Analysis and Measurement of TCP/IP Performance over GPRS Networks

L. Benetazzo, M. Bertocco, C. Narduzzi, and R. Tittoto

Università di Padova, Department of Information Engineering,
Via Gradenigo 6/a – 35131 Padova, Italy
Phone: +39 049 827 7500; Fax: +39 049 827 7699
{luigino.benetazzo,matteo.bertocco,claudio.narduzzi,
ronny.tittoto}@unipd.it

Abstract. The paper presents an experimental analysis of some link layer characteristics of GPRS (General Packet Radio Service) mobile networks and discusses their relationship to TCP/IP performances. Mobile networks implement sophisticated mechanisms in order to provide mobility and capacity in a wide area. Unfortunately, large delays along the radio path may be introduced, thus reducing throughput in data transfers. The paper focuses its attention on a set of phenomena (excess queuing, delay spikes, redundant retransmissions and bandwidth oscillations) that were observed in real mobile GPRS systems. Analyses are carried out on a set of TCP traffic traces captured by passive monitoring at the main protocol interfaces. Several aspects related to TCP performances and throughput are then discussed and results obtained are compared with simulation studies and characterisations already available in the literature. Optimal network configurations and tuning of various parameters are also discussed.

1 Introduction

The need for heterogeneous interoperability of IP (Internet Protocol) across broadly different technologies has attracted considerable research interest, in particular with regards to wireless technology. While the 3G mobile standard [1] represents an example of a full-scale IP implementation designed to provide various customer services, 2.5 generation mobile wireless networks, such as GPRS [2] and EDGE [1], provide data services and Internet accessibility through the IP protocol, in addition to more traditional voice services on circuit-switched networks.

GPRS networks are deployed world-wide and their performance evaluation and optimisation is an active research area. In particular, the use of the TCP protocol [3], [4] over such networks has stimulated a number of studies and analyses aimed at the improvement of performances in heterogeneous networks. TCP performances are strictly related to network characteristics and are known to be often affected by implementation details and network configuration. For these reasons, it is important to assess how well different services are or can be implemented, by measuring significant network parameters. Furthermore, continuous network monitoring is required to detect possible dimensioning problems or failures during normal network operation.

Several aspects of network behaviour have been analysed in [5], and in the references therein, by means of multi-layer tracing at various interfaces. The analysis allows to understand network performance and to discover possible improper configurations. Network characterisation and further TCP performance analysis have also been presented in [6], where specific aspects, such as link buffering, are investigated. The approach followed by the authors in [5] and [6] was to implement the full GPRS connection, keeping both endpoints within reach in their laboratory. In this way, artificial traffic could be generated in a controlled manner.

From the viewpoint of a mobile network operator, traffic can be traced by probing only at some intermediate interfaces along the TCP connection path, since both endpoints of the connections are unreachable. Hence, operators usually have no access to the whole end-to-end connection and have to evaluate overall TCP performances by processing data gathered at their own monitoring points. This is known as passive monitoring and, in spite of some difficulties, it is a powerful tool for network troubleshooting.

This paper presents an analysis of traces captured by passive monitoring of real GPRS traffic at the interfaces typically available to a network operator. The possibility of detecting and quantifying relevant behaviours is discussed and demonstrated, by showing how various TCP anomalous behaviours due to mobile link characteristics can be evidenced. Results are compared with previously published work, where relevant. The proposed analysis could provide useful tools to improve quality of service from the end user point of view.

The next section introduces GPRS link layer characteristics, whereas Sect. 3 describes several details regarding the measurement set-up and observed phenomena on real networks. Final remarks presented in Sect. 4 conclude the paper.

2   GPRS Link Layer Characteristics

Limited data rates, high latency, delay variation, bursty losses and bandwidth oscillations are important GPRS link layer characteristics, as described in [7]. Bandwidth-Delay Product (BDP) values between 4 and 10 kBytes are typical, therefore these networks can be classified as LTNs (Long Thin Networks) [8].

