Measurements of TCP Performance over UMTS Networks in Near-Ideal Conditions

Martin Kohlwes
Followap GmbH
Niederkasseler Lohweg 187
D-40547, Dusseldorf, Germany

Janne Riihijärvi and Petri Mähönen
Department of Wireless Networks
Aachen University (RWTH Aachen)
Kackertstrasse 9, D-52072, Aachen, Germany
Email: jar@mobnets.rwth-aachen.de

***Abstract*— In this paper we report our results from a long series of various measurements on the behavior of TCP over a UMTS wireless channel performed in two different UTMS networks. We conclude that at least in good conditions TCP throughput is close to theoretical maximum, and that RTT is fairly stable. Practically no packet losses were detected, and spurious retransmissions were extremely rare. No performance benefit was observed when TCP retransmission timer modifications, such as the Eifel and FRTO algorithms were used.**

I. Introduction

The third generation cellular systems, such as cdma2000 and UMTS promise to provide much improved support for data communications, compared to the second generation systems (although GSM evolutions, such as EDGE, have considerably bridged the gap). However, there is still only limited understanding, mainly based on simulations on how TCP in particular will perform over a UMTS channel. Studies performed in, for example, GSM and GPRS networks have indicated that in those channels TCP is vulnerable to packet losses and "delay spikes" [1], causing unnecessary retransmissions. Thus it is of interest to try to ascertain, whether similar, TCP hostile conditions are to be expected in the future UMTS networks.

To answer this question we performed a long series of measurements using two already functional UMTS networks in Germany, one in the Alcatel's 3G reality centre, in Stuttgart, and one in Aachen, operated commercially by Vodafone. Of course, while these measurements provide interesting insight on TCP performance over UMTS, they will certainly not be the last word on the issue. Measurements in more realistic conditions (i.e. with larger and dynamic user populations) are definitely needed.

II. Measurement Setup

A. Networks Used

As mentioned in the introduction, the two networks used throughout the measurements were the Alcatel 3G reality centre test network, and the Vodafone UMTS network in Aachen. Both networks operated in the FDD mode, using spreading factor 8 for the downlink, and 16 for the uplink, resulting in the radio bit rates of 384 kbps and 64 kbps, respectively. In the Alcatel network the maximum number of RLC retransmissions was set to 15, while for the Vodafone network the setting was not known.

This work was in part funded by DFG (Deutsche Forschungsgemeinschaft), and RWTH Aachen.

B. User Equipment

Already in, say, wireless LANs we have witnessed that wireless interfaces designed by different manufacturers can exhibit surprisingly different behavior in similar conditions. To see whether this is so for the available UMTS user equipment, measurements were performed using three different handset models, by Samsung, Nokia, and Motorola. In the case of the Nokia user equipment, commercially available, off-the-shelf model was used, while in the cases of Samsung and Motorola, the user equipment was still in the late testing phase.

In all measurements the client was running on a laptop connected to the UE, running Windows XP with standard networking settings, unless noted otherwise. WinDump (the Windows version of tcpdump) was used in all measurements to capture the packet trace of the TCP connections for further analysis.

C. Server Setup

Two types of TCP connections were used in the measurements. First, connections between two UEs were used. While this is an interesting case to study, most of the TCP traffic will probably be related to accessing services in the fixed network. Thus, the bulk of the measurements were performed in a scenario, where a dedicated Linux-based server, running multiple TCP-based services was contacted from a client running on a laptop connected to the UE.

To study the behavior of the TCP state machine in more detail, small modifications to the Linux kernel were done, enabling the logging of various internal variables, such as the congestion window size, the threshold value, and so on. In addition to this data, tcpdump was used to obtain packet traces for further analysis during all measurements.

III. Measurements Performed

Below is a list of measurement types performed, grouped by the application used for traffic generation.

0-7803-8887-9/05/$20.00 (c)2005 IEEE

A. *iperf*

The effects of MSS and window size settings were studied using iperf. Three MSS settings (548, 996 and 1360 bytes), and numerous window sizes between 2 and 64 MSSs were used.

B. *HTTP*

The behavior of HTTP transfers (certainly an important application area in the future) was studied using two types of webpages, and varying the number of parallel connections established by the browser. The webpage types considered were "photo albums" of varying size, consisting of a single page with images of constant size and mirrors of actual news portals, with large numbers of embedded objects of varying sizes.

