation of physical channels to circuit-switched services and packet-oriented data traffic is done according to the "capacity on demand" principle. Simulation results in [8] have shown that the performance of today's GPRS-based services (WAP, WWW, e-mail) does not increase dramatically, when few fixed instead of on-demand PDCHs are used. If the cell capacity is dimensioned with a low blocking probability, e.g., 1 %, for voice services, the use of on-demand PDCHs makes sense. On the other hand operators might allocate fixed PDCHs. With the evolution towards EGPRS sufficient bitrates for real-time or conversational applications are available. These services do need a certain QoS which can only be guaranteed by fixed PDCHs. According to this, two different scenarios have been simulated and the corresponding dimensioning rules for on-demand and fixed PDCH configurations have been applied.

Table 1: 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 (transf.: 9.4)	$4.67 \cdot 10^9$ (transf.: 5.2)
e-mail Parameter	**Distribution**	**Mean**	**Variance**
e-mail size (lower 80 %) [byte]	$\log_2$-normal	1700 (transf.: 10.0)	$5.5 \cdot 10^6$ (transf.: 2.13)
e-mail size (upper 20 %) [byte]	$\log_2$-normal	16000 (transf.: 9.5)	$71.3 \cdot 10^9$ (transf.: 12.8)
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 (transf.: 6.34)	$4.1 \cdot 10^3$ (transf.: 0.71)
Size of ‘Content’ packet [byte]	$\log_2$-normal	511.0 (transf.: 8.6)	$3.63 \cdot 10^5$ (transf.: 1.55)

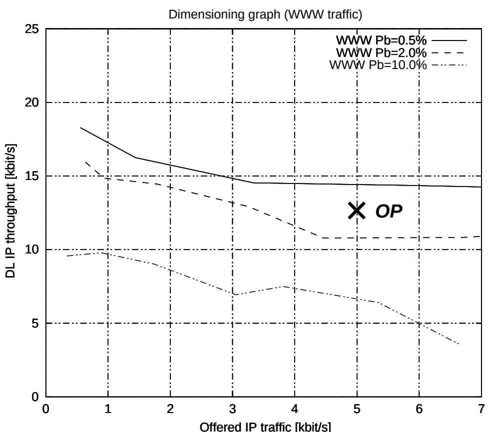

Figure 2: downlink IP throughput per user (WWW traffic)

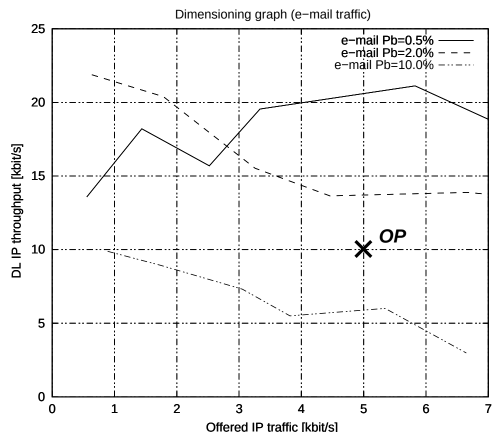

Figure 3: downlink IP throughput per user (e-mail traffic)

4.1. GPRS Configuration with on-demand PDCHs

A traffic mix containing WAP (60%), e-mail (27%) and WWW (13%) sessions, which is typical for today, is carried by a GPRS system with up to 8 on-demand PDCHs. As shown in [3] these networks can be dimensioned with the parameter *blocking probability* (P_b). It represents the probability that an incoming speech call can not be accepted by the network due to congestion. Depending on the predicted number of users and the desired system performance, the dimensioning rule indicates, whether the offered speech traffic corresponding to the P_b value is acceptable for the GPRS traffic or not. If the current P_b is not acceptable, another transmitter-receiver unit (TRX) has to be added to extend cell capacity. To do so, five general steps are necessary:

Step 1: estimate P_b for the related TRX scenario

Step 2: predict the average number of expected users per cell

Step 3: calculate the total offered IP traffic per cell

Step 4: define the desired GPRS QoS parameters the operator wants to guarantee

Step 5: assign the *operating point* (OP) to the dimensioning graphs and choose the curve, representing the acceptable P_b

