Paper submitted to

International Conference on Wireless Networks

June 24 – 27, 2002, Las Vegas, Nevada, USA

Authors: Silke Heier, Andreas Kemper, Jan-Oliver Rock

Title: **Performance Characteristics of UMTS for the Mobile Internet Access**

Keywords: UMTS, UTRA, Mobile Radio, Internet, Performance

Address: Aachen University of Technology (RWTH)
Communication Networks (CN)
Kopernikusstr. 16
52074 Aachen, Germany

Tel.: +49-241-80-28576

Fax.: +49-241-88-90380

E-Mail: {she|kem}@comnets.rwth-aachen.de

WWW: http://www.comnets.rwth-aachen.de/˜she

Performance Characteristics of UMTS for the Mobile Internet Access

Silke Heier, Andreas Kemper, Jan-Oliver Rock

Communication Networks
Aachen University of Technology
Kopernikusstr. 16, 52074 Aachen, Germany
E-mail: {she|kem}@comnets.rwth-aachen.de

Universal Mobile Telecommunication System (UMTS) as a 3rd generation mobile telecommunications system is designed to provide mobile Internet access to the user. In this paper, the *World Wide Web* (WWW) browsing session as one mobile data service application is examined for performance analysis by simulations. A WWW traffic model is presented as the basis of a WWW load generator. This traffic load generator integrated into an event driven, bit accurate *UMTS Radio Interface Simulator* (URIS) is presented. The simulator is based on a detailed implementation of the standardized UMTS protocol stack extended by the *Transmission Control Protocol* (TCP) and the *Internet Protocol* (IP). Simulation results of quality of service parameters like delay and throughput depict the performance of mobile Internet access over UMTS. Of special concern is the interaction of TCP retransmissions with the *Radio Link Control* (RLC) *Automatic Repeat Request* (ARQ) mechanisms.

1 Introduction

In the context of exponential growth of the Internet market, growing demand of data services is also expected for mobile users. Third generation mobile radio networks like *Universal Mobile Telecommunications System* UMTS will provide data services with higher data rates than presently available in second generation cellular systems. UMTS will support bit rates of up to 144 kbit/s in rural areas, 384 kbit/s in hot-spots and up to 2 Mbit/s in indoor scenarios. Hence, the mobile Internet access will be one of the key services which should guarantee a successful start of UMTS operation.

This paper will investigate the performance of *World Wide Web* (WWW) browsing sessions over UMTS from the mobile user point of view. Due to the nature of radio access systems different conditions namely limited bandwidth, higher delays and error rates at the radio interface have to be taken into account compared to fixed network access. Thus, a modeling of the WWW traffic as well as the UMTS radio interface is mandatory for performance evaluation purposes.

First, WWW browsing session traffic characteristics are outlined and mapped with identified parameter sets on a WWW traffic model. The traffic model is the basis for a WWW load generator which was developed by [3]. Following, the UMTS radio interface simulation environment called *UMTS Radio Interface Simulator* (URIS) is presented. This event driven, bit accurate simulator models the UMTS air interface protocol stack in accordance with the UMTS protocol specifications of Release 4. The performance evaluation considers the detailed UMTS protocol implementation as well as the traffic load generation for WWW browsing sessions including *Transmission Control Protocol*/*Internet Protocol* (TCP/IP) modeling. Thus, it is feasible to present the performance of a TCP application in an unreliable mobile environment with higher error rates concerning its packet retransmission interaction on TCP and on radio link control level. Simulation results will show which performance in terms of quality of service parameters a mobile user will experience while surfing the Internet.

Following this introduction, Sec. 2 gives a description of the WWW browsing session and the used protocols. Furthermore the WWW traffic analysis and the traffic modeling are described here. Sec. 3 presents the used simulation methodology and the simulation environment. Sec. 4 and Sec. 5 illustrate the modeling of the TCP/IP and *Radio Link Control* (RLC) protocols. Performance evaluation, simulation results and discussion is done in Sec. 6 and Sec. 7. The paper concludes in Sec. 8 with a summary.