To allow a better understanding of the issues that will be discussed in this paper, some network elements and a number of protocols are briefly introduced. Fig. 1 illustrates the most relevant elements of a GPRS network from a performance monitoring point of view. During a data transfer between a mobile terminal (MT) and an Internet server, packets are processed by the Gateway GPRS Support Node (GGSN), the Serving GPRS Support Node (SGSN), the Base Station Controller (BSC) and the Base Transceiver Station (BTS). GGSN handles connectivity to external networks, whereas SGSN deals with mobility over a single routing area and controls the data flow toward BSC over the *Gb* interface. Finally BSC handles the radio resource scheduling, medium access and data transmission toward MTs over the *Abis* interface. A detailed description of the GPRS system can be found in [2].

Fig. 2 shows the user plane GPRS protocol stack. Protocols that will be considered in this paper have been evidenced with a different colour. The Radio Link Control protocol (RLC) provides acknowledged or unacknowledged data transfer between MTs and BSCs. The Logical Link Control protocol (LLC) provides packet transfer

between MTs and SGSNs; it can operate in both acknowledged or unacknowledged mode, though most GPRS networks operate in the unacknowledged mode. Maximum size of the LLC frames can be configured up to 1556 bytes. The Base Station Subsystem GPRS Protocol (BSSGP) provides data control flow between BSC and SGSN; in fact, it is used by BSCs to communicate buffer status and leak rate for the downlink flow. Finally, MT and Internet hosts communicate by means of the TCP/IP protocol suite.
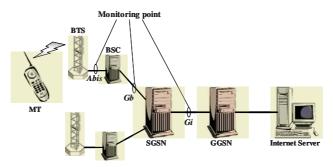

Fig. 1. GPRS network elements involved in a data connection and measurement probe points location
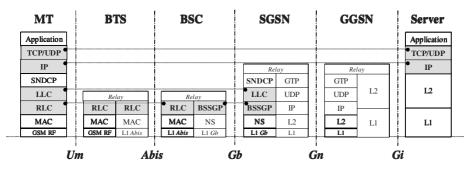

Fig. 2. The GPRS protocol stack

Mobile wireless networks exhibit several differences from traditional fixed networks which, in some cases, affect TCP performances. Data rate is the most important parameter to take into account. Typical data rates for 2.5G systems (GPRS and EDGE) are between 10kbps and 40kbps. In addition, actual data rates change dynamically: users connected to a cell can ask for resources or release them, thus increasing or reducing the available bandwidth, which is also influenced by the distance from the base station. Furthermore, users can move from a cell to another, experiencing abrupt changes in bandwidth. These phenomena, often referred to as bandwidth oscillations, may cause TCP malfunctions and degrade throughput.

Latency is another important characteristic. For GPRS radio links this is typically high, due to the mechanisms implemented to provide mobility, capacity in a wide area and reliability along the radio path. For the latter purpose, the RLC protocol [9] uses the acknowledged mode, implementing specific ARQ (Automatic Repeat reQuest)

and FEC (Forward Error Correction) techniques, which ensure reliability at the cost of increased latency.

Delay along a wireless path also exhibits large variations, due to changing radio conditions and distances between MTs and BTSs. Delay variations compromise data transfer quality and appear like jitter at higher level protocol layers. Latency can increase abruptly and these variations could generate TCP malfunctions such as redundant retransmissions, as will be seen in the next sections.

GPRS networks do not exhibit high packet loss rates and values typically found in normal operational conditions (e.g. no handover) do not affect TCP performances.

3   Monitoring TCP/IP over Mobile Wireless Networks

Mobile operators need to ensure that network elements are configured properly in order to guarantee best performances for their customers. For this reason it is fundamental to evaluate network behaviour by means of distributed measurement devices implementing specific analysis algorithms. Monitoring is usually carried out by passive measurements, i.e., by observing existing network traffic, rather than by generating dedicated test flows. Tektronix NET-GPRS is a monitoring system that enables operators to troubleshoot, optimise and manage GPRS networks by means of protocol analysis, multi-interface multi-protocol procedure trace and statistical measurement. It can monitor different network interfaces by means of several probes, ensuring minimal timestamp uncertainty and synchronisation between different probe points by referencing to GPS (Global Positioning System) time. Data are collected by the various probe boards and sent to a central storage unit where protocol analysis and procedure trace can be performed by specific software tools.