These five steps will now be applied as an example to the particular GPRS scenario. Therefore a P_b of 2% is assumed. An average number of 10 users will generate $225^{kbyte}/_h$ each. This leads to a total offered IP traffic of $5^{kbit}/_s$. A downlink IP throughput per user for WWW applications of 12.5$^{kbit}/_s$ and for e-mail services of 10$^{kbit}/_s$ is assumed. The objective for WAP application response times is set to 2 s. Figures 2 - 4 show the QoS measures produced by the GPRSim for the applications WAP, WWW and e-mail in one radio cell over the total offered IP traffic in the regarded cell. Regarding WWW applications in Figure 2 the OP is lying between the curves for a P_b of 2% and P_b of 0.5%. For the parameter downlink IP throughput per user the curve above indicates the required level. So the position of the OP indicates that the desired QoS can not be provided in a cell dimensioned for a P_b of 2%. Referring to e-mail applications in Figure 3 the OP rests far below the curve for P_b=2%. Hence a P_b of 2% is more than sufficient to provide a downlink IP throughput for e-mail applications of 10$^{kbit}/_s$ within the particular load situation. The OP for WAP services in Figure 4 is located above the curve. Since a shorter application response time is favorable this indicates a satisfying situation. This result does not only outline that another TRX should be added to the particular cell, to be able to provide the specified downlink IP throughput per user of 12.5$^{kbit}/_s$ for WWW applications. It also shows the order of resource sensitivity. E-mail services do not

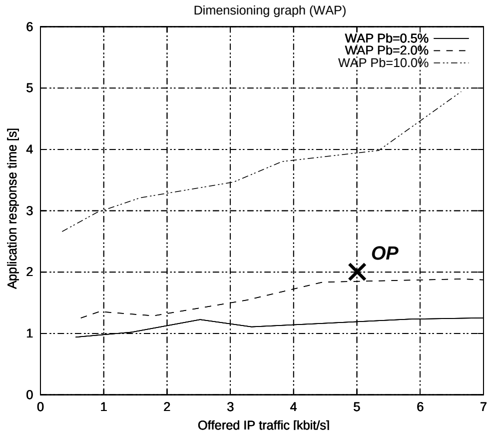

Figure 4: Application response time (WAP traffic)

require much cell capacity. WAP applications do require more resources to reach their goal of QoS while WWW applications represent the bottleneck concerning this particular combination of load situation and QoS requirements.

4.2. EGPRS Configuration with fixed PDCHs

In the next scenario, EGPRS as the enhanced system will be bearer for a future traffic mix. Since evolution scenarios shall be investigated, the regarded traffic mix contains 10% real-time services, while e-mail applications comprise 63% and the remaining part of 27% is WWW. Today's WAP applications will emerge to TCP-based applications, following the WAP 2.0 specification and will show a behavior similar to WWW. To ensure a guaranteed QoS of EGPRS-based applications, GSM physical channels are assigned as fixed PDCHs utilized by the GPRS applications. Therefore the number of fixed PDCHs must be engineered. Thus a reliable relation between the predefined QoS parameter and the offered IP traffic per PDCH is needed. To normalize the offered IP traffic in the graphs to the number of fixed PDCHs, the offered IP traffic has to be divided by the number of fixed PDCHs which has been allocated by EGPRS in the corresponding scenario. The resulting dimensioning graphs are then utilized to apply the four dimensioning rules for fixed PDCH configurations:

Step 1: define the desired QoS parameters the operator wants to guarantee

Step 2: predict the average number of expected users per cell

Step 3: calculate the total offered IP traffic per cell

Step 4: determine the acceptable IP traffic per PDCH by the aid of the dimensioning graphs and calculate the required number of fixed PDCHs:

$$\text{fixed PDCHs} = \frac{\text{total offered IP traffic}}{\text{acceptable traffic per PDCH}}$$

Now these four steps for fixed PDCH configurations are applied to the particular EGPRS scenario. The parameter downlink IP throughput per user is defined as $12.5\,{}^{\text{kbit}}/_{\text{s}}$ for WWW applications, as $10\,{}^{\text{kbit}}/_{\text{s}}$ for e-mail services and as $14.5\,{}^{\text{kbit}}/_{\text{s}}$ for video streaming. The objective for the average downlink IP datagram delay of video streaming applications can be alternatively set to 2 s. In this example, an average number of 10 EGPRS users will generate an increased amount of IP traffic of $2.16\,{}^{Mbyte}/_{h}$ each. This leads to a total offered IP traffic of $48\,{}^{kbit}/_{s}$ per cell.

