Summary Review Documentation for

“Watching Video from Everywhere: a Study of the PPTV Mobile VoD System”

Authors: Z. Li, J. Lin, M. Akodjenou-Jeannin, G. Xie, M. Kaafar, Y. Jin, G. Peng

Reviewer #1

Summary: The paper is a detailed measurement study of video requests and user viewing behavior to PPTV from mobile devices. It contains a wealth of data, for instance, on differences among WiFi and 3G users, skewedness in popularity, and truncated viewing. This data has implications for designing efficient systems and caching strategies for mobile video delivery.

Strengths: Mobile video is predicted to take over the world; this paper is one of the first detailed studies.

Weaknesses: Unclear from the results/analysis really that access from mobile devices all that much, and other than a few nuggets, the data is what you would expect based on intuition and studies of video access from non-mobile devices.

The data is fairly limited (14 days) to convincingly support some of the conclusions and observations, for instance, those around progression of video popularity (Sec 4).

Comments to authors: I like this paper. Even though there isn’t anything terribly surprising in it, there are a few interesting nuggets in the details. I almost wish you had dug deeper into the detail, rather than attempting a broad (but shallow) analysis from many different perspectives (user requests, temporal patterns, video popularity, access source, etc.)

The biggest techical shortcoming is that you don’t seem to be careful about drawing conclusions in the face of limited data. While 14 days is not too short a period, it cannot reliably support some of the conclusions you make from the data. The biggest culprit is Section 4.1, all of the analysis in which seems to be based on data from Dec 4 (after seeding is using Dec 1-3). Results in 4.2 also seem to be based on data from Dec. 4. Did you at least confirm the trends are similar for other days?

The most direct implication of your work is the design of video caching strategies, which your results suggest should be based not just on overall video popularity, but also on time of day, category, and geography. In light of this, I was a little disappointed to see that you didn’t combine all these factors into a caching scheme—one which also accounts perhaps for truncated viewing—and evaluated its effectiveness using your traces. That would have a great way to bring together all of your findings in a constructive manner.

Section 5 seems rushed and is difficult to follow. Even basic things are not explained. What do all the categories correspond to? You talk about category page and search, but what are the others? What category would an embedded link on a different site and clicking from my email correspond to?

Reviewer #2

Summary: This paper presents extensive analysis of data collected from a VoD server based in China. The data analyzed is that of mobile devices accessing videos. The paper has considered tablets as well as phones, access via WiFi as well as 3G. Several meaningful results are extracted from the data, and their implications are presented as well.

Strengths:
1. Work appears to be novel
2. The data is well organized and well presented
3. Appropriate implications of the high level results are presented

Weaknesses: I can identify no major weakness.

Comments to authors: The paper overall does a good job of motivation, data presentation, drawing implications, and comparing with prior work.

In Table-1, does viewing time include pauses? Likely not, since pauses may not be part of your log. Clarify. In Table-1, what does "overseas" mean? Outside of China? In Table-2, what does "WiFi+3G" mean? The same video is viewed from both WiFi and 3G? Clarify.

Sec-2.2 is good in that the paper has identified its own limitations.

In Sec-2.2, I didn’t follow as to how you get the video length from the web-based viewing logs. Do you assume that the web-based viewing results in a complete view?

I liked the fact that the paper draws comparisons in results from previous studies at appropriate junctures.

The result in Table-5 implies that measuring "popularity" by count of number of requests is not quite right. Comments?
Given the result in Table-5, it would be more appropriate to show Fig-5 y-axis weighted by NVT, right?

In your results, is the timestamp at the server? I assume so. In that case, the time-of-day based results need a slightly different perspective in interpreting, since clients come from various time zones. Comment on this aspect.

In 3.5, a WiFi-3G user is one who accessed at least once using WiFi and at least once using 3G, over the duration of your logs, right? Clarify.

