

Response to Review Comments

Desktop and Mobile Web Page Comparison: Characteristics, Trends, and Implications

Original Manuscript Number: COMMAG-13-00421

T. Johnson and P. Seeling

January 2014

We would like to initially thank the editor and the reviewers for their very timely and in-depth review of our original manuscript. In the revision that accompanies this response document, we have carefully considered the suggestions and comments provided and believe that the revised version adds to the readability for the COMMAG audience, thanks to the feedback provided by the reviewers. For convenience, we have included the reviewers' comments here in italics followed by our response.

Guest Editor

Comments to the Author:

Dear Author, the paper is really an interesting one, but as you can also see from the comments of the three reviewers it is not really ready to be published in IEEE Communication Magazine for which we require first class papers. However if you read carefully read and consider all the great feedback that the reviewers have provided I do believe it will become a great paper for the readers.

We are thankful for the opportunity to improve this manuscript based on the thorough review comments and the overall encouraging feedback. The revised version takes the reviewer feedback into account as additionally outlined in this response document. We added more data "lines" for web pages that were present throughout the evaluation period and additionally refreshed the underlying dataset up to the end of 2013 to maintain currency for the readers. We could not consider some reviewer suggestions, such as extended evaluations of additional characteristics due to the tight space constraints for IEEE Communications Magazine manuscripts. The space that would be required to perform these additional evaluations would require an additional manuscript dedicated to the extended set of characteristics.

Reviewer 1

Comments to the Author:

typo: s/lesser extend/lesser extent/

Thank you for finding this detail, we have corrected this mistake.

Reviewer 2

Comments to the Author:

The topic of the paper, evaluating the trends of desktop and mobile web pages, is interesting. Especially the mobile dimension should make this study different from prior longitudinal studies. (I have not checked very carefully the prior studies so I don't really know how much they speak about the mobile aspects). However, there are several weak points in the present manuscript:

We are grateful for the detailed feedback comments provided by the reviewer and their provision in such a timely review process.

1. The paper claims that it is more up-to-date than prior studies. While that is probably true by looking at the dates of the sample it would have been important to compare how the results of this study differ from the prior studies e.g. from [4] or [5]. Such a comparison would have provided evidence to the authors statement that things are now different.

2. The major problem I find with the study methodology is that by blindly using the data in httparchive.org the authors draw conclusions which are probably wrong. As the authors observe the httparchive.org dataset had two major changes during the study period as the criteria for page selection were changed. Therefore I don't think it is fair to calculate the average growth by just taking the difference between the measurements in the first and last study dates and conclude that it shows the average growth rate. I suggest that the authors would (in addition?) take a smaller subset of pages that have been available since the beginning of the longitudinal study and see if they can see similar trends in that data. The two big jumps in the data can have a major corrupting effect.

We thank the reviewer for the in-depth observations concerning the quality of the underlying dataset. As outlined in the manuscript, we did not just utilize the data provided online, but performed additionally processing for each measurement point to make the comparison fair at that point in time. We have added an additional statement to that behalf. We are thankful for the suggestion to include web pages that were present throughout the observation period in addition to the performed evaluation. We re-evaluated the complete dataset and have re-worked the manuscript to incorporate the findings for the 46 web pages that were consistent in the longitudinal study, including a description in Section II describing the dataset. The evaluation of the Theil index, as this reviewer mentions below, aims at providing the evaluation of the impact of the dataset changes due to changes in the web pages selected by httparchive.org.

3. Actually by looking at httparchive.org it seems that the authors have not done anything really radical, just analyzed the mobile data available on the site separately. While science does not need to be complicated having access to such good dataset would perhaps have allowed more thorough analysis and firmer conclusions.

We whole-heartedly agree with the reviewer on the quality and amounts of data that this dataset provides to the community. Taking into consideration the confines of the IEEE Communications Magazine's limitations

with respect to space, we chose to provide the interested reader with a comparative overview of a commonly considered subset of duplicable characteristics of web pages. Certain other characteristics are not readily reproducible, as they would require additional network-centric and time-variant metrics not provided. We are in the process of performing additional evaluations that consider other metrics; however, those considerations are beyond the scope of this submission and require their own manuscript.

4. While energy spending is definitely important for mobile pages I feel the authors emphasize it too much in the intro as their study does not deal with energy at all. Simply focusing on trends as they do in their measurements should be enough for a good and interesting paper.

We agree with the reviewer that the emphasis on energy savings in the introduction was too strong and have hence de-emphasized the paragraph merging it with the previous one.

5. By looking at httparchive.org it seems that the mobile data corresponds to iPhone pages and desktop to IE. It would have been good to state that (as other browsers may render the pages in a different way)
We are thankful for pointing out this shortcoming and have added a clarifying statement in the revised version of the manuscript.