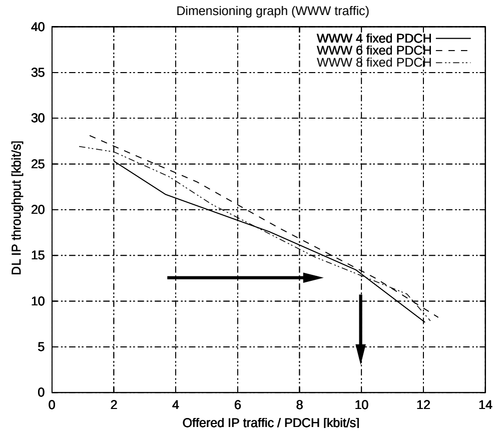

Figure 5: Dimensioning graph (WWW traffic)

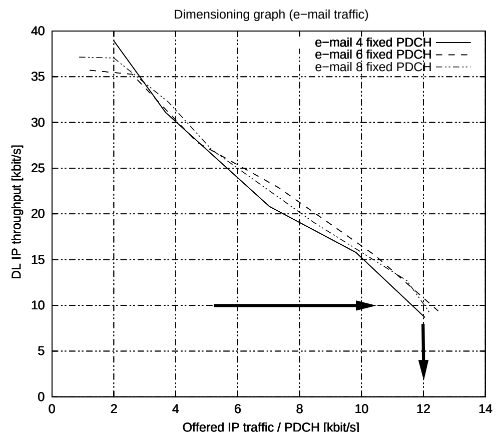

Figure 6: Dimensioning graph (SMTP traffic)

By normalizing the Figures 5 - 8, the different curves are reduced to a single dimensioning graph. In Figure 5 the predefined QoS parameter for WWW applications allows a limit of $10\,{}^{kbit/s}/_{PDCH}$. This results in:

$$({}^{48kbit}/_{s})\Big/\left(10\,\frac{kbit/s}{PDCH}\right) = 4.8$$

Since an integer value of PDCHs has to be found, 5 fixed PDCHs would be needed. Thus 5 fixed PDCHs were sufficient to handle the traffic load, so that for WWW applications an average downlink IP throughput per user of $12.5\,{}^{\text{kbit}}/_{\text{s}}$ can be provided.

Figure 6 indicates that for e-mail applications the traffic limit lies at $12\,{}^{kbit/s}/_{PDCH}$. This results in exactly 4 required fixed PDCHs. Thus, restrictions concerning e-mail traffic are less stringent. The remaining part of the traffic mix is video streaming. Figure 7 shows the downlink IP throughput per user as the decisive QoS value. The required data rate for the video stream chosen for this simulation is $14.5\,{}^{\text{kbit}}/_{\text{s}}$. It is also the predefined goal of the dimensioning rule. It can now be observed that only a traffic value of $9\,{}^{kbit/s}/_{PDCH}$ is allowed. Dividing the calculated total IP traffic per cell by this value, a number of 5.3 remains. This implies that 6 fixed PDCHs must be reserved for EGPRS to fulfill the performance goal set for streaming applications. In Figure 8 the downlink IP datagram delay is alternatively used for real-time services. Applying the dimensioning rules to these values, a traffic limit of $8\,{}^{kbit/s}/_{PDCH}$ results. This leads to exactly 6 required fixed PDCHs if a mean delay of

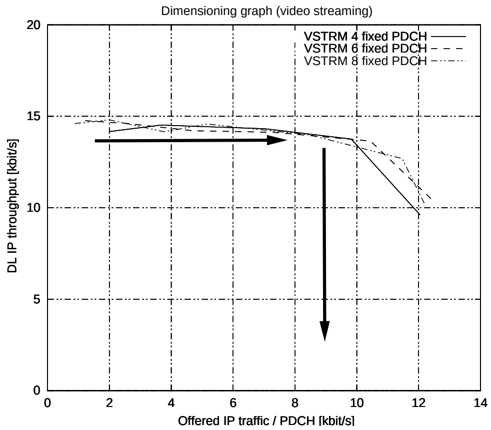

Figure 7: Dimensioning graph (streaming throughput)

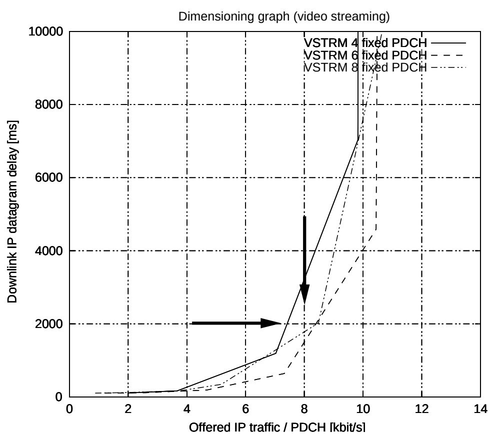

Figure 8: Dimensioning graph (streaming delay)

2 seconds is aimed at. This is a more stringent restriction than the limitation concerning the throughput performance criterion as calculated above.