C. *FTP*

Bulk traffic flow behavior was measured using large FTP transfers. In this case also mobility aspect was studied by performing measurements with the UE located in a moving car.

D. *Interactive Traffic*

While not performed in such a systematic manner as the ones mentioned above, packet traces of a number of interactive ssh- and X11-sessions were recorded, together with the TCP state machine internal variables.

E. *TCP Modifications*

In the final phase of the measurements, the behavior of the FRTO [2] and Eifel [3] algorithms were studied in the bulk traffic flow context.

IV. Results

In this section we shall go through the main results from the measurements.

A. *Overall Notes*

The UMTS channel proved to be extremely reliable. If the UE was stationary, no packet losses were observed, and even in the case of a mobile UE, very few losses were observed. The price to be paid of this reliability is, of course, the slightly "spread out" RTT distribution, caused by the variable number of RLC retransmissions. Handovers were present as visible "spikes" in RTT (Fig. 1), but in all cases the RTT grew slowly enough as to prevent spurious retransmissions taking place. Naturally, the handovers also become visible in the retransmission timeout counter, see Fig. 2.

It was reassuring to observe that UMTS works very well with the MSS and window sizes used by default in the main operating systems. While slight improvements to throughput can be achieved by tuning the parameters, end users not wishing to do this will not be punished by excessively low throughput. Also the subjective quality of UMTS channel in interactive use was quite fair. Main obstruction to Ethernet-like quality in interactive applications is definitely the higher RTT. As can be seen from Fig. 3, the "raw" RTT as measured by the ping-program is comparable to that of a modem connection, and definitely below the values typical to GSM connections.

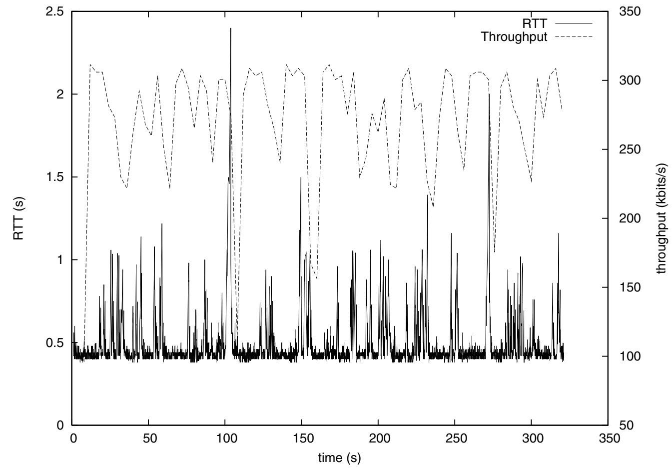

Fig. 1. RTT and throughput during an FTP transfer in a moving car, with two handovers visible at around $t = 100\,\mathrm{s}$ and $t = 270\,\mathrm{s}$.

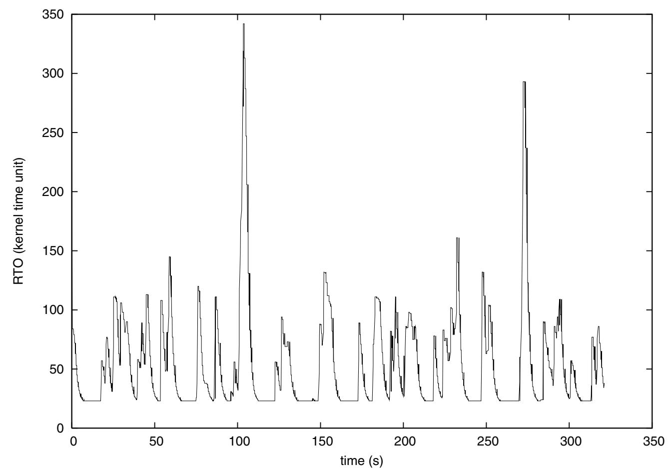

Fig. 2. The TCP RTO threshold during the FTP transfer in a moving car.

B. *Number of Connections, MSS and Window Size*

Especially in the case of HTTP with small object sizes, it was again observed that using multiple parallel connections makes sense from the user point of view, at least when the bit-pipe carrying the TCP connection is not full. In Fig. 4 the cumulative count of bytes transferred from a "photo album" website of eight pictures is shown. It is obvious that the browser using only a single TCP connection requires substantially longer time to transfer the contents of the whole webpage compared to the browser using multiple parallel connections.

