[3] Adaptive Bitrate Protocols. http://en.wikipedia.org/wiki/Adaptive_bit_rate, 2011.

[4] AT&T SXSW Press Release. www.att.com/Common/docs/SXSW_Network%20Fact_Sheet.doc, March 2011.

[5] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2010Ú2015. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html, February 2011.

[6] FFmpeg. http://www.ffmpeg.org/, 2011.

[7] HTTP Live Streaming. http://tools.ietf.org/html/draft-pantos-http-live-streaming-06, 2011.

[8] iOS Developer Library. http://developer.apple.com/library/ios/#documenation/networkinginternet/conceptual/streamingmediaguide/UsingHTTPLiveStreaming/UsingHTTPLiveStreaming.html, 2011.

[9] S. Akhshabi, A. Begen, and C. Dovrolis. An Experimental Evaluation of Rate-Adaptation Algorithms in Adaptive Streaming over HTTP. In *MMSys'11*, San Jose, USA, 2011.

[10] J. Chesterfield, R. Charkravorty, J. Crowcroft, P. Radriguez, and S. Banerjee. Experiences with multimedia streaming over 2.5G and 3G Network. In *BroadNets'04*, Chicago, USA, 2004.

[11] J. Erman, A. Gerber, M. T. Hajiaghayi, D. Pei, S. Sen, and O. Spatscheck. To Cache or not to Cache: The 3G case. In *IEEE Internet Computing*, 2011.

[12] J. Erman, A. Gerber, M. T. Hajiaghayi, D. Pei, and O. Spatscheck. Network-Aware Forward Caching. In *WWW'09*, Madrid, Spain, 2009.

[13] H. Falaki, D. Lymberopoulos, R. Mahajan, S. Kandula, and D. Estrin. A first look at traffic on smartphones. In *Proc. IMC*, 2010.

[14] H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin. Diversity in Smartphone Usage. In *MobiSys'10*, San Francisco, USA, 2010.

[15] A. Gember, A. Anand, and A. Akella. A Comparative Study of Handheld and Non-Handheld Traffic in Campus WiFi Networks. In *Proc. Passive and Active Measurement*, 2011.

[16] J. Huang, Q. Xu, B. Tiwana, Z. M. Mao, M. Zhang, and P. Bahl. Anatomizing Application Performance Differences on Smartphones. In *MobiSys '10*, pages 165–178, New York, NY, USA, 2010. ACM.

[17] G. Maier, A. Feldmann, V. Paxson, and M. Allman. On Dominant Characteristics of Residential Broadband Internet Traffic. In *IMC'09*, Chicago, USA, 2009.

[18] G. Maier, F. Schneider, and A. Feldmann. A First Look at Mobile Hand-held Device Traffic. In *PAM'10*, pages 161–170, Berlin, Heidelberg, 2010. Springer-Verlag.

[19] Y. J. Won, B.-C. Park, S.-C. Hong, K. B. Jung, H.-T. Ju, and J. W. Hong. Measurement Analysis of Mobile Data Networks. In *PAM'07*, pages 223–227, Berlin, Heidelberg, 2007. Springer-Verlag.

Summary Review Documentation for

"Over The Top Video: the Gorilla in Cellular Networks"

Authors: J. Erman, A. Gerber, S. Sen, O. Spatscheck, K. Ramakrishnan

Reviewer #1

Strengths: The authors collect data from the link connecting the SGSN and GGSN nodes in a large cellular network, and on two different locations. They analyze it to identify the videos, the protocol they are relayed on, the resolution, as well as cacheability. Very nice study.

Weaknesses: The paper seems to be rushed at times. One figure is repeated the same in two places (5 and 8), the authors never introduce the content providers and make mention to them without and explanation. But these could be fixed in the final version.

Comments to Authors: This is a very nice short paper. The authors use a new source of data, carefully collected for 24 and 48 hours in two different locations in a large cellular network. They collect flow information and packet information (the first 20KB), thus being able to also look into the streaming bit rate of the video, the type of video, its resolution, information that is not easily accessible. Moreover, they are able to see the protocol that carries the video in question and study its properties.