Data analysed in this paper were collected, using this system, from the *Gi*, *Gb* and *Abis* interfaces (see Fig.1) of various real GPRS networks. For this reason, traces do not completely represent neither the sender side, nor the receiver. Although additional assumptions may have to be introduced about the relevance of some factors on the measured values, resulting evaluations are generally quite accurate.

Since packet delay is an important aspect of packet-switched network dynamics, the analysis presented in this paper focuses on transport protocol behaviour and is based on the measurement of RTT (Round Trip Time). Experimental results are discussed with the aim of relating possible causes for RTT variations to buffering and flow control procedures.

A set of over 800 traces have been analysed, accounting for over 500000 LLC frames. Useful information, including TCP sequence numbers, IP addresses, port numbers, flags, etc., are extracted from header packets by the NET-GPRS software tools. The capture time is indicated by a timestamp associated to each packet. Traces are seen as transport level flows, defined by different (source IP address, source port number, destination IP address, destination port number) 4-ples.

Main data services accessed through GPRS networks are web browsing (HTTP and HTTPS), email (POP3 and SMTP) and file transfers (FTP etc.). Preliminary packet analysis showed that TCP accounts for 83% of the total IP traffic. UDP is also frequently used (17%), mainly for Wireless Application Protocol (WAP) services. As reported in Table 1, the best part of the captured traces is related to data downloads. This is confirmed by packet size analysis that showed a large number of small IP

packets, of size between 40 and 60 bytes, in the uplink direction and IP packets larger than 500 bytes in downlink.

Table 1. Percentage of TCP connections for the different type of services.

	HTTP HTTPS	POP3	SMTP	FTP	Others
TCP traces	78,95%	6,62%	2,60%	9,45%	2,38%

Since data transfers in GPRS networks suffer from large delays in the downlink direction, we focused our analysis on large data downloads from Internet servers to MTs. For these traces it is possible to evaluate the partial RTT experienced along the radio tract by evaluating the time elapsed between the capture of a packet and the corresponding acknowledgement (same sequence number) observed in the opposite direction. This is what a sender usually does, for instance, to determine its own retransmission time-out (RTO). In this case, however, the measurement is performed using packets collected at the *Gb* interface; the partial RTT evaluated in this way can give a good estimate of the time elapsed in the BSC, the BTS, the MT and along the air path, allowing to keep trace of the evolving characteristics of the radio path alone.

In order to correctly interpret the behaviour of transmitters and receivers in the analysed connections, a preliminary examination was carried out in the attempt to detect the various TCP implementations used by the hosts involved. For the majority of traces behaviour was very similar way to that expected for SACK and New Reno implementations [10]. Predictably, some differences were observed, and can be related to proprietary implementations of TCP in specific operating systems.

3.1 Partial Round Trip Time Measurements

Several partial RTT values were obtained for each trace considered in the analysis. Each captured acknowledgement in the uplink direction was considered together with its corresponding packet and a partial RTT was calculated, as the difference between the respective timestamps, in every non-ambiguous [11] case.

Fig. 3 presents a histogram of the measured RTTs obtained analysing a set of over 200 TCP connections. These traces are all HTTP downloads of 43 kbytes and have been captured in a test plant with perfect radio conditions. As can be noticed, values are typically greater than 1 s, with a mean RTT of 1.5 s. Values are considerably high if compared with normal Internet wired connections, anyway they agree with measurements previously reported in the literature [6].