The behavior of observed throughput as a function of the window size behaved in a similar manner to fixed networks. No adverse effects were noted when using a large window size. In Fig. 5 the dependency of the throughput on the window size for various MTU values is depicted from iperf measurements in the Vodafone network.

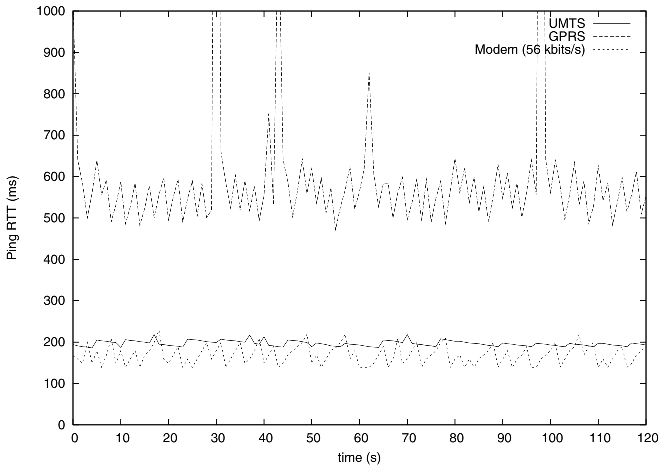

Fig. 3. The “raw” RTTs of UMTS, GSM and 56kbps modem channels as measured using ping.

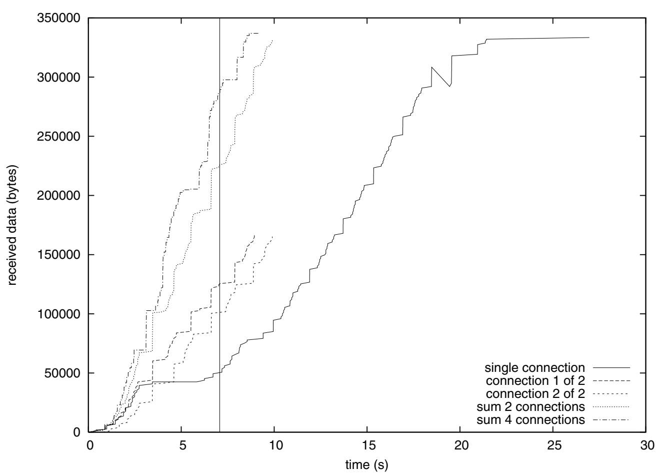

Fig. 4. The amount of received data as a function of time during download of a web page using varying number of connections. The vertical line marks the shortest possible theoretical transfer time ignoring higher-layer protocol overhead.

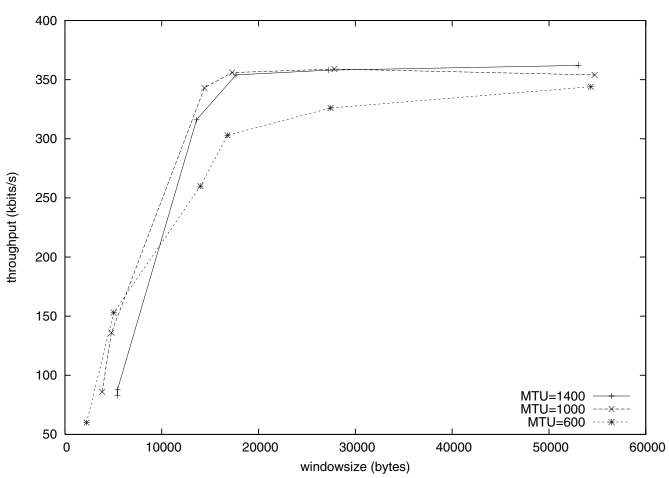

Fig. 5. Throughput as a function of window size for various MTU values.

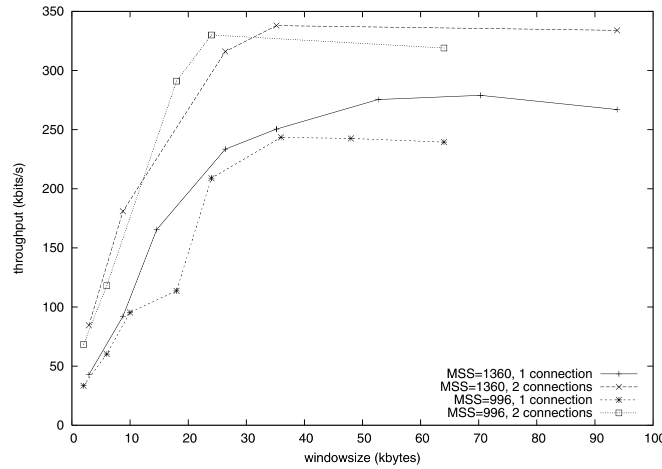

Fig. 6. Throughput as a function of window size for various MSS values, using one or two parallel connections.