2 WWW Traffic Model

Surfing the Internet is a WWW browsing session which is running on the TCP/IP protocol stack. A WWW traffic model is necessary for simulative examinations of the performance of data services of mobile radio networks. In the following, model parameters of a WWW application and its mathematical description for generating protocol specific traffic are presented.

Related documentation can be found in [2, 5, 10]. The main part of the later implementation is based on the work of [2, 3].

A typical WWW browsing session consists of a sequence of page requests. These pages contain a number of objects with a dedicated object size each. During a page request, several packets for each object may be generated which means that the page request constitutes of a bursty sequence of packets. The bursty-ness during the page request is a characteristic feature of packet transmission. After the page has entirely arrived at the terminal, the user is consuming a certain amount of time for studying the information. This time interval is called reading time. Hence, the following must be modeled:

- Session arrival process: The arrival of session set-ups to the network is modeled as a Poisson process. It is important to note that this process only generates the time instants when a session request begins and is independent of the session termination,
- Session length is modeled implicitly by the number of page requests during the session,
- Number of page requests per session,
- Reading time between page requests: The reading time starts when the last packet of the page request was completely received by the user. The reading time ends when the user makes a request for the next page,
- Number of objects within a page request,
- Inter arrival time between objects,
- Size of an object.

Tab. 1 gives an overview of the distributions and their parameters describing the WWW traffic model.

3 UMTS Radio Interface Model

For this study, the simulation tool URIS was developed. This simulation environment is used to investigate, optimize and develop features of the radio interface protocol stack. In addition, it offers the opportunity of capacity and quality of service evaluation by simulations of various scenarios. The simulator is a pure software solution in the programming language C++. The simulation model is implemented with the help of a powerful C++ class library which was developed by the Chair of Communication Networks and is called the *SDL Performance Evaluation Tool Class Library* (SPEETCL) [12]. This generic, object oriented library is well suited for telecommunication network simulation purposes and can be used in event driven, bit accurate simulation environments. The UMTS protocols at the radio interface enhanced by a TCP/IP protocol stack were specified with the *Specification and Description Language* (SDL). To generate an executable out of the SDL phrase notation and the C++ library, a SDL2SPEETCL code generator is used.

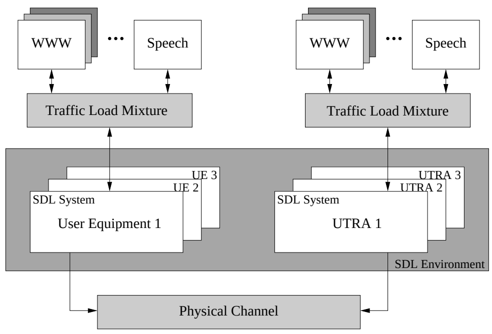


Figure 1: Structure of the URIS Simulator

The software architecture of the URIS simulator is shown in Fig. 1. Up to now, the simulator consists of various traffic generators and a traffic load mixture unit which is used to adjust scenarios with desired load mixtures. The physical channel module models transmission of bursts in radio frames

WWW Parameter	Distribution	Mean	Variance
Session Arrival Rate $[h^{-1}]$	negative exponential	2.0	—
Pages per Session	geometric	20.0	—
Reading Time between Pages [s]	negative exponential	33.0	—
Objects per Page	geometric	2.5	—
Inter Arrival Time between Objects [s]	negative exponential	0.5	—
Page Request Size [byte]	normal	1136.0	80.0
Object Size [byte]	$\log_2$-Erlang-k	$\log_2 2521 \approx 11.3$	$(\log_2 5)^2 = 5.4$

Table 1: Model Parameters of WWW Browsing Session

on the radio interface. This includes discarding of erroneous bursts depending on the error model.

