

Consolidated Review of

Understanding the Effectiveness of Video Ads: A Measurement Study

1. Strengths

The paper uses an impressive data set and indeed performs a rigorous, systematic study, leading to several interesting takeaways.

Massive dataset. It features an impressive dataset.

Well written paper.

Thorough scientific treatment. The work is very careful with its conclusions.

2. Weaknesses

My main concern is that this paper studies the obvious, providing insights into knowledge that already exists

While well written, I thought the descriptions were too wordy and there were several repetitions

None

3. Comments

When I just started reading your paper, I was impressed with the dataset and potential results. However, while I was reading your paper, I realized that this paper studies the obvious -- the "rules" and takeaways that appear in the paper already exist in the advertising industry. The advertisers already know most if not all of these insights, and the ad networks already provide tools that help advertisers perform in-depth study and analysis of their specific viewer's behavior. At the end of the paper, I was left with wondering whether this paper merely exposes the IMC audience to knowledge that any advertiser has direct access to. More specifically, although the paper perhaps gives interesting takeaways to the general audience, those who actually need the information presented in order to take action (e.g., advertisers), already have access to it, and probably in a more tailored and easy-to-leverage interactive tools via the ad networks' tools.

The obvious question is why the industry's metric of "ad completion" is the right metric. Given the Akamai plugin runs on the client, one would imagine it has the ability to go beyond simply tracking whether the ad played to completion. For instance, it could track whether the user muted the sound during the ad or not, whether the user switched to another browser tab while the ad was playing or not, whether there was any user activity (mouse, keyboard) while the ad was playing or not (e.g., if the user went to make a sandwich during the ad). There is a tremendous amount of rich data the plugin could capture that would provide much deeper insights into ad effectiveness than simply whether the ad played to completion.

I understand the industry today focuses on the "ad completion" metric, so the paper's focus on evaluating things with that metric in mind makes sense. But there is the opportunity to educate the industry on better metrics, and I wish the paper had shot for this larger goal. That being said, I absolutely loved the paper given the problem statement the paper set out for itself.

I don't have any technical criticisms -- as far as I can tell, this work has been executed flawlessly. My only "issue" is that I

cannot get myself excited about the goal (help companies make more money from video ads). Surely, there must be more interesting questions that can be answered with this impressive dataset. E.g., do people who live in different geographic locations and/or under different political regimes tend to like similar video content? Is it possible to track changes in public opinion/interests or to predict major sociopolitical events based on video viewings?

Other issues that I have:

- ❖ You state over and over again the importance of causality over correlation. The reader gets it the first time she sees the statement. And perhaps even from your previous work. After a while, looking at the correlations and then the causality analysis for each metric that you study becomes annoying.
- ❖ Introduction - you state that there are other metrics like CTR, but you champion the completion rate. I agree that completion rate is important, but perhaps you should also state that you do not study CTR because you do not have this information (if indeed you do not have it because it is our of the cdn?)
- ❖ 1.2. - contribution (3) you have some statements about "value" to the user. I find it very strange that you decide to discuss value, which is completely subjective, hard to quantify, and you did nothing to validate this argument. You should remove these statements altogether, or at least provide some evidence that they are true.
- ❖ 1.2. - contribution (4) is not clear -- what is "connectivity to viewer" and how do you measure "content"? only later in the paper it becomes a bit clearer, but in the intro it is completely vague.
- ❖ Section 1.1 - Hulu actually does surveys in large scale (not sure how many replies they get, but at least they try).
- ❖ Some results/plots in the paper have no clear takeaway. E.g., Figure 3, Figure 10 (and entire section 5.3 for that matter), and section 5.1.1.

One thing you don't mention is if you can see if users change the tab, or switch to another application when the ad is playing I am not sure how you provide a GUID to identify each user. Does the user login to get the service. If not, the most you can do is get the computer in which the user is logged in, correct? Why are post-roll ads even present. It seems that nobody will really watch an ad that is played at the end of the program. In Table 2, what are the units? Section 4.1 seems excessive. It is simply describing the correlation coefficient, which I would think is pretty well known. You mention that your earlier work showed that users with worse connectivity showed more patience, while this work shows the opposite. Do you have any theories as to why this the case? I didn't understand the point about video content. To look at video content, you look at the completion rate of the video, which is the percentage of ads shown with the video that completed. You see large variations in completion rate. I am not sure how that correlates with the interestingness of the content. It is possible that some videos have higher number of ads, and therefore high completion rate (by your definition)

A minor comment about mid-roll versus pre-roll or post-roll ads: Yes, it makes sense that mid-roll ads are more likely to complete. But they are also more likely to annoy the hell out of the viewer, which may very well have an effect on his/her future visits to this site. E.g., I prefer sites that do not play mid-roll ads. Is there a way to capture "viewer annoyance"? Would it be feasible/make sense to compare how often sites that favor mid-roll ads are visited relative to sites that favor pre-roll ads?

4. Summary from PC Discussion

This paper was not discussed at the PC meeting.

5. Authors' Response

Reviewer comments are italicized and our responses follow.

1) *"My main concern is that this paper studies the obvious, providing insights into knowledge that already exists."*

We do not believe that any of these results are obvious. However, it is true that some of the results are known as folklore or known based on less rigorous analysis within the ad industry. Our work is the first *scientifically rigorous* analysis that we are aware of that provides quantitative insights based on a large data set measured across multiple video genres, content providers, geographies, connection types, and viewing platforms.

2) *"I agree that completion rate is important, but perhaps you should also state that you do not study CTR because you do not have this information."*

We have added two lines to Section 1.1 to clarify that our current data set does not allow us to measure and analyze CTRs or survey responses.

3) *"contribution (3) you have some statements about "value" to the user. I find it very strange that you decide to discuss value..."*

We have reworded it more clearly using cost-benefit rather than using "value" and emphasized that it is a plausible reason.

4) *"Some results/plots in the paper have no clear takeaway. E.g., Figure 3, Figure 10 (and entire section 5.3 for that matter), and section 5.1.1. "*

These charts provide a sense of the data set, example, Figure 3 is CDF of the video lengths. We shortened the discussion of these figures to the appropriate level.

5) *"One thing you don't mention is if you can see if users change the tab, or switch to another application when the ad is playing."*

If the user directs attention away from the video or ad by switching tabs, or the application, or indeed by performing a different activity while the ad is playing, we are not able to directly detect that event in our current data set. We have now clarified this limitation in Section 2.2.

6) *"I am not sure how you provide a GUID to identify each user."*

In most implementations, the GUID is tied to the device or desktop of the viewer. Thus, we cannot detect instances where one user watches video on another user's device. We have now added a clarification in the Section 2.2.

7) *"In Table 2, what are the units?"*

Time units of minutes are now clearly specified in the table.

8) *"Section 4.1 seems excessive."*

We have shortened the description of Kendall correlation.

9) *"To look at video content, you look at the completion rate of the video, which is the percentage of ads shown with the video that completed."*

The completion rate of a video is the completion rate of ads embedded in the video, not the completion rate of the video itself. This point was not clear and we have now defined it more clearly.

10) *Would it be feasible/make sense to compare how often sites that favor mid-roll ads are visited relative to sites that favor pre-roll ads?*

This is an interesting direction for future work but we do not study it in this paper.