When David helps Goliath: The Case for 3G OnLoading

Narseo Vallina-Rodriguez†, Vijay Erramilli‡, Yan Grunenberger‡, Laszlo Gyarmati‡
Nikolaos Laoutaris‡, Rade Stanojevic‡, Konstantina Papagiannaki‡
University of Cambridge, Cambridge, UK†
Telefonica Research, Barcelona (Spain)‡

ABSTRACT

Access link can often be the bottleneck for application performance. In this paper, we propose to augment wired connections using cellular ones, that we term "3G onloading (3GOL)". 3GOL utilizes available mobile devices and already-paid-for data volumes to augment and improve performance of applications on wired network. We motivate 3GOL by understanding bottlenecks present in the wired and the cellular networks. In order to understand the potential benefits of 3GOL, we conduct active experiments using mobile devices. We show that capacity gains can scale linearly with the number of devices on the downlink while also seeing improvements on the uplink. Using real traces we show how video on demand can benefit with 3GOL, even when volume caps are in place. We design 3GOL as an over the top service, and highlight research challenges.

Categories and Subject Descriptors

C.0 [**Computer Systems Organization**]: System architectures

General Terms

Design

Keywords

Cellular, Wired, Networks, 4G, 3G, Onloading, DSL

1. INTRODUCTION

The last few years have witnessed a large scale adoption of powerful, bandwidth hungry[1] smart phones that have posed serious provisioning challenges for carriers worldwide. In response to those challenges, a fair amount of research has gone into proposals for mobile data offloading [5, 10]—offload data *from* the resource constrained cellular medium *to* the (relatively) resource rich wired medium through femtocells and ubiquitous WiFi.

[1]Mobile data volume is projected to be 3.6 Exabytes per month [2].

Hotnets '12, October 29–30, 2012, Seattle, WA, USA.

Such cellular networks offloading strategies assume that the cellular network is constantly under strain while the wired network has abundant resources to assist. The reality is that the wired network itself has bottlenecks. The first and foremost bottleneck is the access link [12]. Moreover typical wired broadband comes with far less bandwidth in the uplink than the downlink (by design), essentially constraining the development and adoption of applications that source content from within the home. While a number of operators are in the process of deploying fiber in order to remove the last mile bottleneck, such upgrades are often lengthy, costly, and primarily targeted at locations where the operator expects to recover the cost, making the availability of such technology sparse. On the other hand, extension of the 3G network is an ongoing activity and specific regions are adopting LTE before they become target areas for higher speed broadband deployment [8].

In this work we make the following contributions:
1. We identify cases when the resource-constrained cellular network (David) can help the typically assumed resource-rich wired network (Goliath).
2. We propose 3G/4G Onloading (3GOL), as a way to improve performance for residential users, for those applications that are bottlenecked by the wired network. 3GOL is used to augment existing connections in the household by moving part of the traffic onto the cellular infrastructure.
3. Using 2 different data sets and active experiments, we quantify the potential for 3G Onloading to i) augment the primary broadband connection, ii) improve application performance and the resulting load on the cellular network, iii) operate within the limits of today's capped cellular data plans, while quantifying the resulting traffic "onloaded" to the cellular network.
4. We show that throughput augmentation could scale linearly with the number of 3G devices in the downlink, but will in all likelihood be limited by the 3G up-

link technology. Gains are significant reaching 3x in the downlink and 12x in the uplink even with a small number of 3G devices.
5. We sketch a system architecture that could enable cellular onloading both as a network provided service, as well as over the top.
6. We outline the research challenges that would need to be addressed in developing the end to end solution, and that would need to take into account cell tower scheduling policies, channel acquisition delays, increased cellular latency (as opposed to broadband link latency), as well as volume capped cellular data plans.

We conclude that 3G Onloading could be a useful technology for home users with slow broadband plans (either due to the network deployment, or choice).