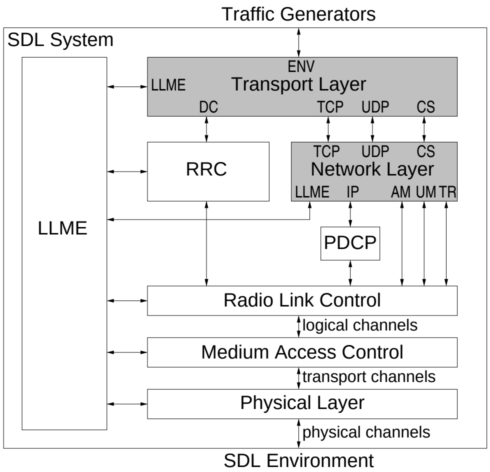


Figure 2: URIS SDL System Structure

The core of the simulator are the modules *User Equipment* (UE) and *UMTS Terrestrial Radio Access* (UTRA) which are built formally similar. Each single UE and UTRA is represented in URIS as an SDL system which contains the protocol implementation of the layers. Fig. 2 gives an overview of the protocol structure in a generic SDL system.

The complex protocols like *Medium Access Control* (MAC), RLC, *Packet Data Convergence Protocol* (PDCP) and *Radio Resource Control* (RRC) based on UMTS Release 4 and TCP, *User Datagram Protocol* (UDP) and IP are specified in SDL. Each layer is represented by an appropriate SDL block which contains the layer specific processes. Usual simulator approaches model protocols and functions on basis of abstractions and simplifications. The aim of URIS is a detailed, bit accurate implementation of the standardized protocols. This offers the opportunity to determine the performance of UMTS in a realistic manner.

Since SDL offers object oriented programming properties, they were used to build a generic protocol stack in conformance of the OSI reference model. Each URIS SDL system follows the inheritance tree shown in Fig. 3.

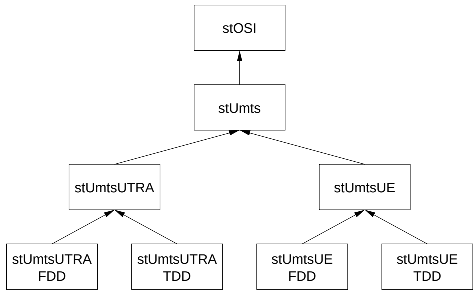


Figure 3: URIS SDL Inheritance Tree

All system types are derived from a base system type called *stOSI*. This base system type includes all methods and functions to build up a protocol stack including the layers. The generic structure of each layer is shown in Fig. 4. The base system type includes manager/entity concepts and *Service Access Points* (SAP) as well as the handling and processing of *Interface Data Units* (IDU), *Protocol Data Units* (PDU) and *Service Data Units* (SDU). The layer specific procedures are mapped on the entities.

Each SDL system inherits the members of the base type, redefines them or adds new ones. The SDL system type *stUmts* includes blocks, processes, procedures and signal definitions that are common to all UMTS radio interface nodes, i.e. UE and UTRA. Functionality which is specific to UE or UTRA is contained within the respective derived system type, e.g. *stUmtsUE* and *stUmtsUTRA*.

The advantage of the object oriented concept is greater flexibility when functionality of the whole system or a subtree is edited. All changes will automatically be available in derived types.