By comparison, Fig. 4 shows measured partial RTTs in the case of bulk download traces gathered from real traffic. In this case different services are considered, such as HTTP, FTP and POP3. The majority of the MTs analyzed uses 4 or 3 timeslots in downlink and 1 timeslot in uplink. Also, code scheme 2 (CS-2) encoding is widely used since it provides good performances if compared with CS-3 and 4. For these mobiles the bandwidth-delay product (BDP) is approximately equal to 10 kbytes.

The figure refers to 4795 measures divided in two main groups: the first group, represented by white bars, refers to all RTT estimates obtained from TCP traces with a receiver advertised window up to the BDP (10 kbytes in this case), which is suggested as the reference value [7]. The second group, represented by black bars,

refers to the remaining traces, where the receiver advertised window was greater than the BDP. It should be emphasised that the histogram has been obtained by analysing only large bulk transfers of at least 20 kbytes; in this way, it is ensured that the size of the transmitter congestion window for the second group of traces is as close as practically possible to the maximum allowed value.

As can be seen from the figure, large window sizes imply larger RTT values. For long TCP sessions the transmitter fills the buffer located at the BSC excessively, leading to partial RTTs that can be 4 to 6 times longer. This confirms the excessive queuing phenomenon described in [6], where a number of consequences are briefly described (RTT inflation, inflated retransmission timer value, problems of stale data and high recovery time).
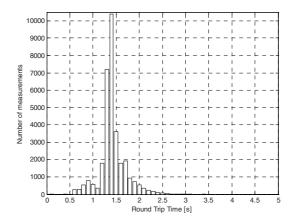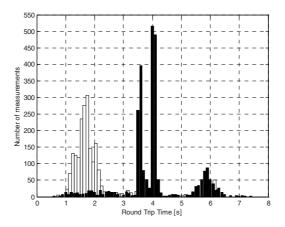

Fig. 3. Partial RTT evaluated analyzing a set of over 200 HTTP downloads of 43 kbytes

Fig. 4. Partial RTT evaluated analyzing a set of over 120 TCP downloads of at least 20 kbytes. Black bars refers to measurements collected from traces with advertised window greater than 10 kbytes, white bars refers to traces with advertised window up to 10 kbytes.
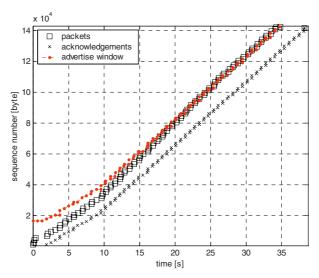
Fig. 5. A TCP trace example (advertised window equal to 16 kbytes); delay gets larger as the congestion window increases
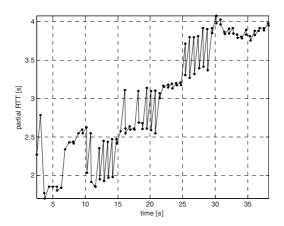

Fig. 6. Partial RTT estimates for the trace of Fig. 5

Fig. 5 shows a TCP trace, depicted using squares to represent packets and "X"s to represent acknowledgements. The x-axis shows the packet timestamp (capture time) expressed in seconds. The y-axis shows the packet sequence number expressed in bytes. Sequence numbers and timestamps are scaled, so that the starting value appears to be always equal to 0 for each connection. The dot-dash line represents the receiver advertised window.

Fig. 6 presents the corresponding partial RTT measurements as functions of the capture time. It can be noticed that delay gets larger as the transmitter increases its congestion window. The maximum allowed size is set to 16 kbytes, therefore the congestion window can reach values greater than the optimal BDP.

Further analysis of other traces confirmed that best performances are achieved if the maximum congestion window is set to the BDP; in practice, 10 kbytes is a good value for the case of a 4-timeslot MT using CS-2. Higher values must be used in the case of higher encoding schemes, CS-3 and CS-4, or for MTs that can support more timeslots.