2. 3G ONLOADING: THE CONTEXT

Consider user Alice who has a wired connection with 6.7Mbps downlink and 2.7Mbps uplink. These speeds are the average broadband speeds per connection reported by Netalyzr [9] and similar numbers are reported in [12]. Say Alice has (along with others in her household) three devices with mobile data access (3G or LTE), that can provide downlink speeds 2.0Mbps and uplink 1.5Mbps (shown later as the average throughput in our experiments). Using three devices, Alice can augment her downlink capacity by 6Mbps (nearly double) and uplink capacity by 4.55Mbps (nearly triple). If Alice has LTE, her downlink/uplink capacities can increase by approx. 36.75Mbps and 15.18Mbps [6] respectively, the augmented capacity is 6.44x and 6.62x the original speeds. Our work falls in the space of bandwidth aggregation schemes, such as [13, 1, 7], with a key difference being that we are not trying to use multiple connectivity options on the same device, but use cellular networks to augment home broadband connection.

2.1 The Network

Consider the different bottlenecks in the cellular network. First, cellular spectrum is a shared medium with a technology dependent capacity. The effective throughput that a user can receive from the cellular network will depend on the capabilities of his/her handset, the capabilities of the cell tower, the quality of the connection between the user and the cell tower, and the load on the tower. Furthermore, cellular base stations are typically connected with a 40-50 Mbps link to the Internet and that could become the bottleneck in high load situations[2]. But how pronounced can it be?

If we assume a neighborhood with a population density [3] of 35000 hab/km^2 and cells of radius 200m, we get approx. 4375 hab/cell. If we further assume 4 people per household, with 80% ADSL penetration, we get around 875 ADSL connections per cell. If the average downlink speed is 6.7Mbps, we get overall ADSL downlink capacity to be 5,863Mbps, compared to 42Mbps for state of the art cells – almost two orders of magnitude difference.

Consequently, 3GOL *cannot* assist all wired connections and at all times. It can, however, assist selected applications at *particular* times when they become constrained at the wired broadband connection. For well provisioned wired deployments (like fiber), 3GOL may never be required. For less well provisioned wired deployments, 3GOL is likely to be used occasionally and only for bandwidth hungry applications, like video. However, given the dramatic throughput increase provided even by a small number of devices, 3GOL could have a significant role to play in overcoming the throughput limitations of last mile access.

2.2 The Economics

Using the cellular network to speedup broadband bottlenecked applications could be significant given todays and future 3G/4G downlink and uplink speeds (shown also in the next section). However, cellular data plans typically come with data volume caps. For example, in the US, the capacity of the service plans ranges between 250MB to 12GB with monthly subscription fees from $15 up to $80 (AT&T, Verizon, Sprint). Service plans in other countries have similar characteristics both in terms of capacities and fees.

Once the cap is crossed, the operators apply two diverse policies: volume pay or throttling. In case of volume pay, operators charge users for each additional unit of volume used. Operators in the US use this mechanism predominantly. The unit of above-the-threshold usage ranges from 1MB through 250MB to 1GB with a fee in the range of $10 and $60 for every additional GB.

The second widely adopted scheme is throttling where operators decrease the connection speed of the user significantly (*e.g.*, by an order of magnitude) after crossing the cap. The total volume that a user can transmit is however unlimited in case of the throttling policy, for no additional fees. Throttling is used in the UK, Germany, and other European countries.

Given the cost and resource constraints of 3G/4G networks, such pricing schemes are unlikely to change significantly. 3GOL needs to be able to incorporate data caps in its formulation so as to exploit the already-paid-for volume, not to degrade the primary cellular data service performance or lead to extra cost for the user.

2.3 The Service

Finally, we look into what 3G Onloading would be as a service. We see two different ways for deployment: i) as a network integrated service, or ii) over the top. In the former case, a network operator could decide to

[2]Note that LTE, HSDPA and HSUPA are shared channels with TDM on top of UMTS. Proprietary algorithms allocate time slots to the connected users.
[3]http://www.demographia.com/db-citydenshist.htm