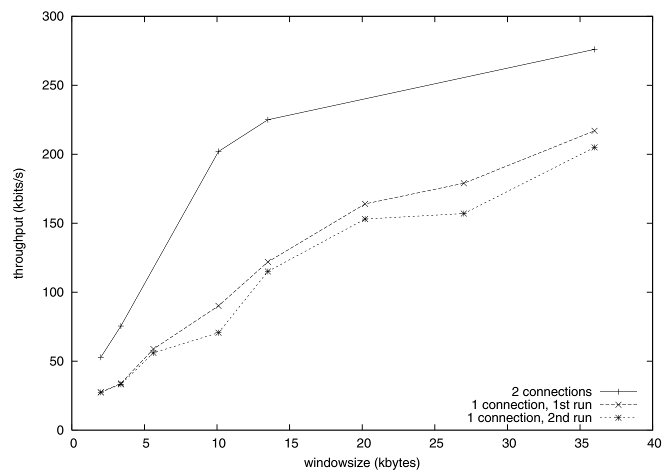

Fig. 7. Dependence of throughput on the window size for small MSS value, with one and two parallel connections. For illustration of the relatively high consistency of the UMTS channel, two individual runs of transfers using a single connection are shown.

From the iperf measurements in the Alcatel network similar conclusion was drawn over all window sizes. In Fig. 6 the throughput as measured by iperf is plotted as the function of the window size, using MSS values of 996 and 1360 bytes, and with one and two parallel connections. The benefit of multiple parallel connections is clearly visible, as is the use of larger MSS size (albeit the latter effect is more difficult to discern).

Similar conclusions hold also for smaller MSS sizes. Fig. 7 shows the dependency of throughput on the window size and the number of connections when MSS value of 548 bytes was used.

C. Transfers Between Two UEs

As mentioned earlier, measurements were also made concerning the behavior of data transfer over UMTS between two UEs. As an example of the results obtained, Fig. 8 shows an excerpt of the throughput and RTT plots measured from an FTP transfer of a 3.45 MB file between two mobile

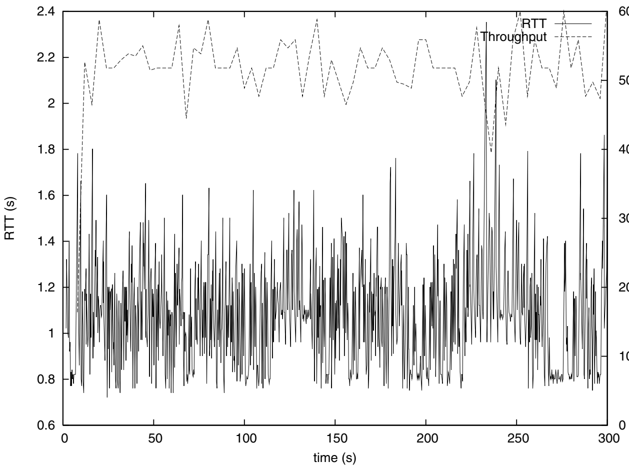

Fig. 8. The RTT and throughput during an FTP transfer between two mobile stations.

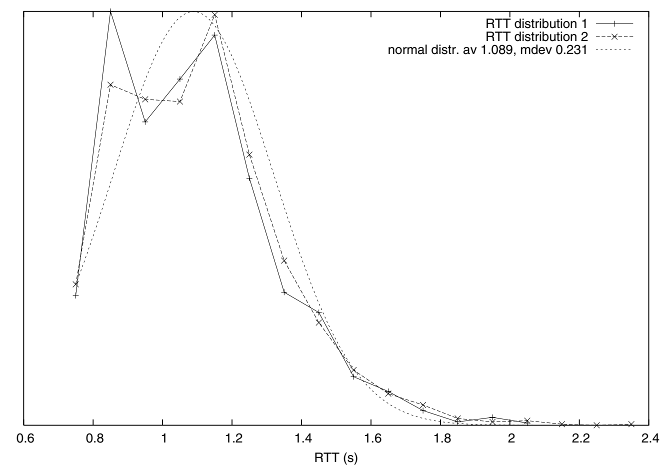

Fig. 9. The distribution of RTTs during the FTP transfer between two mobile stations.

