Passive Tomography of a 3G Network: Challenges and Opportunities

Fabio Ricciato*, Francesco Vacirca*, Wolfgang Fleischer†, Johannes Motz ‡ and Markus Rupp§
*Forschungszentrum Telekommunikation Wien (ftw.)
†mobilkom austria AG&CoKG ‡Kapsch CarrierCom §Technical University of Vienna

***Abstract*— Several applications of practical interest stem from the capability to monitor and store packet-level traces in a 3G network. Among them, the possibility to infer and locate network problems (e.g. persistent shortage of capacity, or equipment misfunctioning), in the core and radio sections, without direct access to the equipments. This approach yields strong practical benefits, given the costs and complexity of accessing network equipments, especially in the Radio Access Network. At the same time, it exposes practical issues - e.g. the need to dynamically locate the traffic sources (Mobile Stations) - and theoretical problems - e.g. inferring congested cells from Routing-Area level TCP measurements. We report on our work-in-progress aimed at implementing such mechanisms on top of an advanced monitoring system now deployed in an operational network.**

I. Introduction

Public wide-area wireless networks are now migrating towards third-generation systems (3G), designed to support packet-switched data services. Europe has adopted the Universal Mobile Telecommunication System (UMTS), developed by 3GPP as an evolution of GSM. A 3G network includes two main sections: a packet-switched Core Network (CN), which is based on IP, and one or more Radio Access Network (RAN). Along with the UMTS RAN (UTRAN) based on W-CDMA, several operators maintain a parallel GPRS RAN evolved from the legacy GSM radio. This structure is sketched in Figure 1. Several UMTS networks became operational since 2003 while first deployments of GPRS date back to 2000. Since then, the growing popularity of 3G terminals and services has extended the coverage of Internet access to the geographic area, and 3G networks are becoming key components of the global Internet in Europe. However the 3G environment is still under evolution, at least along the following dimensions:

- subscriber population and traffic volumes;
- terminal capabilities and relative penetration of the various terminal types (handsets, laptops with 3G card, etc.);
- service portfolio and tariffs offered by the operators.

Furthermore, technological upgrades are still in the agenda of many operators: EDGE in the GPRS RAN, HSPDA in the UMTS RAN, IMS in the CN [1]. All these aspects collectively build a potential for changes in the global traffic that can occur at the macroscopic scale (network-wide) and in a relatively short time frame. Hence, the ability to accurately and extensively monitor the network state and to early detect drifts in performance and/or local troubles is a fundamental pillar of the network operation and optimization process.

Contact author: Fabio Ricciato (ricciato@ftw.at)

Monitoring a wide-area network is not an easy task. First, the number of elements is large and they are spread geographically. Secondly, for most practical purposes it is necessary to access configuration parameters (e.g. provisioned banwidth), logs and counters from several network elements, with different software and from different vendors, and considerable costs, complexity and complications are found in practice where it comes to extraction, gathering and correlation of such heterogeneous data. In summary, installing and maintaining a monitoring infrastructure with direct access to the network elements of the production network is very expensive. Furthermore, the quality and the granularity of the data available from the equipments (e.g. built-in counters) is often poor and/or inadequate for in-depth analysis of the network state.

In our research we are exploring the feasibility of monitoring a production 3G network exclusively by passive sniffing packets on few key CN links, without direct access to the equipments. This approach yields a number of practical benefits, and opens new further opportunities for improving the engineering and operation of a real network. For example, the combination of location data extracted from the signaling frames with TCP performance indicators (e.g. retransmissions and RTTs) estimated with previously developed methods allows to monitor the actual perfomances of the whole RAN, and to spot the need for local radio re-optimization intervention, without direct access to the RAN equipments. This approach is similar in principle to passive network tomography [2] [3]. Another value point of passive monitoring, when coupled with trace storage, is the possibility to perform post-mortem analysis of network troubles. Applied to 3G, it allows for pioneering research directions unexplored so far, like the assessment of the potential impact of undesired traffic (e.g. worm infections, DoS attacks) onto the functionally-complex 3G infrastructure (an instructive case for the wired network is reported in [4]), or the analysis of signaling traffic to reveal buggy terminals or network misbehaviour on the control-plane. In summary, while such approaches are not novel in the general sense, nonetheless their application to the specific context of 3G networks reveals new unexplored facets.