3.2   Link Level Buffering at the BSC and Additional Delays

It has already been reported that substantial buffering capacity is provided within the GPRS infrastructure [6]. Measurements presented here show that the element involved is primarily the BSC. Precise transit delay estimates can be obtained by monitoring at the same time the *Gb* and the *Abis* interfaces. In this case delay within the BSC is simply given by the difference between the timestamps that the monitoring device associates to the same downlink LLC frame at the two interfaces.
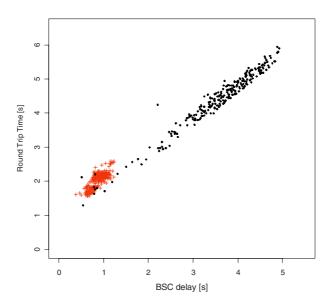
In Fig. 7 two scatter plots of the BSC transit delay versus the partial RTT evaluated at the *Gb* interface are reported. Both measurements have been obtained analysing 500-kbytes TCP downloads (4-timeslot MT using CS-2). Dots refer to the case of a large window size (advertised window equal to 64 kbytes). As can be noticed, partial RTTs and correspondent delay queues are extremely high and highly correlated. In fact, a correlation coefficient equal to 0,98 is obtained as a remarkable proof that round trip delay is mainly conditioned by buffering at the BSC. For real traces with large congestion window size (greater than 10 kbytes) delay values of about 3 or 4 s are typical and buffering can reach values of 24 kbytes and more.

The second plot of Fig. 7, depicted using '+', shows how correlation decreases in the case of a smaller advertised window (8 kbytes). A correlation coefficient equal to 0,74 is obtained and RTT estimates are partially conditioned by other aspects such as radio transmission, resource allocation, etc.

The GPRS network constantly controls buffering at the BSC for each single MT and for the total amount of traffic. Specific flow control messages (FCM) are exchanged between the BSC and SGSN by means of the BSSGP protocol. Basically, the flow control algorithms try to keep a constant buffer level at the BSC by delaying LLC frames at the SGSN. FCMs are sent periodically from the BSC to the SGSN, informing it of the estimated downlink leak rate (R) calculated by the RLC protocol and of the maximum buffer size (B_{max}).

Detailed analysis of the current flow control implementations showed that better performances are achieved using small size LLCs. In fact, with LLCs of approximately 500 bytes the system ensures less variable delays at the BSC buffer, reducing jitter seen by higher layer protocols. Since GPRS networks provide transparent link layer fragmentation, performed both by SNDCP and RLC protocols, these size limitations do not need to reflect on the IP MTU (Maximum Transfer Unit). Further considerations regarding IP datagram maximum size will be discussed later in this paper.
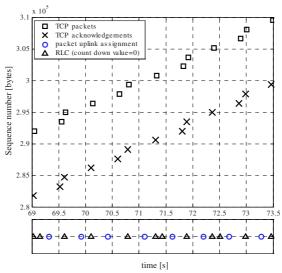
Buffering at the BSC is not the only cause of high latency in GPRS networks. For example, considering Fig. 7 it can be seen that packets spend approximately 1 s along the radio tract and at the MT. This has to be taken into account when computing the partial RTT observed at the *Gb* interface. A rough assessment of the way this delay builds up can be given as follows.

Fig. 7. Scatter plot of the delay queue versus the partial RTT evaluated at the *Gb* interface; dots, '.' refers to a trace with large TCP transmitter window values; '+' refers to a trace with TCP transmitter window limited to 8 kbytes

Data of Fig. 7 refer to TCP connections using IP datagrams of 1500 bytes. Since CS-2 and 4 timeslots are used for both cases, data packet radio transmission in the downlink direction takes approximately 350 ms [12], [13]. A corresponding 52-byte TCP acknowledgement is transmitted uplink, using a single timeslot, in approximately 50 ms. The remaining 600 ms are due to the use of delayed acknowledgement (200 ms), packet buffering and processing time at the various network elements; a delay contribution may also come from radio resource requests that periodically occur for the uplink transmission, as discussed later on.

