A novel experimental framework to investigate context-aware solutions for opportunistic M2M content delivery

Pavan Kumar*, Pietro Lungaro+, Vikrant Nandakumar* and Zary Segall*+

*Department of Computer Science, University of Maryland, Baltimore County
+KTH Mobile Service Lab, Kista, Stockholm

pavan4@umbc.edu, pietro@kth.se, vikrant1@umbc.edu, segall@kth.se

1 Background

In the last couple years we have been witnessing an unprecedented adoption of mobile data services. Key factors behind this success are both the development of powerful user devices such as smartphones and tablets, capable of delivering unprecedented levels of user experience, and flat rate pricing, which has made mobile Internet finally affordable to wider consumer segments. While on one hand the success of mobile data was somehow predictable, this was not the case for what concerns its rate of adoption and the associated traffic growth in mobile network.

Mobile operators expected a much gradual penetration and significantly milder traffic growth [1]. iPhone alone took within one and half years about 14% of the smartphones market share, while the Android platform is rapidly catching up. Moreover, each of this new smartphones is considered loading the network with approximately 60 times more data than previous “feature” phones. This combination has caught mobile operators off-guard, with under-dimensioned networks, in most of the cases leading to severe congestion problems [1].

Together with smartphones, a novel way of delivering and consuming mobile content has rapidly emerged: the apps paradigm. The success of the apps can be quantified in terms of both number of downloads from the various “app stores” and in terms of user face-time: recent surveys, e.g. [2], clearly show that apps have rapidly become the main channel for wireless content consumption. Similar to other publish/subscribe mechanisms, apps fetch information from their remote content servers via either push or pull mechanisms and use the mobile network for completing the content delivery to the user terminals. In many cases the transferred content is characterized by some degrees of tolerance to delay and many data transfers are performed as “background” operations.

With increasing numbers of apps per phone, and with smartphones gaining considerable shares of the device market, it is not difficult to understand the extent of the additional traffic volumes injected into the mobile networks.

In order to mitigate the effects of this novel “machine-to-machine” (M2M) type of traffic we proposed and investigated in [3]-[6] a set of context-aware solutions for opportunistic content delivery. The basic idea is to optimize content delivery, in terms of investments in terminal and network resources, by exploiting the times in which there is excess of resources available at the BSs. Since a large portion of the transferred content has some degrees of tolerance to delays, operators can opportunistically wait for users to roam closer to BSs and/or to less loaded cells and then perform a more efficient content delivery.

While apps can be considered as the present manifestation of rapidly growing M2M traffic trends, a large number of key players within the telco industry seem to be unanimous in foreseeing, for the next few years, the dramatic growth of sensors’ generated traffic in cellular systems. This can be seen as another type of M2M traffic, different from apps in terms of the amount of data generated by individual sensors and most likely stressing more the uplink side of the communication, e.g. from the sensors and/or data collection units towards data fusion centers [7]. Understanding the potential impact of M2M traffic on current network dimensioning and architectures is of paramount importance for the success and sustainability of these novel services in future wireless mobile networks.

2 Problem

With the aid of a novel experimental framework, this paper aims at quantifying the impact on network and terminal performances of both the aforementioned M2M traffic types. Developed within the COSEM project, the C2L living laboratory “test bed” allows you to remotely control a number of Android smartphones, injecting in real cellular network synthetic workloads and collecting results, on all monitored variables, from the deployed terminals. By mixing synthetic workloads representing human and M2M traffic loads, the C2L testbed allows you to anticipate the potential impact e.g. on content delivery delays and terminal energy costs for different penetration levels of apps’ and sensors’ traffic in future networks.

Moreover, by considering in the workload scheduling module of the testbed context information concerning both network and terminal contexts, this paper takes a step forward in quantifying the potential benefits of our proposed context-aware solutions [3]-[6].

3 The C2L Test bed

The testbed architecture consists of three separate components which interoperate with each other. Controller Gateway Servlet (CGS), Active Phone Database (APD) and the Results Database (RD) constitute the server elements. The CGS acts as a gateway for all incoming and outgoing information between the server and the mobile nodes. CGS handles the requests sent from/to the terminals

such as sending network statistics (push), getting workload information (pull) from the Results Database (RD). The APD is the snapshot of the current state of the mobile nodes. The RD contains the active workload deployment details along with the results obtained by running each one of them. Each mobile node has two components viz. Mobile Monitoring Service (MMS) and the Workload Generation Engine (WGE). The MMS is responsible to keep the APD updated with vital information about various parameters such as cell id, location, signal strength, battery capacity, battery voltage, battery drain percentage, network statistics, and application usage information etc. and the WGE is invoked by the user to start running workloads on the terminal. The entire system is exposed to the user through a web-based interface which has provisions to discover terminals, control them remotely, generate workload, display results, querying the Results Database (RD) and obtaining processed statistics.
The test bed works seamlessly and independently of any service provider or device manufacturer. The system can manage mobile nodes spanning across different remote locations and allows different combinations of heterogeneous workloads. These run in the background, without interrupting the user, while a workload scheduling feature additionally permits the users to deploy workloads at a future time instance. The system is secure with device level and user level role-based authentication and complexity scales efficiently on demand to accommodate a large number of mobile nodes.