Thus, in this planning scenario, real-time data services are most sensitive with regards to resource requirements. The next application type following are the WWW services with its kind of interactive character in terms of feedback from several web pages during one single session. Last e-mail services, which do neither require high throughput performance nor short IP datagram delays, are following. Focusing on streaming services, requirements resulting from the downlink IP datagram delay as the decisive QoS parameter are more stringent than the one resulting from the downlink IP throughput per user. This indicates that the delay is more sensitive concerning resource requirements than the pure throughput is.

The characteristics of the different applications can be supported by QoS management [1]. The transmission of streaming data may be privileged on the expense of,e.g., e-mail traffic. On the one hand the curves in the e-mail dimensioning graph are shifted to the left so that less traffic is allowed per PDCH to meet the e-mail requirements. But on the other hand the curves for video streaming are shifted to the right, so that more traffic per PDCH is allowed and the higher streaming requirements are still met. Since the most stringent application specifies the needed number of fixed PDCHs, the increased requirements for e-mail do not have any effect up to a certain limit while the decreasing demands of video streaming results in a smaller number of fixed PDCHs to be allocated for EGPRS. This may lead to less TRXs needed to serve one cell and therewith in lower cost for the radio equipment.

5. CONCLUSIONS

In this paper dimensioning rules for today's GPRS networks carrying WAP dominated data traffic as well as for emerging EGPRS networks being bearer for multimedia driven data applications are presented. Regarding GPRS on-demand PDCH configurations dimensioning graphs based on the simulation results for different speech traffic load scenarios can be used to determine whether a new TRX has to be installed to serve the GPRS traffic with a desired QoS or not. The dimensioning graphs for EGPRS configurations with fixed PDCHs are based on simulation results for different load scenarios and profits by nearly linear correlation between downlink IP throughput and the offered IP traffic in the radio cell. With these dimensioning graphs the relationship between desired QoS, offered traffic and needed radio capacity can be estimated. QoS management may be one possibility to overcome the restrictions. With the privileged data transmission of the most resource sensitive application, the requirements concerning radio network equipment can be reduced, or vice versa the performance can be increased.

6. References

[1] P. Stuckmann, "Quality of Service Management in GPRS-based Radio Access Networks," *Telecommunication Systems 19:3, Kluwer Academic Publishers*, pp. 515–546, 2002.

[2] S. Ni and S. Häggmann, "GPRS Performance Estimation in a GSM Voice and GPRS Shared Resource System," in *Proc. of IEEE Wireless Communication and Networking Conference (WCNC '99)*, pp. 1417–1421, 1999.

[3] P. Stuckmann and O. Paul, "Dimensioning GSM/GPRS Networks for Circuit- and Packet-Switched Services," in *Proceedings of the 10th Symposium on Wireless Personal Multimedia Communications*, (Aalborg, Denmark), pp. 597–602, September 2001.

[4] P. Stuckmann, "Simulation Environment GPRSim: Tool for Performance Analysis, Capacity Planning and QoS Enhancement in GPRS/EDGE Networks," Technical Report, `http://www.comnets.rwth-aachen.de/~pst`.

[5] M. Arlitt and C. Williamson, "A synthetic workload model for internet mosaic traffic," in *Proceedings of the 1995 Summer Computer Simulation Conference*, (Ottawa, Canada), pp. 24–26, July 1995.

[6] V. Paxson, "Empirically-derived analytic models of wide-area tcp connections," *IEEE/ACM Transactions on Networking*, vol. 2 (4), pp. 316–336, Aug. 1994.

[7] P. Stuckmann, H. Finck, and T. Bahls, "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), 2001.

[8] P. Stuckmann and F. Müller, "GPRS Radio Network Capacity Considering Coexisting Circuit-Switched Traffic Sources," in *European Conference on Wireless Technology (ECWT 2000) Conference Proceedings*, (Paris, France), Oct. 2000.