Another important issue regarding TCP over GPRS is related to delayed acknowledgements. In fact, considering current performances of 4-timeslot MTs, practical considerations reported in [14] suggest that a small IP MTU (e.g. 576 bytes) should be used. However, analysed traces have shown that, even with these MTU sizes, the IP datagram can in some cases occupy the radio link for longer than the delayed ACK timeout, causing a considerable increase of the total RTT. A simple solution could be to acknowledge every segment [8]. It is the authors' opinion that ack-every-segment can be applied for the total duration of the TCP connection in a GPRS network. In fact, this rather drastic solution could also partially solve the problem related to periodic radio resource request and the consequent acknowledgement compression described in the next subsection.

Fig. 8. A TCP download example and relative RLC signaling showing the periodic release and assignment of the uplink radio link.

3.3 RTT Variability

Partial RTT estimates obtained from the majority of the traces analysed in this paper exhibit large variability. This can be observed, for instance, in Fig. 6, where RTTs oscillate about the local mean value with an amplitude equal to 1 s. This behaviour is very common in real GPRS traces and can be defined as transport layer jitter. The causes of this phenomenon are various and are not only related to flow control mechanisms discussed above (Sub. 3.2).

A not uncommon phenomenon, also observed in [5] and [6], is acknowledgement compression, a situation where acknowledgements seem to be released in bursts. The main cause of this behaviour is related to the frequent release of the uplink radio resources by the MT. Resources on the uplink air interface are released by applying the countdown procedure [9]. The mobile station continuously keeps track of the number of RLC data blocks in the transmission queue and, when there is only a certain number left, initiates the countdown procedure. This is done by decreasing the so-called countdown value (CV) for each RLC block sent uplink, the very last block being sent on reaching zero. If data have been correctly received, the BSC then releases the radio link.

It is important to note that the countdown procedure is irreversible. When a MT has only small IP packets to transmit in uplink, such as TCP acknowledgements, the count down procedure is frequently started, so that radio link releases and requests alternate during the whole length of the connection. An example is shown in Fig. 8, where basic RLC signalling is plotted along the TCP trace. Triangles represent uplink RLC data blocks with CV equal to 0: they indicate a subsequent radio link release performed by the BSC. Packet Uplink Assignments (PUAS), depicted using circles, are always present after these RLC blocks, proving that the monitored MT requested

back the radio resources. This mechanism is the main cause of acknowledgement compression and large RTT variability.

Other causes for RTT variations are periodic radio resource uplink losses and reassignments and burst reorganisations due to buffering and fragmentation performed by the RLC protocol.

An example of burst reorganisation is given in Fig. 9 where the burst structure of the transmitted TCP packets, as observed at the monitoring point, does not correspond to that of TCP acknowledgements returning from the receiver (MT). In view of the behaviour of TCP receivers, acknowledgements can be assumed to reflect the actual burst structure of packets as they are received. The figure shows that the last packets of a TCP burst can be buffered and finally depart with the first packets of the following burst. This phenomenon is mainly related to the RLC window size [9] used in the downlink direction and can be observed for older GPRS MTs. Analysis on real traffic has shown that this anomalous behaviour is extremely rare in recent networks where RLC window size is typically equal or greater than 192 radio blocks.

Both phenomena described here, together with high delay variation due to buffering at the BSC, are significant causes of bad TCP retransmission timeouts (RTO) and redundant retransmissions (RR) described in detail in the following section.
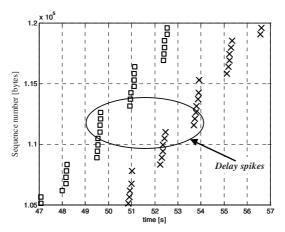