Interesting findings are:
- majority of video is carried over HLS
- the transmission of the video is done over varying bit rates that can be observed through the trace, indicating the actual need for streaming rate adaptation to wireless conditions
- most content today is streamed at rather low streaming rates
- 40% of the videos never complete
- given that the majority of video traffic is transported over HTTP, the question of cacheability is asked and the authors simulate the actual gains one could expect from such a solution by replaying the collected traces.

One question I had is whether the authors know if there are any content acceleration middleboxes in the path observed. That would be essential to understand since then your infrastructure monitors the behavior of the middlebox and not the origin video server.

More detailed comments:
- Section 2.1: I assume that the capacity for data collection is 9 TB and not 3 (otherwise text disagrees with Table1)
- Figure 3 says that it displays normalized volume, but in that case I would expect the x-axis to be between 0 and 1 or 0 and 100%. In your final version, you need to describe the metrics displayed in the figures more carefully.
- The paragraph below Figure 5 points to Figure 4, but that figure does not contain the information discussed.
- Video providers are discussed in page 4 without any prior reference and only appear in the next page. You need to reorganize the paper for better exposition.
- I was actually surprised that the top resolution covers only 3% of the videos. How many resolutions do you see being used in total?

All in all, I think this is a very decent attempt for this first study of its kind. The writing is rushed but I think the content is there and it's interesting. I would be in favor of acceptance.

Reviewer #2

Strengths: The paper presents characteristics of a large volume of video traffic on a cellular network. There are some interesting data points, e.g., the distribution of different types of video streaming methods.

Weaknesses: The data seems not well analyzed, even for a short paper and some of the results are too anecdotal. The paper is poorly written in parts.

Comments to Authors: Video streaming over cellular networks is definitely a growing part of traffic, and this analysis from a large cellular network is certainly of great interest. However, I found the paper to have inadequately analyzed, what certainly is a very rich dataset.

As an example, Figure 2 shows an anecdotal download using the PD-byterange method. The results are strange and unexplained in the paper. Does this kind of download happen a lot, or is it just a rare occurrence?

Fig 5 and Fig 8 appear to be identical.

Some of the results presented in the paper, i.e., choice of video adaptation rates used by different applications, are not really actionable. The only part of the paper that I personally found particularly useful was the cacheability discussion.

Overall, this seems like a large dataset that requires more interesting analysis and the authors probably have barely scratched the surface with this work.

Reviewer #3

Strengths: Important topic.

Weaknesses: This reader is somewhat concerned with the representativeness of the results from this study, due to various indications shown in the paper. For example, Section 2.3 stated that video objects were replayed into the popular ffmpeg tool to extract out the needed information, and "In our data set ffmpeg was able to parse 45% of the video objects".

Another concern is the rather different ratio of uncacheable data between PD and HLS (8% vs 78%), which was stated but not clearly explained. The paper explained this is largely due to encryption of HLP. But how should one view this result together with the Section 2.3 statement in the above ("was able to parse 45% of the video objects")?

Section 4 stated that " the impact of video caching on the RAN is limited". Is this referring to the above result, or something else?

Comments to Authors: None.

Reviewer #4

Strengths: This is a classic network measurement paper. The measurements are strong, interesting, and provide some new insights.

Weaknesses: The paper tends, at times, to feel like a data dump.

At times some of the underlying issues of data quality seem a little brushed under the carpet. For instance, how do the authors account for a possible bias caused by some brands of devices more represented than others in their dataset, in comparison to other wireless networks?

Comments to Authors: The paper seems to be along the same lines as "P2P, the Gorilla in the Cable (2003)" by some of the same authors. Although the topic is actually different, the similarity of titles leads one to want to know the relationship between the two papers.

There are a few statements that are just a little off. For instance, the two datasets were collected at the same time, but one is longer than the other. I guess it is just the same start time? The data collector had 3 TB storage, but the traces where both over 8 TB? Please clarify.