In para before Fig-10 and in Fig-10, one set appears to be a superset of the other. Or do you mean the first group as those who accessed between 1 and 49 videos? If so, this needs a restatement.

Reviewer #3

Summary: This paper studies PPTV, one of the most popular mobile VoD system. It presents a lot of useful findings which can guide the optimization of the system. It’s one of the best submissions in my batch.

Strengths: This paper offers a new perspective into the widely-used PPTV system covering millions of users. It makes many novel insights which are highly valuable for content providers and advertisers.

Weaknesses: Writing needs to be improved.

Comments to authors: PPTV is a popular mobile VoD system in China. This paper presents an empirical study on PPTV. I very much enjoy reading the paper since it reveals many interesting facts about the behavior of a large-scale VoD system. Many findings can directly be used to optimize content caching and recommendation. Although some of the findings, e.g. video popularity and network impact on viewing behavior, are not the first-of-its-kind, there are still significant contributions in presenting the quantitative results.

The paper largely focuses on how to improve the caching strategy. Although the results seem to be helpful, can they really be exploited to dramatically improve caching? Implementing all the tricks seems to be overly complicated. One major pitfall is that the authors didn’t show how much cache hit rate increase can be derived from such optimization.

At several places, you mentioned that the results in this paper are similar to those observed in the Internet VoD systems. How many of your findings are really unique to mobile VoD?

Most results are closely tied to the PPTV system. Can they be generalized to other similar systems?

First paragraph on p10, I can’t follow how you estimate video upload time. The estimation method seems to be highly questionable.

Reviewer #4

Summary: PPTV is a large VoD system, and this paper examines the behavior of mobile users of the system, broken down by various categories, such as type of device, type of network access, and content type. The paper does a reasonably good job of analyzing PPTV’s mobile viewing, but doesn’t really place it in any context, nor do the lessons learned from this really seem new or useful.

Strengths: The paper analyzes a fairly large amount of data from a real system, and provides some insight into the traffic patterns at a large VoD service.

Weaknesses: The take-aways from the paper are either not surprising or already known. For example, TV series and movies are popular, or that people skim longer videos, etc., are not really that new.

Comments to authors: This paper takes a large amount of data, and while it is commendable that you’re making this analysis from a real, commercial system public, the analysis doesn’t really have much context or depth to it. There’s also some issue about slicing it up the way you do that seems to be too finely-slice - why does the mobile behavior require special consideration from the content provider, and what, if anything, changes in the PPTV environment as a result of the mobile viewing support? The paper really presents a lot of data, and while the analysis produces a lot of numbers and graphs, it feels to be short on depth. For example, when considering viewing time, the higher scores for WiFi are not entirely unexpected. An obvious question would be what’s driving this difference - is it the quality of the 3G connections, or is it an issue of pricing, and users not wanting to spend their possibly-limited data budgets watching this content? The WiFi viewership is presumably fixed, but is there any information on what’s happening with 3G? These kinds of questions could add more depth and provide more insight, but they never get raised in the paper.

There’s also a fair bit of analysis where the data is analyzed, the graphs are shown, and the discussion rightly points out that this behavior has already been identified previously in other systems. That’s good confirmation, but it also shows the lack of new contribution in some of these analyses. One of the conclusions in the paper is that caching the first parts of videos would be helpful, rather than necessarily doing whole-video caching. Isn’t that already known, and done in practice? The tendency to only preview videos has driven behavior in competing systems already, and is even user-visible. For example, YouTube bursts the first few seconds of a video, and then rate-paces the rest, to avoid sending more data than necessary when users preview the video. Even entire systems of encoding and caching support these kinds of ideas - the naming and caching of 10-second fragments in adaptive streaming systems already does this.

Overall, this paper would have been good if it were the first paper of its kind, but the bar keeps rising over time. The data is useful, but largely confirms what is already known about these kinds of systems. The focus on mobile, instead of being a differentiator, just adds a number of unasked and unanswered questions into the mix.