In this contribution we try to enlight the opportunities of fine-grain monitoring a 3G network, with a focus on the problem of large-scale packet-level performance monitoring. We report on the main technical issues and the open points

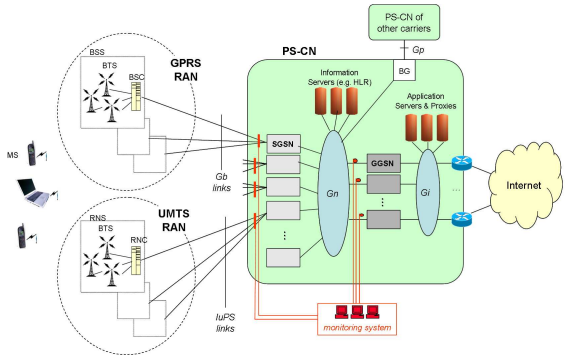

Fig. 1. 3G network structure.

we are encountering in implementing such mechanisms on top of an extensive monitoring system that was developed in a previous project, METAWIN [5], and is now deployed in the operational network of mobilkom austria. Such advanced features are meant to be used experimentally on a production basis by the operator as they are developed and stabilized.

This work is part of the research project DARWIN (Data Analysis and Reporting for WIreless Networks [5]), run in close collaboration between industry and academia. The project partners are: mobilkom austria AG&CoKG (the leading mobile communications operator in Austria, EU), Kapsch CarrierCom (provider of equipments and network engineering services), the Forschungszentrum Telekommunikation Wien (a private/public co-founded research center) and the Technical University of Vienna. The achievements of METAWIN and DARWIN are prominent examples of the factual value, on both the scientific and industrial sides, produced by the concrete collaboration between industry and research institutions.

II. Monitoring Setting

The reference network scenario is depicted in Figure 1 (for details on the GPRS/UMTS network structure see e.g. [1]). As any other access network, the 3G network has a hierarchical tree-like deployment. The Mobile Station (MS) and base stations are geographically distributed (nation-wide). Going up in the hierarchy (first BSC/RNC, then SGSN, ultimately GGSN) the level of concentration increases, involving progressively smaller number of equipments and physical sites. In a typical network there are relatively few SGSNs and even fewer GGSNs. Therefore it is possible to capture the whole data traffic from home subscribers on a small number of Gn/Gi links. However, for some application it is required to access the Gb/IuPS interface near the SGSNs. In any case, monitoring each interface requires the acquisition system to be capable of parsing and interpreting the full 3G protocol stack, for both the signaling and user plane (see [1, pp. 94, 202]).

The development of a large-scale passive monitoring system, including a parser for the whole PS-CN protocol stack, and its deployment in the operational network were accomplished within the METAWIN project [5], enaling us to passively monitor all CN interfaces (Gi, Gn, Gb, IuPS). Frames are captured with DAG cards and recorded with GPS synchronized time-stamps. For privacy requirements traces are anonymized by hashing all fields related to user identity at the lower 3G layers (IMSI, MSISDN, etc.), while the user payload above the TCP/IP layer is removed.

As discussed afterwards, one key aspect in several monitoring applications is the ability to refer each packet to the corresponding MS and to its current radio location, i.e. the cell or, at higher granularity, the so called Routing Area (RA). It is possible on the Gn interface to refer packets to MS, but not to its location. In fact, when exchanging data traffic MS are required to establish a logical connection with the GGSN, the so called “PDP-context”, which is maintaned for the whole activity period (ranging from sub-second to several hours). The PDP-context is similar to a modem dial-up connection in ISP networks. It is an important entity in the 3G dynamics, since several functions are performed on a per-PDPcontext basis (e.g. the assignment of IP address, generation of billing ticket) and important parameters of interest (e.g. APN, QoS parameters) can be sniffed only during the PDP-context opening procedure. Our monitoring system tracks these signaling frames and maintains state for each active PDP-context. It is able to refer each arriving packet on Gn to the corresponding PDP-context and hence to the originating MS.