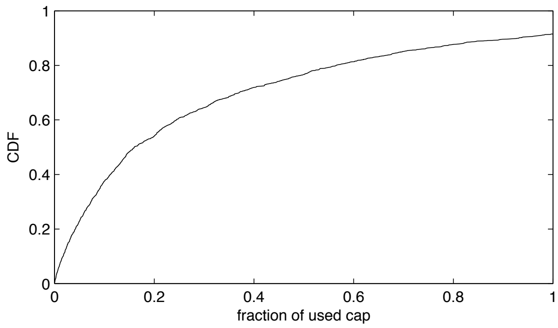


Figure 1: The distribution of fraction of used cap. 40% of customers use less than 10% of their cap. 75% of customers use less than 50% of the cap.

allow 3G Onloading to take effect i) only when the 3G network is lightly utilized, and ii) potentially not count the usage towards the data caps. In this kind of scenario, the problem of deciding when and for how long to use the 3G devices for assistance are designed into the service. In areas where one operator offers *both* broadband and mobile services, 3GOL could be used to encourage bundled purchase of both services. If unlimited use is a concern, then the operator could further enable 3GOL for fractions of the download, *e.g.*, until a video playout buffer is sufficiently filled for the first few MB (a la Comcast's Powerboost feature [3] that uses additional but unused capacity already available on the DSL link).

However, one can easily envision an application that could enable 3GOL over the top. In this second scenario, the cellular network can be used whenever the broadband connection imposes bottlenecks to application performance (that could also occur when the cellular network is highly utilized), but would need to do so within the constraints of the cellular data plan caps. Constraining the use of 3GOL within data caps also helps disjoint the service between the wired and mobile provider; the operators can be different and as long as 3GOL is transparent, the wired provider can benefit (by having traffic offloaded), the user can benefit by experiencing better performance and utilizing the already-paid-for volume.

To understand how cellular data plan caps can limit the efficacy of 3GOL, we analyzed data collected from a medium sized MVNO in Europe. The data set contains all voice, sms, and data accesses performed by its nearly 1 million customers. In Fig. 1 we show the distribution of data usage as a fraction of the contracted cap. Surprisingly, we find that 40% of the customers use less than 10% of their cap, and 75% of the customers use less than 50%. Therefore, even within capped cellular data plans 3GOL could be of interest.

Finally, an important question is which applications could make use of 3GOL. Given the channel acquisition delay in 3G (on the order of 2 sec [11]), and increased latency, 3GOL could provide benefit to applications that download/upload large amounts of volume and require no inter-activity. Bulk transfers and video applications downloading content that would require more than 2 seconds over the broadband connection are very viable targets for 3GOL. We again stress here that 3GOL may not be suitable for all networks and all scenarios.

3. FEASIBILITY OF 3GOL

In this section we present evidence that 3G Onloading can indeed meet the goals outlined in previous sections. To do this we answer 3 specific questions. 1) How much additional throughput could a residential user expect from 3G devices in his home?, 2) How much could that additional throughput help applications that are bottlenecked by the broadband network today?, and 3) What would that mean for the cellular network in terms of load?. Our results show that 3GOL could relieve application performance bottlenecks at the edge of the wired network, but needs to be done carefully so as to avoid overloading the cellular network.

3.1 3G Bandwidth Augmentation

The uplink/downlink rates obtained by mobile devices depend on multiple factors as discussed above. To glean the available bandwidth under real operating conditions, we programmed 10 high-end terminal devices (Galaxy S II with MIMO HSDPA Category 20 and HSUPA Category 6 with a 10GB mobile data plan) to download and upload 2MB files from/to a server, via the cellular interface. Measurements were performed in a major western European city in 5 different locations and at different times of day. One would expect that the degree of augmentation possible will be higher for rural areas and during the night. Our locations were selected accordingly to test this assumption.

Given that cellular networks have middleboxes [14] and dynamic IP address assignment, we used `wget` to estimate downlink bandwidth using a random 2MB file to avoid bias by proxies with compression capabilities. Uplink bandwidth was tested with `iperf`, although we also used `netperf` for validation and got similar results. Our experiments ran in cycles. We first activated one 3G device and performed uplink/downlink throughput and latency measurements. We repeated such measurements 4 times in sequence. Every 20 minutes we introduced a new device and ran the same measurements for all active devices in parallel. All devices were synchronized using NTP. The size of the file and number of repetitions were carefully selected so as to allow us collecting measurements in five locations without exceeding the data plan cap. For each experiment, we also recorded the base station serving each smartphone, as well as its signal strength.