4 Example of experimental results

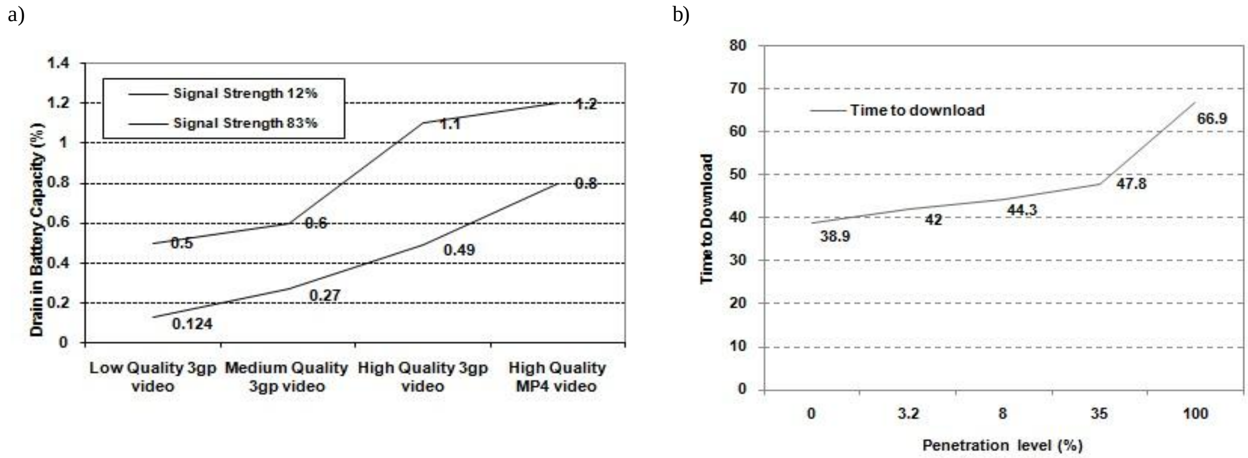

Fig. 1: Example of experimental results. In sub-figure a), the drain in battery capacity (in percentage) is shown for a content delivery (video) of increasing quality and file sizes, when the transmission is performed with terminals experiencing various signal strengths. In sub-figure b), the total duration of a video download is shown for different penetration levels of sensor traffic that is simultaneously served in the same cell. The penetration levels and sensor traffic composition are extracted from [7] and [8].

While in the final paper an extensive set of experimental results will be discussed and analyzed, in this section we focus on two specific study cases. In Fig.1a we show the energy costs associated with receiving video files of different qualities. The costs in terms of battery drainage can be significant, if signal strength is not considered when deliverying the content to the user terminals. By exploiting this context information, and waiting for user terminals to roam closer to the BSs, operators can reduce the overall terminal energy costs by 75%, for low quality videos, and 33% for high quality .
In Fig.1b we instead present a study case in which two terminals are co-located within the same cell. While the workload in the first terminal represents human generated traffic (the dowload of a small file), the second one injects in the network a syntethic workload corresponding to different levels of sensor traffic activity in the same cell. The penetration levels considered in the experiments are relative to the network dimensioning assumed in [7], while the sensor mix and traffic associated with each sensor type are based on the scenarios presented in [8]. The results show that with increasing sensor traffic, the performances experienced by the users progressively deteriorate, reaching almost 100% increased download duration when as many sensors as mobile users are assumed to be colocated within coverage of the same BS (100% penetration).

References

[1] Roben Farzad, "AT&T's iPhone Mess", article on Businessweek, 15th of February 2010, available online at: **http://www.businessweek.com/magazine/content/10_07/b4166034389519.htm**
[2] Zokem report on smartphone usage, 14th of February 2011. Available online at: **http://www.zokem.com/2011/02/gsma-partners-with-zokem-at-the-biggest-mobile-event-of-the-year-to-report-the-latest-in-mobile-usage/**
[3] P. Lungaro, Z. Segall and J. Zander, "ContextShift: a Model for Efficient Delivery of Content in Mobile Networks", in Proceedings of the Wireless Communications & Networking Conference (WCNC10), Sydney, April 2010.
[4] P. Lungaro, Z. Segall and J. Zander, "Predictive and Context-based Multimedia Content Delivery in Future Cellular Networks", in Proceedings of the Vehicular Technology Conference (VTC10), Taipei, May 2010.
[5] P. Lungaro, Z. Segall and J. Zander, "Context-aware Resource Schedulers for Opportunistic Content Delivery in Cellular Networks", in Proceedings of the 3rd international conference on Communication Theory, Reliability, and Quality of Service (CTRQ10), Athens, June 2010.
[6] P. Lungaro, Z. Segall and J. Zander, "ActiveCast - a network and user aware mobile content delivery system", in Proceedings of the 2nd international conference on Ubiquitus and Future Networks (ICUFN), June 2010, Jeju Island, Korea
[7] J. Krnic and S. Krco "Impact of WSN Applications' Generated Traffic", in Proceedings of PIMRC 2008.
[8] .J Marron, D. Minder and the Embedded WiSeNts Consortium, "Embedded WiSeNts IST-004400 – Application Scenarios", 2006-02-21.