Consolidated Review of

Understanding the Super-sized Traffic of the Super Bowl

1. Strengths

The studied dataset is fairly recent, incorporates LTE network aspects, has some very interesting outage events (power cut in the middle of the game) that all shed insights into how users react and operate. This additional information sets the paper apart from the prior work in [13].

The dataset investigated in the paper is very interesting and is likely representative of the type of data that is more likely to be observed in the future. The paper has some interesting insights about the data.

- video traffic is the largest share, and the large portion is live streaming of the game itself (!)
- caching is unlikely to help for live streaming, because users seldom stream the video simultaneously
- surprisingly upload traffic is about the same as download traffic since users upload photos and videos to social networking sites. This says that subscribers need to do some provisioning for upload traffic
- built-in delay tolerance can help mitigate the congestion caused by upload traffic
- it makes a case for re-positioning the antenna in different locations according to the time;

Well written.

2. Weaknesses

My main comment is that it doesn't do a good job of comparing with previous work. The previous sigmetrics work on sporting events also analysis performance and throughput at such events, and discusses why the performance is poor (they found that it was because of the difficulty of resource allocation at the control plane of 3G). In this paper, the authors claim that their analysis is different because they are looking at LTE and not 3G. But there is no discussion on what is fundamentally different between 3G and LTE and how that affects their analysis. Is it the case that there are no performance bottlenecks in LTE? Is LTE the reason for the increased amount of video traffic? The bandwidth may still be bottleneck-ed by the backhaul, correct? If there is bottleneck, what is the reason? there is no performance analysis of the networks, so I wasn't able to get much in terms of how the landscape has changed with the introduction of LTE.

The paper could have shed some light on how the LTE network reacted to these large-scale events, in terms of congestion levels, signaling load, number of devices (and types of devices) that connected. While one can appreciate the issue of user privacy and the corporate need for the authors to not reveal how this cellular network performed under high load, not supplying even coarse-grained information seems very unsatisfying. Perhaps, the authors are reserving this ammunition for another paper?

The lack of access to the Wi-Fi data may make some of the results about the cellular data less definitive (e.g., maybe the Wi-Fi was quite bad for some users so they turned to cellular, but was good for other users, so they didn't). It may be hard to generalize the paper's insights to other datasets.

3. Comments

I found this paper to be interesting. There have been few studies of on-site traffic and characteristics at large sporting events and this paper adds valuable new material to this under-studied space. That said the work is a small delta compared to prior work. E.g., [13] also studies a sporting event. The authors claim that the difference is that they study LTE where the prior study focused on 3G. It's not clear what aspects of the setting change due to this and how these changes drive new findings. To address this comment, the authors would need to spend more time outlining exactly what [13] has, how the current observations are different, and why? Unlike [13] this paper does not show any performance issues. Which made me wonder if none arose, or they simply were not examined; I would be surprised if it were the former, and would have loved to see an analysis of performance issues users faced. My final comment is about the mitigation techniques: Consider the idea of delaying traffic - while it could lead to outcomes that are "globally optimal" it could leave individual users unhappy. Also it is not clear users would want to wait for downloads or uploads, e.g., downloading a play to watch a replay. Waiting on the order of a few minutes, as you suggest, just seems to be not very useful in this context. SO while your solution to delay traffic is interesting, I'm not sure it is practical.

The key insights from this paper are in line with what one would expect from prior studies (e.g. [13]) but there is sufficient new contributions as well to merit acceptance. The traffic volume graphs that suggest peaking activity during the game outage as well as before the game is interesting. Specifically, the traffic peaks relating to cloud sync activity is something new. Is this related to how the users/device-makers configure their sync settings on their smartphones? Surely these settings can be adapted to react to network conditions - this is where it would have been useful for the authors to show network congestion levels (to infer if such cues would have helped these apps delay their uploads). The fact that just 6 users occupied a large fraction of the bandwidth is not surprising, given that the system was configured as a large DAS for coverage purposes. What was the maximum throughput of the network observed? The authors present only normalized numbers. What was the throughput distribution across the cells? The highly variable nature of the traffic and the authors' suggestion to apply delay to uplink traffic is a point well-taken. However, this is not new, and one can easily imagine this to be true given the nature of this event. The paper does not shed light on how the Wi-Fi networks behave at the same time in the same venue. This is just an observation and is not a weakness of the paper. It would have been nice to correlate this information with cellular network performance, but that might be asking for too much given that this is a 6-page paper.

It will be useful to separate smartphone traffic from Tablets, to get a sense of how people use the network at such events. what about https traffic? what percentage was https traffic and how does that affect your analysis (for example, you cant look at content sharing in the https case) You talk about using IP address DNS hostname to identify photo and document synchronization, but is that complete. Can you identify all of those apps using your technique? Figure 2 has 3 legends but only a single line Table 1, 56.5% was http traffic, but it need not be web browsing. almost all apps use http

- While a lot of traffic is transmitted over the cellular network, it would be interesting to see how much traffic belongs to each user. In particular, are small subsets of users responsible for a large portion of traffic?
- Something that perhaps cannot be captured in the data is that when a user streams a video or downloads information, the user may share that information verbally with friends/neighbors with whom the users is sitting, or show the video to those friends/neighbors.
- The number of simultaneous users at any given time was about 1200 out of maybe 50,000, and a maximum of 26 users at any time were simultaneously downloading video. In some sense then the proposed system design considerations are for when there are not that many simultaneous users of the system (out of the total potential users). How would the design considerations change if there were more users? Do the authors expect that in the future that there will be more simultaneous users? Or is the percentage observed typical?
- The authors analyze the traffic from the provider perspective, but it would be interesting to understand how the users experience: did they find the cellular network performance sufficient? Would they have been willing to tolerate delay? Were they actually using the Wi-Fi much more?
- In Section 5, the authors say that "Constraint 2 ensures that all data was delivered within the deadline." Where do the deadlines come from? How do they get set?

4. Summary from PC Discussion

The paper was accepted without discussion.

5. Authors' Response

To address the comments of the reviews about the comparison to previous work, we have updated Section 2 with more detailed comparisons. Specifically, for [13] we have added more discussion about the fundamental differences between the 3G network studied previously and the LTE networks studied here. However, a key difference of [13] with our work is that we have not focused our analysis on the RAN performance bottlenecks in this case.

Regarding the question about providing some performance analysis of the event, while we agree this would be highly desirable, it was not the focus of our analysis in this case. In addition, the detailed performance metrics, such as network congestion, cannot be readily measured from the passively collected data set we had available. For instance, some of the performance issues users faced which are application behavior (e.g., stalls) cannot easily be measured by passive measurements in the network.

For HTTPS traffic, in Table 1 we show the percentage of HTTPS traffic in the data analyzed. HTTPS has minimal impact to our analysis as we can map IP addresses back to the original content provider using DNS hostname. The HTTP traffic does not have to be just web browsing and we will update the paper to reflect that.