There is a general statement about the number of subscribers covered in the study, but it isn't made clear if this is a total of which we see a sample (in each area), or if this is the number sampled.

Why does Figure 9 go above 100%?

There are quite a few small bugs and typos in the paper.

Reviewer #5

Strengths: - Novelty: new aspect of mobile cellular network usage captured for the first time, with some important new insights (on the amount of incomplete videos, and the relative spread of their popularity).
- Level of detail: practitioners and people working on detailed protocols of video encoding could find interesting bits.
that there are a couple of anecdotal evidences it is nice to monitor how YouTube works in a systematic fashion.

Weaknesses: - Not much, some limitations could be mentioned a bit more clearly, and naturally it would be awesome to understand spatial property as well, but there is so much one can do in 6 pages.

- The authors could have contrasted their findings with wire networks to emphasize the new characteristics found in this context.

Comments to Authors: This paper is I think a no-brainer accept. It is not revolutionary but it provides detailed insights into video traffic on cellular network, by using collected measurements and presenting a very clear set of numbers. Among them one that stands out particularly is that a large part of the videos are not completed (although it is only seen on a part of the data set), and that caching is not going to reduce the traffic much, although it will be effective for the most popular videos.

This is well done and contains quite clear description of video encoding characteristics. Seems a great match for a short paper.

A few points to improve:

1. I think that it is important to put some of the results in perspective of your limitation. In particular, the fact that only PD videos are examined from the completion standpoint is (1) only establishing this result on a part of the data set, (2) may be biased (some of these videos may take more time to load as they retrieve from multiple sources and needs reassembling, and hence may create too much impatience in the users). This point ought to be clarified.

2. When you mentioned caching, it could also be the case that caching improves delay, which seems to be an important thing given the relative fraction of uncompleted download indicating that users are affected by delay. It would be nice to see that point discussed.

3. I am surprised not to see any discussion of previous findings of videos on other networks. I imagine that the rate could be quite different (although it depends on which year we consider). But would the popularity be the same? Would the amount of abandons and video lengths be different?

4. The restriction to remove ad-video seems appropriate, but it seems also a bit incomplete (you mention you remove another domain, but isnt there more like these ones)? Could you provide some numbers (based on duration) to justify that no more are present?

This seems in particular important as you indicate later that the most popular videos are ads.

Response from the Authors

We thank the reviewers for their constructive comments.

We have addressed all the minor issues in the paper and made the appropriate adjustments (e.g. we have removed the duplicate figure pointed out in the review).

While we have done a careful example of the sample video behaviors in Section 2, this was more to illustrate the behavior than to come to our conclusions based on them. For example, we have observed that HLS sessions, in general, show their ability to adapt to changing conditions. Figure 1 is an example to illustrate this more carefully. The issue of duplicate downloads, as

highlighted in Figure 2 is not an exception, but a common behavior across multiple applications that we have observed on multiple platforms (device OSs). We have explained this anomalous behavior better in Section 2, and also described how the behavior can be reproduced. To further establish that this is not just based on the observation of a few example situations, Figure 8 shows that for a significant percentage of videos more than 100% of the content is downloaded by the end-point. We believe that this adequately addresses the issue brought up by the fourth reviewer.

In addition to providing a novel characterization of video traffic on 3G networks, we also make useful observations about how a carrier can appropriately handle this class of traffic. This study is also important for application developers to understand their impact on the network and to more optimally use the cellular network.

This study and observations are based on traffic generated by millions of users in a large tier-1 network. The large footprint of the data set gives us a lot of confidence about the generality of our main findings. While there may be some idiosyncrasies e.g., the device mix studied, we have looked at other platforms with different devices and OS mixes and many of these conclusions still hold. Finally, we have also reminded readers that the results naturally may not be completely representative of the video traffic in every cellular network, since the combination of wireless devices, the types of content providers and the behavior of users might be different.

We leave the comparison of our results on cellular networks with video characteristics on wireline network for future work. Thanks for the suggestion