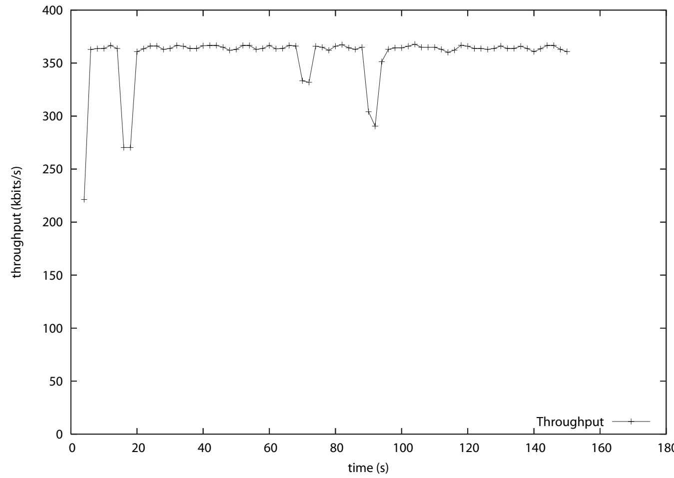

Fig. 10. Throughput during an FTP transfer as measured using Nokia handset.

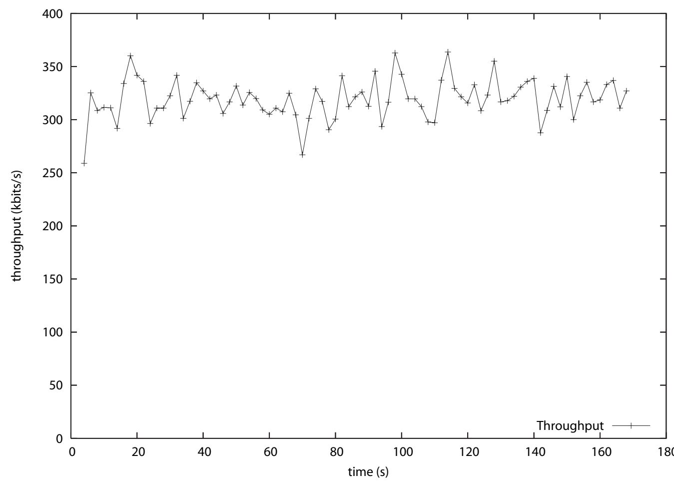

Fig. 11. Throughput during an FTP transfer using Motorola handset.

stations. Being limited by the slower uplink speed, the average throughput is reduced to 52.95 kbit/s, a value understandably much lower compared to download from the fixed network.

The high variability in the RTT measured is also striking. Fig. 9 shows the distribution of the RTT times as measured at the two mobile stations, together with normal distribution of same average and mean deviation. Two peaks are clearly visible in the RTT distributions, the origin of which could not be conclusively determined. Our working hypothesis is, that the peaks correspond to different RLC retransmission numbers the packet in question has been subject to.

D. Different User Equipments

Considerable differences were observed in terms of the RTT behavior between UEs from different manufacturers. Namely, the RTT distribution variance appeared to change from one UE to the other, leading, of course, also to notable changes in the achievable throughput.

To characterize these differences more quantitatively, we shall now present the throughput and RTT graphs corresponding to large FTP transfers with different UEs. For concreteness, the results from the Nokia and Motorola mobile phones are shown. The behavior of Samsung phone was quite similar to that of Nokia one. First, Fig. 10 and Fig. 11 show the throughput as measured using the Nokia and Motorola UEs, respectively. Obviously, the throughput with Nokia UE is more steady.

The connection of the difference in throughputs provided and the RTT can clearly be inferred from Fig. 12 and Fig. 13 where the RTT plots corresponding again to Nokia and Motorola UEs, respectively, are shown. Very large window size was used at this stage of the measurements, leading to rather larger average RTTs than encountered in the graphs above.