Fig. 9. Burst reorganisations due to buffering and fragmentation

3.4 Redundant Retransmissions

Several simulation results are available in the literature [15] regarding TCP/IP behaviour over GPRS networks. The most common phenomenon observed during these simulations regards redundant retransmissions (RRs) caused by high round trip delays and bad retransmission time-outs. Simulations show that the principal cause of RTO expiration are delay spikes, which are usually associated to dynamic allocation of resources and mobility related mechanisms. Analysis of real traces presented in this paper showed that a third cause is related to the phenomena described in Sect. 3.2 and 3.3.
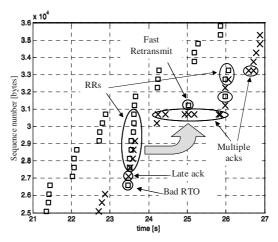
Fig. 10. Typical GPRS bad retransmission time-out and successive related RRs

In order to determine the exact percentage of RRs, it would be necessary to track both sender and receiver traces [16]. Unfortunately, this is not possible for a mobile operator and for this reason RRs evaluation is hard and cannot be sufficiently precise if implemented in a passive monitoring device. A simple yet effective solution for this problem could be the calculation of the ratio between the total number of sent TCP packets and the net number of packets excluding retransmissions. This ratio can be computed easily for each monitored TCP connection and traces exhibiting high ratio values should be analysed in greater detail.

An example of RRs is given in Fig. 10. In this case the transmitter congestion window reaches values around 5 kbytes and phenomena similar to the ones described in the previous sections can be observed. Between the beginning of the connection and 22 s the TCP sender progressively adjusts its RTO to a value approximately equal to 2 s. As shown in the figure, this situation is interrupted after a delay spike that occurs at around 23 s. At this point a first packet is retransmitted after a RTO; this is actually a redundant retransmission, as proved by the fact that a "late acknowledgement" is observed just a few milliseconds later. In this case, because of burst reorganisation, the acknowledgement arrived with the first packet of the successive burst. Typically, after a RTO the sender enters the slow start phase so that, for each successively received acknowledgement, it retransmits successive packets. A total number of 7 retransmitted packets have been reached in Fig. 9. Since the packets had already been received at the MT, multiple acknowledgements are generated, as evidenced by the "x" between 24 s and 26 s. These make the sender enter the fast retransmit phase, whereby other packets are unnecessarily retransmitted.

While the results reported agree with the behaviour described in the literature, they show that delay spikes and RRs can also be related to burst reorganisation, rather than to resource allocation and mobility. Other traces showing similar phenomena have been analysed. It must be pointed out that, after the RRs, the transmitter reverted to a stable and regular transmission. This is due to the fact that the sender enters the fast retransmit phase only when three multiple acknowledgements are received. In other words, even a large number of RRs compromises TCP performances only in a

relatively limited period, since TCP mechanisms ensure a further regulation in subsequent traffic.

RRs inevitably compromise TCP performances [16]. In order to develop non-invasive tools capable of correct performance evaluation, RRs need to be considered since they represent important aspects related to TCP/IP over mobile networks.

3.5 Bandwidth Changes

Users can get variable resources as a function of the radio path characteristics, distance and number of MTs simultaneously accessing the same cell. These variations can be observed also by monitoring TCP/IP traffic. Captured traces related to real GPRS traffic exhibit these behaviours. Examples are given in Fig. 11. In this case the traces shown refer to large file transfers and present abrupt bandwidth changes and high delay variations.

Other traces with similar characteristics have been observed, especially regarding long file transfers lasting more than 30 seconds. This behaviour is strongly related to the mobile mechanisms implemented to share the limited resources available along the radio path. Handoff operations can also lead to abrupt delay variations with total RTT that can reach even 10 s and longer.

When link characteristics change abruptly, long delay spikes and packet loss are frequently observed. However, in the majority of cases analysed during this work, these phenomena were not accompanied by redundant retransmissions. The GPRS traces presented also demonstrate how bandwidth oscillations can occur frequently during a single download. It can be noticed that, because of the fast variations, it would be difficult to evaluate bottleneck bandwidth, especially when monitoring TCP/IP traffic by a non-invasive technique.
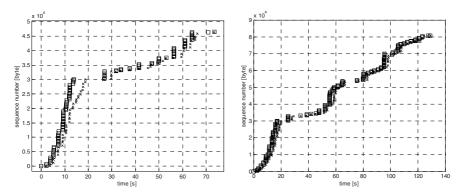

Fig. 11. Bandwidth oscillations due to resource sharing and handoff procedure, real GPRS traffic.