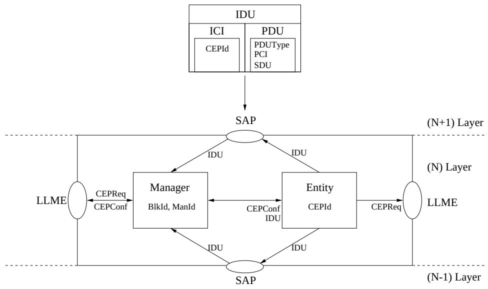


Figure 4: Generic URIS SDL Layer Structure

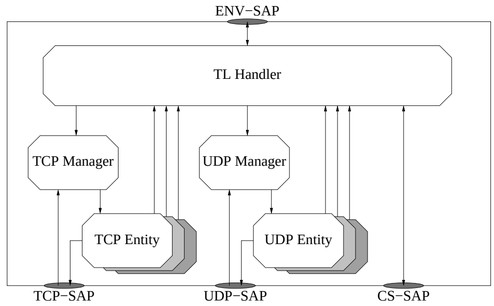


Figure 5: SDL Model of the Transport Layer

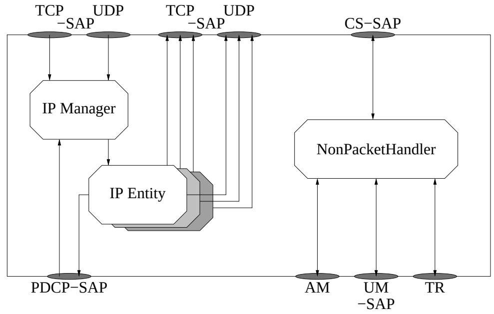


Figure 6: SDL Model of the Network Layer

4 TCP/IP Protocol Suite

The modern Internet is mainly based on TCP and IP. In order to provide Internet services to the mobile user, the TCP/IP protocol suite has to be adapted to the UMTS protocols. The UMTS radio interface protocol stack offers this functionality by providing the PDCP.

TCP provides overlying protocols with a reliable connection. The following services are offered:

- Retransmission of lost packets,
- Recovery from out-of-order delivery,
- Flow control.

The TCP implementation realized in URIS is based on the so called "Reno" TCP stack and uses the following flow control mechanisms:

- Slow start and congestion avoidance,
- Fast retransmit and fast recovery,
- Delayed acknowledgments,
- Selective acknowledgments [4, 9].

The SDL model of the transport layer is shown in Fig. 5. The transport layer offers TCP and UDP and bypasses circuit switched data like speech traffic. Every TCP or UDP session is handled by a dedicated entity.

IP is a connection-less and unreliable protocol that operates on a best-effort basis. This means that IP packets may be lost, out-of-order or even duplicated without IP handling these situations. This has to be done by higher layers. In URIS, the IP protocol implementation currently performs data encapsulation.

The SDL model of the network layer is shown in Fig. 6. The network layer bypasses circuit switched data like speech traffic. Every IP session is handled by a dedicated entity.

5 Radio Link Control Protocol

The RLC realizes segmentation and retransmission services for both user and control data. Fig. 7 shows the structure of the RLC layer as it is laid out in [1].

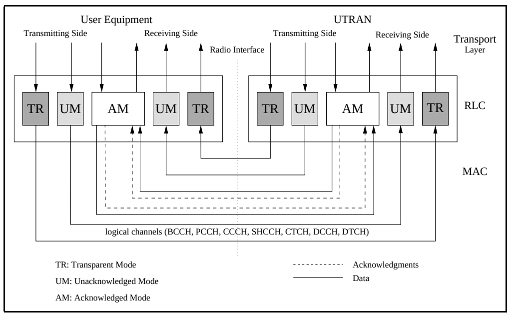


Figure 7: RLC Structure according to 3GPP 25.322

The RLC protocol provides three different data transfer services to the adjacent layer:

- *Transparent data transfer service mode* (TR),
- *Unacknowledged data transfer service mode* (UM),
- *Acknowledged data transfer service mode* (AM).

The TR mode is an unidirectional service typically used for broadcast or paging services, where it is not necessary to guarantee an error-free transmission. A use for transmission of streaming data (e.g. audio or video) is also feasible, especially if real-time transmission is more important than error reliability. A dropping mechanism prevents delivery of already expired PDUs.

The UM is an unidirectional service typically used for streaming applications (streaming class) where it is not necessary to guarantee an error-free transmission. Voice over IP (conversational class) is also a feasible service conveyed by the UM. The UM transmits higher layer data packets without guaranteeing delivery to the peer entity. By using a sequence-number check function in the receiving entity, the UM is capable of detecting missing RLC PDUs, but error recovery is not performed. A dropping mechanism prevents transmission of already expired PDUs.

The AM transmits higher layer data packets with guaranteed delivery to the peer entity. Therefore, it is mainly used for traffic of interactive class or background class type. *Automatic Repeat Request* (ARQ) mechanisms are applied for correction of transmission errors. It is possible for higher layers to request a transmission confirmation from the AM. A Go-Back-N ARQ has been implemented as a first approach for AM. Segmentation and reassembling as well as concatenation are fully implemented as specified in [1].