We wish to reiterate that as all of the UE models used were either quite new to the market, or in late testing phase, these results should not be taken as any kind of indication to the quality or performance of future mobile stations of any manufacturer. Instead, they serve to remind that any complex wireless technology can be implemented near-optimally, and

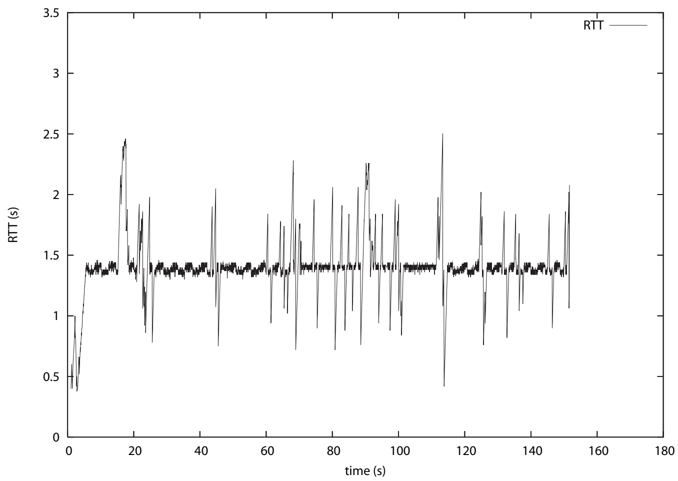

Fig. 12. RTT during FTP transfer using Nokia handset.

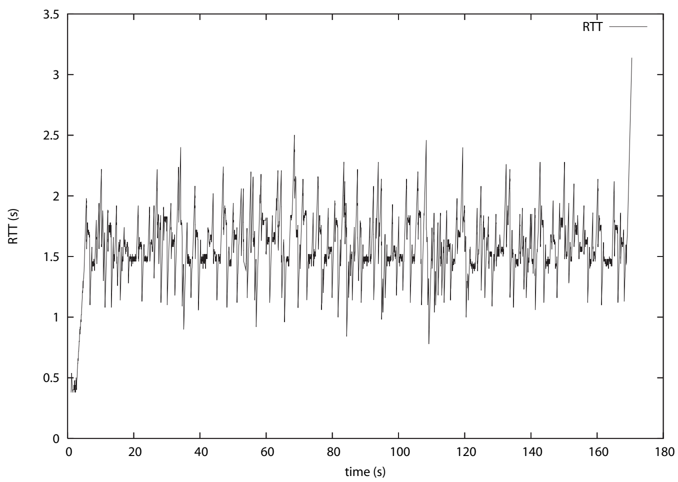

Fig. 13. RTT corresponding to FTP transfer with Motorola handset.

with little differences between manufacturers, only after considerable development effort, and gaining of experience.

E. TCP Modifications

Numerous measurements were performed with FRTO and Eifel algorithms enabled, but no change in TCP behavior compared to the standard implementations were observed. In fact, extra logging facilities added to the Linux kernel allowed us to monitor that, for example, the FRTO recovery algorithms for recoverable spurious retransmissions were not called at all. Thus, for TCP over UMTS, at least in good conditions, no advantages seem to exist in modifying the existing TCP retransmission timer implementations.

V. Conclusion and Discussion

We have seen that in the conditions the measurements were carried out UMTS offered a stable data channel for TCP to operate in. High packet loss rates and highly varying RTT commonly attributed to wireless channels were not observed. Thus, UMTS certainly has the potential to offer high-speed (compared to second generation cellular systems before most recent GPRS evolutions) Internet connectivity using standard protocols and settings.

Whether this potential continues to be realized in commercial networks in the future is an interesting question. In our measurements, no other users were to our knowledge present in the network. Due to the design of the UMTS system, these conclusions should also hold for networks used by not a single, but a small group of users. However, at present it is unclear what takes place when networks become truly crowded. If the operators were to deploy stringent enough call admission control procedures, the answer to this question would be without practical meaning. The commercial realities, on the other hand, make it unlikely that these expensive networks will be run with only a few users per cell. Therefore, the performance of the future UMTS networks should be examined when the time is right (with new measurements, of course).

Acknowledgment

We would like to thank all the people at Alcatel's Stuttgart campus and at Vodafone for their assistance during our measurements.

References

[1] A. Gurtov, "Effect of Delays on TCP Performance," Proceedings of IFIP Personal Wireless Communications (PWC'01), August 2001.
[2] R. Ludwig, and K. Sklower, "The Eifel retransmission timer," ACM SIGCOMM Computer Communication Review, vol. 30, pp. 17–27, July 2000.
[3] P. Sarolahti, M. Kojo, and K. Raatikainen, "F-RTO: An Enhanced Recovery Algorithm for TCP Retransmission Timeouts," ACM SIGCOMM Computer Communication Review, vol. 33, April 2003.