4   Conclusions

This article discusses a set of phenomena that can be observed by analysing real TCP traces over new generation mobile networks. The work points to some problems related to TCP/IP and GPRS wireless links. The purpose is to provide experimental evidence of behaviours that have already been noticed in the literature by means of simulations, but also to describe some peculiar behaviours in detail and propose optimal network configurations that allow to avoid them.

Further work has to be done in order to develop measurement techniques that could help network operators discover possible anomalies and improve performances. In particular, the relationship between delay variability and flow control is currently under closer investigation. The intention is to achieve a more precise evaluation of this type of phenomena.

Traces reported here refer to real GPRS traffic obtained from various mobile operators. Results are in agreement with the analyses in [5], [6], proving that passive monitoring can effectively provide operators with the tools for an accurate assessment of network performances and optimisation. It is hoped to extend the analysis to UMTS traces as soon as possible, with the intention of discovering differences and improvements related to this new technology.

References

1. H.Kaaranen, A.Ahtianinen, L.Laitinen, S.Naghian, V.Niemi: UMTS Networks: Architecture, Mobility and Services. John Wiley & Sons, 2001.
2. G.Brasche, B.Walke: Concepts, Services, and Protocols of the New GSM Phase 2+ General Packet Radio Service. IEEE Communications Magazine, August 1997.
3. J.B.Postel: Transmission Control Protocol. RFC 793, September 1981.
4. W.R.Stevens: TCP/IP Illustrated, Volume 1: The Protocols. Addison-Wesley, 1994.
5. A.Gurtov, M.Passoja, O.Aalto, M.Raitola: Multi-Layer Protocol Tracing in a GPRS Network. Proceedings of the IEEE Vehicular Technology Conference (Fall VTC2002), Vancouver, Canada, Sepember 2002
6. R. Chakravorty and I. Pratt: Performance Issues with General Packet Radio Service. Journal of Communications and Networks (JCN) – Special Issue on "Evolving from 3G deployment to 4G definition", pages 266–281, Vol. 4, No. 2, December 2002
7. H.Inamura, G.Montenegro, R.Ludwig, A.Gurtov, F.Khafizov: TCP over Second (2.5G) and Third (3G) Generation Wireless Networks. IETF RFC 3481, February 2003.
8. G.Montenegro, S.Dawkins, M.Kojo, V.Magret, N.Vaidya: Long Thin Networks. IETF RFC 2757, January 2000.
9. 3GPP TS 04.60: Radio Link Control/ Medium Access Control (RLC/MAC) protocol. Release 1999, January 2002.
10. K.Fall, S.Floyd: Simulation-based Comparisons of Tahoe, Reno, and SACK TCP. Computer Communication Review, Vol. 26 No. 3, pp. 5–21, July 1996.
11. V.Paxson, M.Allman: Computing TCP's Retransmission Timer. RFC 2988, Proposed Standard, November 2000.

12. M.Meyer: TCP Performance over GPRS. IEEE Wireless Communications and Networking Conference, New Orleans, LA, September 1999.
13. U.Vornefeld: Analytical Performance Evaluation of Mobile Internet Access via GPRS Networks. In Proc. of the European Wireless '02, pp. 712–718, Florence, Italy, February 2002.
14. S.Dawkins, G.Montenegro, M.Kojo, V.Magret: End-to-end Performance Implications of Slow Links. BCP 48, IETF RFC 3150, July 2001.
15. A.Gurtov, R.Ludwig: Evaluating the Eifel Algorithm for TCP in a GPRS Network. In Proceedings of European Wireless, Florence, Italy, February 2002.
16. V.Paxson: End-to-End Internet Packet Dynamics. IEEE/ACM Transactions on Networking, Vol.7, No.3, pp. 277–292, June 1999.