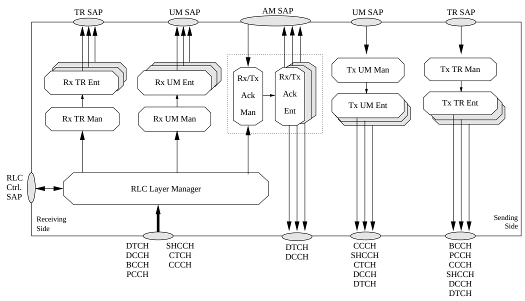


Figure 8: SDL Model of the RLC Layer

6 Simulation Scenarios

The main parameters concerning *Quality of Service* (QoS) which are affected by RLC and TCP/IP, are delay and throughput. Two measurements were made during the simulations:

1. Packet Delay: Time from sending a traffic load packet to the layer until correct reception of the packet by the traffic load receiver,

2. Packet Throughput: Size of the traffic load packet divided by the packet delay. The throughput is given in bytes per second [byte/s].

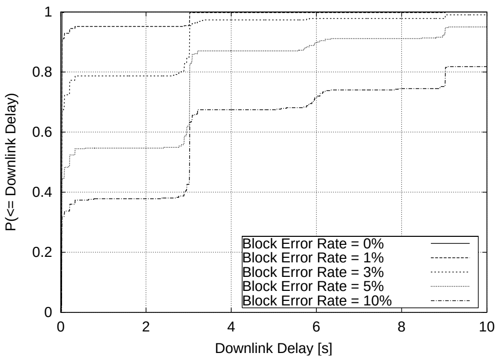
Simulations of a WWW browsing session running on TCP/IP in RLC UM or RLC AM were executed. The cumulative distribution function of the measured packet delay and the complementary cumulative distribution function of the measured packet throughput have been calculated. Measurements have been made for increasing block error rates from 0% up to 10%. For a better understanding of the WWW simulation results, additional simulations with a *Constant Bit Rate* (CBR) were executed. Tab. 2 shows the parameters of RLC and TCP/IP used for these simulations.

7 Simulation Results

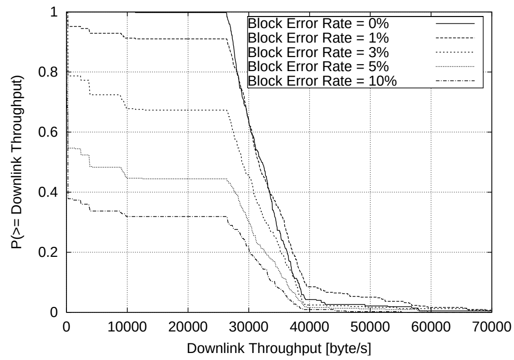
The results of the simulations are given in Fig. 9. As can be seen a minimum delay is always present, caused by the constant transmission delay of the physical layer. Block error rates of 0% represent the ideally achievable throughput. The throughput is not limited by the underlying *Dedicated Channel* (DCH) since the offered load is always less than the channel capacity. For the CBR scenario a minimum of four sessions is always active which results in a source bit rate of 25.6 kbyte/s.

Parameter	Value	Value
Traffic Generator	CBR	VBR
Session Arrival Rate [h^{-1}]	10	2
Session Data Rate [kbyte/s]	6.4	Tbl. 1
TTI Length [s]	0.01	0.04
DCH Capacity [kbyte/s]	600	260
RLC TxWinSize [PDU]	1024	1024
Max. PDU Retransmissions	40	40
Status Prohibit Timer [s]	0.01	0.01
Poll Timer [s]	0.5	1.0
Discard Timer [s]	10	10
Max. TCP Segment [byte]	512	512
Max. TCP Window [kbyte]	16	16
Min. TCP RTO [s]	3	3
Max. TCP RTO [s]	64	64
Block Error Rate [%]	0, 1, 3, 5, 10	