6. The discussion of Theil index is hard to follow. I first understood that with Theil index they compare how similar the desktop and mobile versions of the sites are, but by looking at Fig 2 with separate curves for mobile and desktop sites this cannot be the case. I am lost here and don't understand what this comparison is all about.

7. The data has some strange sections. E.g. in Fig 3a there are two major drops (in the middle of fixed curve) and a bump close to end. Discussing what these are would be interesting. Same in 3b close to end. Are they just random measurement errors or what?

We added an additional statement concerning the variability inherent in the dataset, as the additional measurement points to the end of 2013 show that these individual outliers constitute the overall dataset variability.

8. On the last line of page 9 I don't understand what the authors mean by saying that increased use of frameworks has an influence on bytes per image

We have re-worded the original statement, making it more clear that we refer to the common theming differences in modern content management systems employed by providers of these web pages.

9. Fig 7 is a bit vaguely described. How are the curves for handsets, tablets, smartphones and North America estimated? Moreover, my personal experience over the study period is that popular mobile websites (in Finland) have become faster not slower.

Reviewer: 3

Comments to the Author

First, I appreciate characterization papers - I always learn something new!

We are thankful for the timely review and the encouragement provided by the reviewer.

I feel the abstract and introduction conflate web content and internet content in the mobile space - and that's a hot topic about how much these two things interact. The study clearly deals with web content on mobile, but what might be more interesting is the content actually being consumed on mobile and the paper doesn't really give any insight into that.

We agree with the reviewer on the exciting avenue of taking mobile user interactions with the web into account. We acknowledge that this paper does not analyze the content consumption behavior of mobile web users, but has instead a focus on the web page characteristics differences. Approaches that such an analysis would have to take are likely either detailed studies based on network provider level details or limited trials with human subjects in controlled settings; both probably having additional IRB requirements. Instead this manuscript provides a comparison of web page characteristics, which in future works could be utilized to analyze comparatively how usage differences between “fixed” and mobile web and web page characteristics interplay.

III.A offers "continued increase in the average number of web objects requested likely results in additional networking overheads (such as connections setups, or DNS resolutions) and could have significant negative power consumption impacts". I really get nervous about phrases including "likely" and "could" when you are making the underlying argument of the paper. Do these things matter, and can you quantify how much? Common techniques such as persistent connections, spdy, and speculative DNS can impact this data more than is immediately obvious. What's the right metric here - power? latency? speedindex? bandwidth charges?

We acknowledge that the mentioning of these overheads was overall not clear enough, especially considering the limitations of this dataset (initial page view) and lack of lower level experimentation details. We hence have removed the statement and focus on the overall trend description in this manuscript, leaving detailed analysis of these characteristics for future works.

in III.C you say "overall trend for both request origins is rather linear and close" and I have to say I think that's the most interesting conclusion in the paper - I would love to see it highlighted better! Similarly you say that the "average amount of bytes per image is fairly close" which I also found to be a non-intuitive and interesting finding. (and one I suspect doesn't hold up when talking about mobile apps instead of mobile web apps - but it would be a great study to read).

Correct, likely mobile applications will target mobile delivery and consumption modalities directly and hence can optimize composition and delivery of data; we are considering this in our own related research

as well. As noted above, the additional limitation is that of initial page views without any prior caching, which with items such as the popular mobile jquery version is almost a given in non-initial views and potentially across page views on a mobile browser as well. We worked out the similar size per object more in the manuscript, as we agree that this is quite an interesting observation that deserved more highlighting than given in the original version of our manuscript.

in III.D you talk about "items cannot be cached effectively [...] due to limits below a day" and I think you may be missing a lot of the utility of caching. Many scenarios see most of their cache hits within that time frame (i.e. same session) as different portions of a site are thoroughly explored. Consider the number of transactions involved in that vs the small number involved in checking back on a different day for updates - short term cachability can be a serious boon.

We acknowledge the reviewer's motivation to evaluate the short-term caching differently. We have modified the valuation and focused on the description of the trends visible, as more details than provided in the dataset would be required to evaluate the caching in sessions.

in IV the discussion of download time and wait time doesn't seem to consider the more interesting metrics the industry uses for these things beyond total download page time. Time to first byte, time to first render, time to screen readable, and webpagetest.org speedindex are far more interesting metrics that don't necessarily track linearly against download time. Again, what is it about download time that is that interesting? I'm also concerned that the calculations are basically based on bandwidth calculations when web page download time is dominated by RTT more than bandwidth most of the time - so these headline numbers aren't very reliable but they are likely to stick with the reader as the take away of the article.