Reviewer #5

Summary: This paper studies a large scale mobile VoD system accessed through WiFi and 3G. The authors study access patterns and popularity of individual videos and contrast these findings for the two different access technologies and previously reported results on non-mobile VoD systems. They further provide implications on the design of the VoD system and appropriate caching architectures.

Strengths:

- nice analysis of the data set at hand
- comparison between WiFi, 3G and the type of handset
- implications that could be relevant to VoD system designers and caching architectures

**Weaknesses:**
- the work is rather incremental compared to existing literature in the domain
- there is nothing in the current manuscript that actually demonstrates benefits for VoD distribution to mobile clients when using the findings reported in the paper
- no justification behind using a small sample of the data set

**Comments to authors:** All in all this is a nice measurement study. The authors have a very detailed data set and they are able to analyze and contrast different metrics across access technologies, types of devices and previously reported work.

Having read the paper in detail I was however left wondering how important the identified differences are in reality. The authors do a really good job at pointing out implications but there is nothing in the current manuscript that takes those observations and shows that they actually matter in reality when it comes to the performance and efficiency of VoD systems and caching infrastructures in support of the dissemination of VoD content to mobile devices. To me, this is important, since the work is in an area which is rather over-published.

Second, the authors make a number of statements about abandoned viewing but there are no detailed results on this.

In summary, I think given that the incremental nature of the work, the authors should have spent more time demonstrating that the identified differences, as well as the prediction of content, would indeed lead to a more efficient VoD system. Otherwise, this is yet another VoD measurement paper. Having said that, it could be useful as a reference paper for people working in this space.

**Response from the Authors**

We would like to thank the reviewers for the valuable suggestions that we believe improved this paper. We tried to revise the paper following these suggestions whenever it was possible. In the following, we summarize the main revisions and changes we implemented addressing the main comments:

One of the main comments from the reviewers (and PC discussion) was: the need for an evaluation of the cache hit rate improvement if a strategy would rely on our observations/findings (i.e. caching the top 10% videos and only the first chunks of each video). We addressed this by adding a dedicated section (Current Section 6: Application: Design of an efficient caching system). Using traces driven simulations, we examined the overall performance of different caching strategies, including the one consisting in caching 10% of the popular videos, and only the first chunks. We assumed a virtual global cache system which can be deployed either in a centralized or fully distributed configuration. We consider the cache hit rate and the cache size as the main performance metrics. Videos are divided into chunks with equal lengths and we considered 3 different strategies (Full, Partial and Full caching with limited cache size). We do show that the partial caching strategy outcomes the two other strategies providing a better cache size Vs cache hit rate compromise. Please refer to Section 6 for more details.

A second comment addresses Mobile vs. Wired: We tried along different sections of the paper to compare our findings with results reporting on landline based VoD systems and/or mobile IPTV systems, when applicable. Unfortunately, we didn’t study the similarities/differences that might exist when extracting and examining the PPTV wired version of the logs (a different dataset is needed, and we experienced many challenges to find comparable logs) but instead we pointed out that we keep this as a future work.

As suggested by the first reviewer, the observation period of 14 days is relatively short to assess general behavior of video popularity distribution analysis. We removed such an analysis and discussion from figure 11, and only kept one subfigure in Fig 11 related to the Pareto principle. However, we also confirmed the high correlation between videos early view records and the near future popularity for other days, in addition to Dec.4. Specifically, we added new correlation results for few other days at the end of section 4.3.

We also fixed many typos and edited few paragraphs with minor corrections, to clarify the descriptions as suggested by reviewers C and D when applicable. However, some of the suggestions, although very interesting, are very challenging to incorporate in this paper, and rather have been left as future work (due to time limitation or lack of data). For example, we couldn’t systematically find evidence supporting the reasons we think are behind the different behaviors of WiFi and 3G network users observed in this dataset. We also leave the detailed analysis of the impact of access types (e.g. wired clients, web browsers and mobile clients) on video consumption patterns as future work.