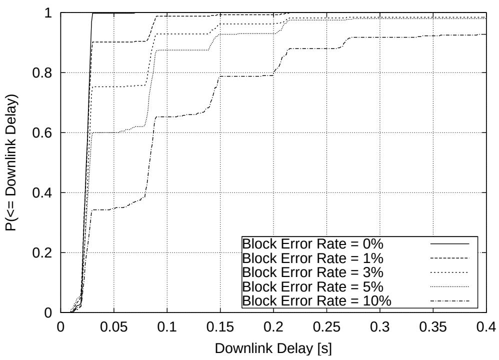
Table 2: Simulation Parameters



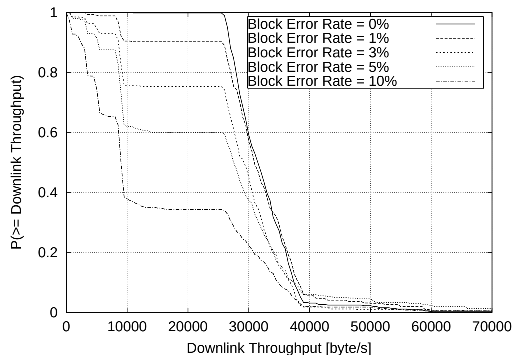
(a) CBR Packet Delay with RLC UM



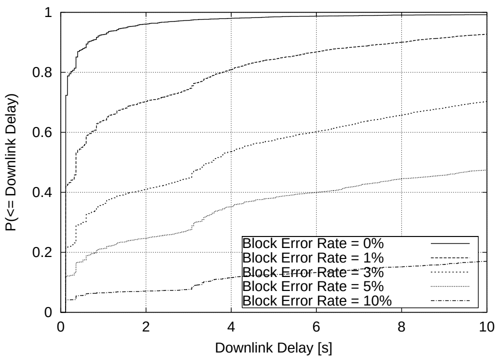
(b) CBR Packet Throughput with RLC UM



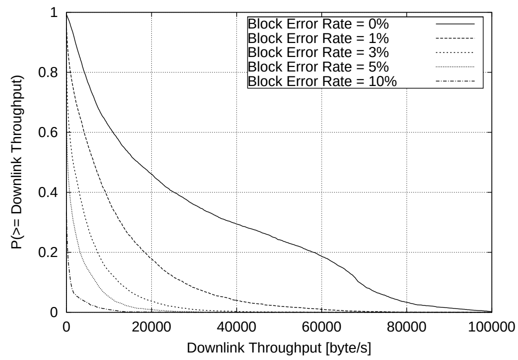
(c) CBR Packet Delay with RLC AM



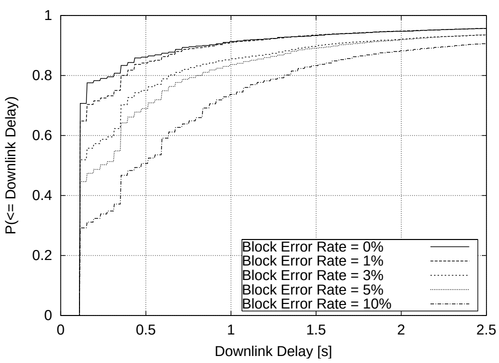
(d) CBR Throughput Delay with RLC AM



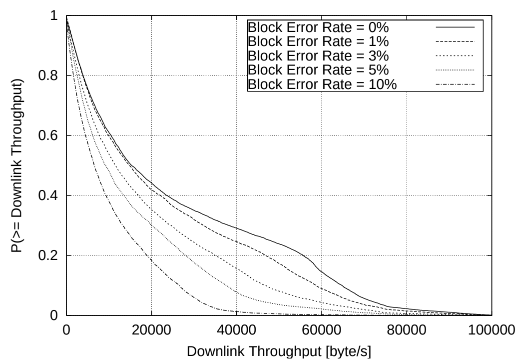
(e) VBR Packet Delay with RLC UM



(f) VBR Packet Throughput with RLC UM



(g) VBR Packet Delay with RLC AM



(h) VBR Packet Throughput with RLC AM

Figure 9: Simulation Results for Traffic Load Packet Delay and Throughput

7.1 Scenario 1: CBR Traffic on RLC UM and RLC AM

The RLC UM has been used to evaluate the performance of TCP's retransmission mechanisms. Fig. 9(a) depicts the TCP delay which shows higher delays with growing packet error rates. It can be noted that the steps of the curves mark the expiration of a TCP *Retransmission Timeout* (RTO). In all simulations a minimum RTO of three seconds has been used. It is obvious, that the retransmission mechanism of TCP is far too slow to handle the heavy packet losses of a radio link. Fig. 9(b) gives the respective throughput which clearly shows heavily reduced throughput for higher error rates. At a block error rate of 10% only 35% of the traffic load packets will achieve a throughput of more than 25.6 kbyte/s which is equivalent to the offered source bit rate. TCP on its own is not able to provide a reliable end-to-end connection in a radio based environment.