Location	Time	DSL (d/u)	Cellular (d/u)	3GOL/DSL (d/u)
1. Densely populated residential area (city center)	1 a.m	3.44/0.30 Mbps	4.37/1.64 Mbps	2.27/6.46
2. Office area at rush hour	4 p.m	4.51/0.47 Mbps	4.56/5.18 Mbps	2.01/12.02
3. Residential area in tourist hotspot	10 p.m	6.72/0.84 Mbps	1.92/1.53 Mbps	1.28/2.82
4. Sparsely populated residential area (suburbs)	1 a.m	2.84/0.45 Mbps	4.67/3.89 Mbps	2.64/9.64
5. Popular shopping center in peak time	2 p.m	n/a	5.31/2.64 Mbps	n/a

Table 1: Description of the different locations used in the active experiments and comparison of the DSL and 3GOL (DSL + 3G) capacity when using 2 devices.

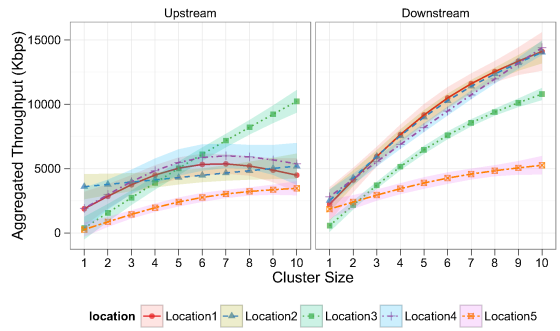


Figure 2: Aggregated throughput (uplink and downlink) for 10 devices

In Fig. 2 we plot the aggregate 3G throughput as a function of the number of active devices for the five locations in Table. 1. Downstream capacity can be augmented by up to 14Mbps in the downlink and 10Mbps in the uplink. In 4 out of 5 locations there was association with 2 different base stations. In the shopping mall, we found devices connecting to 6 different base stations, demonstrating a denser 3G deployment, with load being balanced across the network. Measurements in Location 2 are also taken during peak hours (4pm) but the cellular network seems to be much better provisioned, leading to a linear increase in downlink throughput up to 10 devices.

The behavior we observe on the uplink is significantly different. For 2 out of the 5 locations, we observe a clear plateau in aggregation at 5 devices, equal to nearly 5Mbps, which is the capacity for HSUPA (5.76Mbps). For the shopping mall location, even 10 devices are unable to receive the theoretical maximum, possibly due to congestion on the radio links. Interestingly, though, Location 3 does exceed 5Mbps, and reaching 10Mbps. For that particular location, all 10 devices are primarily using one base station. We believe that for this particular experiment, while all phones were connected to one base station, they were connected to different sectors, thus going beyond the HSUPA capacity per sector[4]. Such an assumption is plausible since Location 3 is a hub for tourists with a large density of cellular infrastructure.

Our results so far demonstrate significant benefits both uplink and downlink even with a small number of 3G devices. However, the majority of uplink throughput augmentation can be obtained with 2-3 devices, whereas downlink throughput seems to scale linearly with up to 10 devices. We also find that such benefits are pronounced in rural areas, but not negligible in urban areas. In fact, *when we contrast the throughput of 3GOL with that of broadband, we find a speedup of 2-5 for downlink, and 10-12 for uplink when using 10 devices, and a respective speedup of 1-3 downlink and 3-12 uplink even when using just 2 devices* (Table 1). These gains are further observed not only for rural but also urban areas, and even during peak hours.