Some level of localization is possible by sniffing on Gb (for GPRS) and on IuPS (for UMTS), with different methods for the two due to different protocol specifications. It is always possible to discriminate exactly the current MS location at the level of Routing Area (a collection of several adjacent cells), for both GPRS and UMTS, also for those MSs that are presently attached to the network but not involved in active traffic exchange (note however they are still source of signaling traffic, e.g. Routing Area Updates). The finer-grain localization, i.e. per cell, is possible in GPRS only for active MSs. Therefore it is possible in GPRS to refer data packets (in downlink and uplink) to the current cell. For UMTS the per-cell localization of data packets is not complete: when a MS moves into a new RA, only the first visited cell is reported to the SGSN and can be therefore sniffed on IuPS. Subsequent cell changes within the RA are not “seen” on IuPS and therefore are not knowledgeable by our monitoring system. This aspect makes the task of revealing cell-level troubles (e.g. congestion) more challenging in UMTS than GPRS, since the cells within a RA are not individually and exactly observable, thus requiring additional inference techniques. Finally, based on the RA and/or local link information it is possible to associate the packets captured near the SGSNs (on GB/IuPS) to the current BSC/RNC, a basis for revealing troubles or other deviations at the level of these equipments.

III. Large-scale performance monitoring in 3G

Similarly to wired network, TCP is the dominant traffic component in the 3G network (more details on 3G traffic composition are given in [6]). Since TCP is closed-loop controlled, its behaviour depends on the conditions of the end-to-end path. This is true for both the microscopic (per connection) and macroscopic levels (per aggregate). Tools have been developed to estimate several TCP performance indicators (e.g.

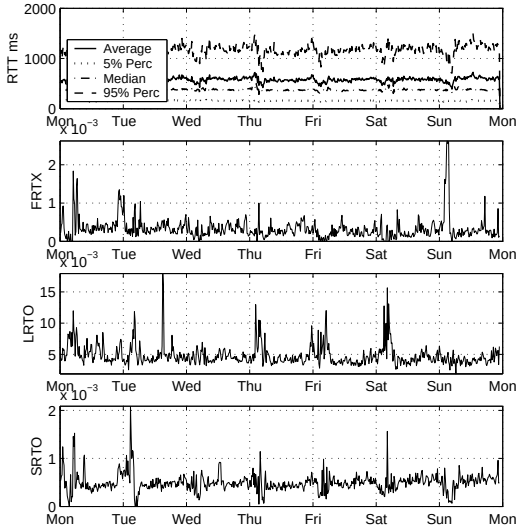

Fig. 2. Measured TCP indicators on UMTS (1 week, 15min bins).