The RLC AM has been used to investigate the interaction of the TCP protocol with the RLC ARQ. Fig. 9(c) shows results of the measured TCP delay for several block error rates. The distinct steps of one *Transmission Time Interval* (TTI) length are caused by retransmissions on RLC protocol level. As expected, the delay imposed by the TCP protocol and the underlying layers increases for higher block error rates. At block error rates up to 5%, 90% of the packets have a delay of less than 0.15 seconds. Fig. 9(d) shows the respective throughput. As a result of higher delays lower data throughput is achieved.

Compared to the RLC UM simulation scenario, the ARQ mechanism of the RLC AM is far more quicker and efficient than the TCP retransmission mechanisms. Due to the heavy packet losses of the radio interface, packet data services which are error sensitive should always use the RLC AM.

For a 10% block error rate, the implemented Go-Back-N ARQ mechanism of the RLC AM is no longer able to provide sufficient data throughput. Due to the unnecessary Go-Back-N retransmissions of correctly received packets even more transmission errors are caused and congestion occurs within the RLC protocol.

To avoid unnecessary retransmissions and congestion a Selective-Reject-ARQ (SR-ARQ) mechanism should be used. The SR-ARQ is currently under study.

7.2 Scenario 2: VBR Traffic on RLC UM and RLC AM

The second scenario will present performance results for *Variable Bit Rate* (VBR) sources, e.g. the WWW browsing session. First, results of a WWW browsing session running over RLC UM are discussed. Fig. 9(e) shows the measured TCP delay and Fig. 9(f) the corresponding throughput. Compared to the CBR scenario, delays of traffic load packets are higher. This is caused by the bursty nature of WWW browsing sessions and the flow control of TCP. The window mechanism of TCP realizes a flow control where packets have to wait since the send window is closed. Window waiting times produce additional delays and lower throughput. In case of CBR this effect is not mentioned since the TCP send window will not be closed anytime. Only packet bursts will cause a closed window. The packet throughput is so much reduced that a TCP connection at block error rates of more than 1% is not feasible. Delays and throughput will cause unsatisfied mobile users waiting too long for WWW pages.

The results for WWW browsing sessions running over RLC AM are more promising. At block error rates up to 5%, 90% of TCP packets have a delay of less than 1.5 seconds (see Fig. 9(g)). The decrease of throughput (Fig. 9(h)) is acceptable for mobile users but the guaranteed RLC radio bearer capacity is left unused since the TCP send window is closed due to bursty traffic, missing acknowledgments and retransmission timeouts.

An interesting effect has been observed at high loss rates (10%). Due to the heavy losses on the radio link, the Go-Back-N ARQ mechanism of the RLC must perform many retransmissions. In the mean time, the RTO for transmitted TCP segments expires which leads to a retransmission by the TCP protocol. In this case, delays get even worse since data packets will be sent twice on different protocol layers which increases the traffic unnecessarily. As a result, duplicate TCP packets will arrive at the TCP receiver since original packets are delivered by the RLC ARQ first and second by TCP retransmission. Resulting duplicate acknowledgments are the reason why TCP's fast retransmit algorithm is triggered. By mistake, TCP assumes congestion and reduces the throughput which results in unused but guaranteed RLC radio bearer capacity.

8 Conclusions

This paper presents a simulator concept which can be used to model the protocol stack and service behavior at the UMTS radio interface in a realistic manner. Due to the detailed and standard related implementation, the simulator can be used to evaluate the performance to be expected by real systems. In a second step the simulator can be used to optimize performance or even to develop advanced protocol algorithms.