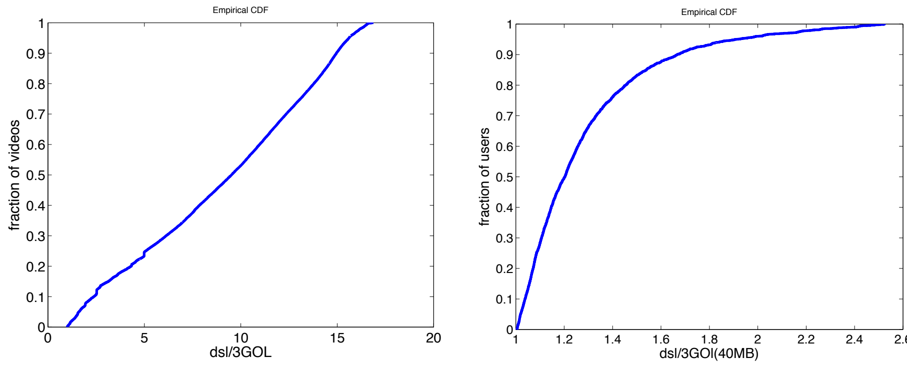
[4]Unfortunately, our devices can only report the base station they connect to, but not the sector.

3.2 Application Performance

Earlier we argued that video would be one of those applications that could benefit from 3GOL. In this section we quantify the performance improvement of video sessions using a session level trace collected from a DSLAM in a major metropolitan city in the same country over 24 hours in April 2011. The represented geographical area would typically be covered with 8-10 cell towers.

The data contains HTTP transactions, (userid, time of request (UTC), URL) from which we separate out all sessions pertaining to video. The data contains the HTTP response headers, from where we can obtain the size of the video files requested. We note that users may not watch the entire video that is requested, however our choice of using the size of the entire file is *conservative* as we shall demonstrate. We observe videos from primarily 4 sources: Youtube, Megavideo, streaming from the premier TV channel in the country, and the rest (Vimeo, Veoh etc.).

Our objective is simple: we want to understand if 3GOL can help in reducing the latency for videos. As we do not have transport level information, we use download times as a proxy for latency. Towards this end, we look into the following scenarios. The first is the latency observed while using a DSL connection of 3Mbps (average speed in our active experiments). We then use 3GOL, with 2 devices (median speed 4.8Mbps aggregated over two devices) plus the 3Mbps of DSL and obtain the latency to download the same video. This is the case where there is no metering of the data volume consumed via 3GOL. In Fig. 3(a), we plot the speedups (defined as $\frac{DSL\ latency}{3GOL\ latency\ 3GOL}$) observed over using 3Mbps. We note that if 3GOL is used without metering, there can be substantial improvements—50% of the videos have speedup by a factor of 10 and below. We note here



(a) (b)

Figure 3: (a) CDFs of improvement of 3GOL(no metering) per video (b) Per user improvements in terms of latency

that as there is a 2 secs delay in obtaining the cellular channel [11], we consider the augmentation effect of 3GOL always *after* the first two seconds.

When 3GOL has to deal with budgets, we study the case of a 'Powerboost'-like acceleration of the video transfer. In our datasets, we find that an end user [5] views 14.12 videos (mean, 6 median, 30.13 std) per day. Therefore, aiming to accelerate *every* video by even 20MB would lead to a daily consumption of 280MB for the 3G data plan, thus being clearly unacceptable. We, thus focus on using 20MB per device per day. With two devices, we can use 40MB per day, using 3GOL, using this 40MB over any/all videos an end-user views. We ask: what can the benefits be *per user*. We plot the savings *per-user* over using just DSL ($\frac{DSL\ latency}{3GOL\ latency\ under\ budgets}$) in Fig. 3(b). We find that 40% of the users can see around 20% speedup, while 5% of the users can see a speedup of 2. We note here that our results are conservative. Video does not need the whole file to be downloaded; it only requires enough to fill the playout buffer for the video to start playing. There is no fixed size of the playout buffer but the general consensus is that it is an order of magnitude smaller than the file size. If the average size of a video is 50MB [4], and the playout buffer is 5MB, then 3GOL can speed up 8 videos (5MB total per video) per day per user—a tangible benefit. From Sec. 3.1, we saw an increase in capacity with more devices, leading to even higher gains. Likewise, with LTE deployed, we can expect much higher benefit to the enduser, *even* with budgets in place.