RTT, retransmissions) by analysing the behaviour of individual connections (micro level) from passive traces (e.g. `tcptrace` [7]). Concepts were also proposed to extract indicators of anomalies at the macro level (e.g. [8]). Network-wide statistics of TCP indicators reassume the status of the entire network: averaged over long time-bins they can be used as synthetic global KPI (Key Performance Indicators), helpful for revealing macroscopic drifts in network-wide performances (e.g. due to slow changes of the global traffic distribution) or for measuring the actual improvement upon technological upgrades (e.g. HSPDA, EDGE). As an example, Figure 2 reports several TCP indicators measured in UMTS for one week in October 2005: the comparison with past measurements [9, Fig. 5] does not reveal any sensible performance shift after one year, despite the considerable increase in traffic volume. When coupled with routing information, these indicators can be used to infer the presence of performance degradation points distant from the monitoring point. This approach, a form of "network tomography" similar to [3], fits nicely to the 3G Core Network due to some key peculiar characteristics: tree-like structure, symmetry of paths, separation between clients (the MS) and servers (on the Gi side). This allows to monitor the actual quality and network performances at the packet-level for the whole CN, and to localize anomalies. Furthermore, the combination of these indicators with the per-MS location data extracted from Gb/IuPS traces (as explained in Sec. II) makes it possible to refer performance indicators to specific radio areas (cells or RA), so as to construct "performance maps" in space and time. From there, one can implement relatively simple mechanisms for detecting areas of persistent or recurrent performance degradation, directly spotting problematic areas. Some degree of statistical inference is required to overcome the lack of complete per-cell resolution in UMTS, where exact maps are available only at the RA level. The further correlation with "load maps" (e.g. number of active users and/or transferred volume per cell) would be helpful for discriminating between possible causes, for example disturbed radio coverage versus shortage of radio capacity, and then trigger local intervention (radio re-optimization). As shown in [9], the global statistics can be biased by a few MSs experiencing very poor transfer conditions. To avoid false alarms, caution is required to discriminate cases where the poor performances are accountable to MS-specific conditions rather than to area-wide status. The problem can be expressed in terms of statistic inference and hypothesis testing, an interesting aspect for research. Additionally, the matching with mobility data will give insight into the usage patterns and TCP performances of moving MSs during handovers, a natural continuation of the work in [10].

IV. Conclusions

It should be evident the revolutionary potential of the proposed approach with respect to the radio optimization process in 3G - at least to those having knowledge of the current practice. We remark that this approach does not require access to the RAN equipments but only to the Gb/IuPS near the SGSNs, usually co-located at few physical sites.

Our ongoing work is in advanced stage. So far, we have demonstrated the possibility to locate a bottleneck link within the Core Network from traces collected at few vantage points on Gn (see [8] [9]). The MS-tracking has been already implemented, separately for GPRS and UMTS, respectively on Gb and IuPS. The extension towards the RAN, i.e. per-cell TCP measurements, required some further enhancements to the analysis tool (a modified version of `tcptrace`) to handle cell information. In the associated poster we report preliminary measurements of TCP Round-Trip-Time (RTT) and Retransmission TimeOut (RTO) frequencies in individual GPRS cells, based on past traces. Our preliminary results indicate that in GPRS the number of MS concurrently active in the same cell is very low. This makes more challenging to discriminate performance degradation due to cell conditions from MS-specific problems. A possible approach would be then to correlate cell measurements across different times (days) and search for signs of "recurrent" performance degradation.

References

[1] J. Bannister, P. Mather, S. Coope. *Convergence Technologies for 3G Networks* . Wiley, 2004.
[2] R. Castro et al. Network tomography: Recent developments. *Journal of Statistical Science*, 19(3), 2004.
[3] V. N. Padmanabhan, L. Qiu, H. Wang. Server-based Inference of Internet Link Lossiness. *Proc. of IEEE Infocom 2003, San Francisco*, April 2003.
[4] C.C. Zou, W. Gong, D. Towsley. Code Red Worm Propagation Modeling and Analysis. *Proc. of CCS'02, Washington, USA.*, Nov 2002.
[5] *METAWIN and DARWIN projects. http://www.ftw.at/ftw/research/projects*.
[6] P. Svoboda et al. Composition of GPRS/UMTS traffi c : snapshots from a live network. *submitted to IPS-MOME*.
[7] Tcptrace 6.6.1, available at http://www.tcptrace.org.
[8] F. Ricciato, W. Fleischer. Bottleneck Detection via Aggregate Rate Analysis : A Real Case in a 3G Network. *accepted to NOMS'06.*
[9] F. Ricciato, F. Vacirca, M. Karner. Bottleneck Detection in UMTS via TCP Passive Monitoring : A Real Case. *Proc. of CoNEXT, Tolouse, France*, October 2005.
[10] F. Vacirca, T. Ziegler, E. Hasenleithner. TCP Spurious Timeout estimation in an operational GPRS/UMTS network. *accepted for Journal of Computer Networks*.