In addition this paper comprises simulative examinations of the performance of WWW surfing over the UMTS radio interface. The simulation results show that a WWW session using TCP and the RLC UM is not feasible to satisfy a mobile user. If packets are lost the TCP protocol infers that there must be congestion in the network between the two peers. In consequence, TCP retransmission leads to a decrease of the send window size which results in less throughput. This assumption is correct in wired environments but TCP is unable to handle the unreliable radio link of mobile users due to the high error rates and TCP's slow error handling. Very high delays and small throughput lead to unsatisfied mobile users. Improvements may be achieved by using the TCP SACK option which allows the receiver to convey more information about its state to the sender, thus increasing the efficiency of the retransmission strategy.

Running a WWW session over a standard TCP and the RLC AM will be an ordinary scenario during the introduction of UMTS. First UEs will rely on standard application solutions which will be adopted from the fixed world. Hence, TCP will run end-to-end including the radio interface. This paper shows that this solution is applicable to satisfy the mobile user but the performance suffers since TCP mechanisms will not efficiently use the guaranteed QoS of the UMTS radio bearers in terms of delay and throughput. It has to be mentioned that effects of the core network are not included in our simulations. The overall performance will get worse since TCP as a end-to-end protocol will experience additional delays and congestion in the core network. The mobile user will face a reduced throughput and longer waiting times for WWW page downloads than in fixed networks. This relies on the weak TCP performance in a mobile environment and can not be influenced by assignment of radio bearers with higher capacity.

9 Acknowledgment

The authors would like to thank Prof. B. Walke of the Chair of Communication Networks for his support and friendly advice to this work. The contribution of S. Graebner is highly appreciated.

References

[1] 3GPP TS 25.322. *Radio Link Control (RLC) Protocol Specification*. Technical Specification V4.1.0, Release 4, 3rd Generation Partnerschip Project, Technical Specification Group Radio Access Network, jun 2001.

[2] Martin F. Arlitt, Carey L. Williamson. *A Synthetic Workload Model for Internet Mosaic Traffic*. In *Proceedings of the 1995 Summer Computer Simulation Conference*, 1995. Proceedings of the 1995 Summer Computer Simulation Conference, Ottawa, Canada, July, pp. 24–26.

[3] Holger Cornelsen. *Entwicklung und Implementierung von Lastgeneratoren für multimediale Sprach-, Audio- und Paketdatenanwendungen*. Master's thesis, Chair of Communication Networks, RWTH Aachen University of Technology, nov 1996.

[4] K. Fall, S. Floyd. *Simulation-based Comparisons of Tahoe, Reno, and SACK TCP*. Lawrence Berkeley National Laboratory, July 1996.

[5] Victor S. Frost, Benjamin Melamed. *Traffic Modeling For Telecommunications Networks*. IEEE Communications Magazine, March 1994. pp. 70–81.

[6] Sebastian Gräbner. *Simulative Performance Evaluation of UMTS Radio Link Control Protocol*. Master's thesis, Chair of Communication Networks, RWTH Aachen University of Technology, mar 2001.

[7] Harri Holma, Antti Toskala, Hrsg. *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications*. John Wiley & Sons Ltd., mar 2001.

[8] Alexander Lätsch. *Design and implementation of a Simulator for UMTS*. Master's thesis,

Chair of Communication Networks, RWTH Aachen University of Technology, jun 1999.

[9] M. Mathis, J. Mahdavi, S. Floyd, A. Romanow. *Selective Acknowledgement Options*. Rfc, Internet Engeneering Task Force, April 1996.

[10] Vern Paxson. *Empirically-Derived Analytic Models of Wide-Area TCP Connections*. IEEE/ACM Transactions on Networking, 2 (4), pp. 316–336, August 1994; ftp://ftp.ee.lbl.gov/papers/WAN-TCP-models.ps.Z, March 1994.

[11] Jan-Oliver Rock. *Simulative Performance Evaluation of IP based Services over UMTS*. Master's thesis, Chair of Communication Networks, RWTH Aachen University of Technology, may 2001.

[12] Matthias Steppler, Matthias Lott. *SPEET - SDL Performance Evaluation Tool*. 8th SDL Forum '97, September 1997.

[13] B. Walke. *Mobile Radio Networks, 2nd Edition*. John Wiley & Sons Ltd., Chichester, England, 2 edition, 2001.