3.3 Impact on 3G Network

So far, we have shown that 3GOL can increase performance for applications bottlenecked by the last mile network. But at what cost for the 3G network? Using the packet trace, we compute the amount of traffic that would be "onloaded" to the 3G network if we were to accelerate the *first* video that could benefit from 3GOL (with a size greater than 750KB, that would require more than 2 seconds on DSL). We further assume we have 2 3G devices (with HSPA+) that can be used to accelerate up to 20MB each per day.

[5] a DSL 'subscriber'

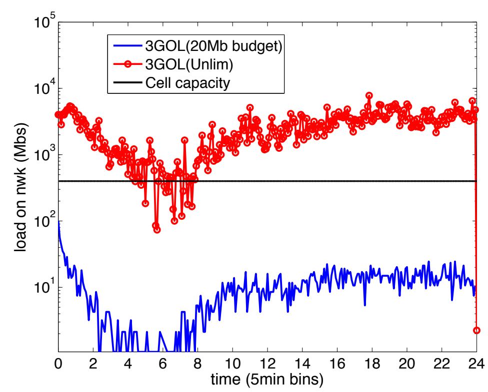


Figure 4: Load on cellular network by using 3GOL (budgeted) and 3GOL (unbudgeted). Y-Axis is logscale. Solid horizontal line is backhaul capacity (40Mbps)

In Fig. 4 we show the traffic onloaded to the cellular network and compare it with the backhaul capacity of the cellular deployment that could cover a similar geographical area (10 base stations with 40Mbps backhauls). We note that if 3GOL would operate without caps, then the 3G network will be guarenteed to be overloaded. However, when 3GOL operates within caps then the additional load introduced on the 3G network could be reasonable. Our active experiments demonstrate that 10Mbps could be available in the cellular backhaul across geographical regions and at different times of day. Note that our results are conservative since we accelerate the entire video transfer, whereas prior work has demonstrated that only 10% of the whole file is downloaded [4]. However, the actual overhead on the network can only be quantified post deployment.

4. SYSTEM SKETCH AND CHALLENGES

In order to develop an architecture for 3GOL, the following factors that are important for deployability: i) the underlying mobile platform can often times be closed (*e.g.*, Apple's iOS), limiting development of ap-

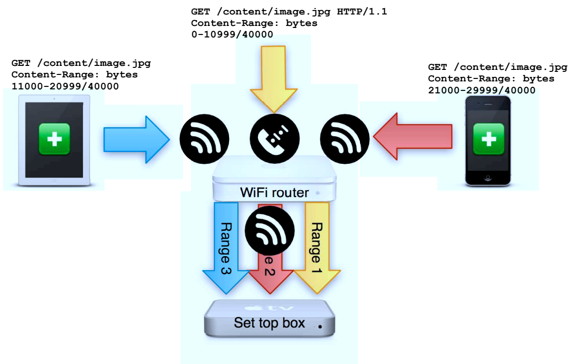


Figure 5: 3GOL architecture with 2 devices and a compliant set top box.

plications that can use the lower layers of the software/network stack, and ii) operations need to be transparent. Given that most traffic today is HTTP, and to meet the aforementioned requirements, we propose to implement 3GOL as an HTTP proxy. We exploit the fact that i) DSL/cable/fiber routers in home today are fitted with WiFi, and ii) most 3G devices today support tethering. The high level architecture is shown in Fig. 5.

We introduce two components: the first one is running on the mobile devices and handles the following: i) detects the available cellular network capacity and advertises it on the WiFi network through a discovery protocol like Bonjour, ii) waits for incoming connections on a TCP port, and iii) pipes incoming connections to a remote host connection established over the 3G network. A second component is i) running a HTTP proxy as a background service, ii) discovers mobile devices present on the WiFi home network and waits for HTTP connections from the applications to be augmented, and iii) communicates with the devices and retrieves the different pieces of content. Specific properties of the requested payload are exploited to parallelize the downloads. The properties can be based on the shape of the content (chunk of video, such in HTTP Live Streaming) or based on the HTTP protocol itself (like HTTP 1.1 Content-range parameters). Note that this makes 3GOL *application agnostic*.

A full implementation of 3GOL is part of our future work. The fundamental challenge would be in the design of an efficient algorithm that would i) decide when to use 3GOL (for which applications and at what point in time), ii) select the 3G devices to assist, and iii) for how long. Such an algorithm will need to take into account i) the channel acquisition delay in 3G, ii) the fact that 3G throughput is difficult to predict in advance, and iii) the different volume caps and level of consumption of different devices, with the goal of being strictly below the cap. Our preliminary work towards that direction indicates that it would be entirely feasible to predict device data consumption across the month and utilize their data plans for 3GOL without exceeding their caps.

5. CONCLUSIONS

In this paper, we have proposed 3G Onloading, a strategy that uses the cellular network to improve application performance in the home, when constrained by the wired network. Our contribution is in studying the feasibility of 3G Onloading and deriving requirements for an end to end system. Using 3 different data sets, and active experiments in urban and rural locations, we have demonstrated that the benefits to application performance could be significant with reasonable overhead to the cellular infrastructure. Such overhead could further be minimal if 3GOL was rolled out as a network integrated service, offered only when the cellular infrastructure is lightly utilized.

Acknowledgements

This work is partly sponsored by FIGARO Consortium funded by the European Comission (project FP7-ICT-258378). Narseo Vallina-Rodriguez is partially funded by a Qualcomm Innovation Fellowship. The authors would like to thank the anonymous reviewers and our shepherd Sylvia Ratnasamy for constructive feedback.

6. REFERENCES

[1] Anirudh Badam, Michael Kaminsky, Dongsu Han, Konstantina Papagiannaki, David G. Andersen, and Srinivasan Seshan. The hare and the tortoise: taming wireless losses by exploiting wired reliability. In *MobiHoc*, 2011.

[2] Cisco. Cisco Mobile Data Traffic Forecast. `http://www.cisco.com`.

[3] Comcast. Comcast Powerboost details, 2012. `http://www.dslreports.com/faq/14520`.

[4] Alessandro Finamore, Marco Mellia, Maurizio M. Munafò, Ruben Torres, and Sanjay G. Rao. Youtube everywhere: impact of device and infrastructure synergies on user experience. In *ACM IMC 2011*.

[5] S. Ha, S. Sen, C. Joe-Won, Y. Im, and M. Chiang. Tube: Time-dependent pricing for mobile data,. In *ACM Sigcmm*, 2012 (to appear).

[6] Junxian Huang, Feng Qian, Alexandre Gerber, Zhuoqing Morley Mao, Subhabrata Sen, and Oliver Spatscheck. A close examination of performance and power characteristics of 4g lte networks. In *ACM MobiSys*, 2012.

[7] Pan Hui, Richard Mortier, Kuang Xu, Jon Crowcroft, and Victor O. K. Li. Sharing airtime with shair avoids wasting time and money. In *HotMobile*, 2009.

[8] idate. FTTH vs LTE. `http://www.idate.org/en/News/FTTH-vs-LTE_747.html`.

[9] Christian Kreibich, Nicholas Weaver, Boris Nechaev, and Vern Paxson. Netalyzr: illuminating the edge network. In *ACM IMC 2010*.

[10] Kyunghan Lee, Joohyun Lee, Yung Yi, Injong Rhee, and Song Chong. Mobile data offloading: how much can wifi deliver? In *ACM Conext 2010*.

[11] Feng Qian, Zhaoguang Wang, Alexandre Gerber, Zhuoqing Morley Mao, Subhabrata Sen, and Oliver Spatscheck. Characterizing radio resource allocation for 3g networks. In *ACM IMC 2010*.

[12] Srikanth Sundaresan, Walter de Donato, Nick Feamster, Renata Teixeira, Sam Crawford, and Antonio Pescapè. Broadband internet performance: a view from the gateway. *ACM SIGCOMM 2011*.

[13] Cheng-Lin Tsao and Raghupathy Sivakumar. On effectively exploiting multiple wireless interfaces in mobile hosts. In *ACM Conext 2009*.

[14] Zhaoguang Wang, Zhiyun Qian, Qiang Xu, Zhuoqing Mao, and Ming Zhang. An untold story of middleboxes in cellular networks. SIGCOMM '11.