PATTERNS OF LARGE SCALE ATTENTION

Mor Naaman
W/ Nir Grinberg, Minsu Park

@informorm
PATTERNS OF LARGE SCALE ATTENTION

Mor Naaman
W/ Nir Grinberg, Minsu Park

@informor
IMMERSIVE RECOMMENDATION SYSTEMS
Veit et al., ICCV 2016
IMMERSIVE RECOMMENDATION SYSTEMS

LOCALLY CONNECTED EXPERIENCES

NEW INTERFACES
LOCALLY CONNECTED EXPERIENCES
Open Questions

Forms of Inputs

CX Connector App

Contexts for Sharing
NEW Interfaces
AOL/CX: FIRST STEPS
Attention.
“A wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.”

— Herbert Simon
Capturing Attention
Concurrents: 29,781
Recirculation: 6%
Engaged Time: 0:51

80% of visitors read until here.
What is Scrobbling?

scrobble: skrob·bul [verb]

To automatically add the tracks you play to your Last.fm profile.

More than 600 devices and players scrobble, including:

- Music
- Video
- Flash
- Traffic
- Windows
- Android
- iOS
- Spotify
- YouTube
- Pandora
- Tidal
- Apple
- Napster
- Amazon
- Kickstarter
- Most Skipped Track

Chief Keef — I Don't Like
Luba Luft. No support, he informed himself. Most androids I’ve known have more vitality and desire to live than I do. 

The encounter with Kadalyi-Polokov had changed his ideas rather massively.

Snapping on his hovercar’s engine, he whisked nippity-nip up into the sky, heading toward the old War Memorial Opera House, where, according to Dave Holden’s notes, he would find Luba Luft this time of the day.

He wondered, now, about her, too.
Horace Silver 5tet - Song For My Father [1968]

Dorian Grey

Subscribe

14,246

503,820
TRENDS IN ATTENTION

competition

opportunities/tools for capture

mined in physical world
TRENDS IN ATTENTION

using finer attention models
PAYING ATTENTION TO ATTENTION
STUDY 1: HOW DO PEOPLE READ (PAY ATTENTION TO) ONLINE MEDIA STORIES?
Concurrents: 29,781
Recirculation: 6%
Engaged Time: 0:51
80% of visitors read until here
Chartbeat
How To Take Your Dog Camping

Spending time together outdoors is pretty much the neatest thing you can do with your dog. Here's how to do it safely.

In the 14 months that we've spent chasing coyotes around Joshua Tree and hiking in mountain lion-infested Sespe Wilderness, I've tagged along on a dirt bike trip to Big Sur twice, and a knife fight at a remote biker bar and climbed most of the mountains in Los Angeles.

For a few months, I've been tending to the Sheep. This mountain sheep was once more common in the wilderness north of Los Angeles and was once more plentiful than we've ever seen it in recent memory.

How To: Make Your PC and Mac Share Stuff Like Best Friends

Networking is stupid. You'd think it'd be possible to get two computers to talk to each other and share stuff between PCs and Macs, but it's not as nearly simple as you'd think. You'd also think it'd be possible to make 'em talk and share stuff like best friends.

You Need

To do this, you'll need a way to connect them. Most of the time, we use a shared drive, but we've also seen direct network shares. If you're using a shared drive, make sure it's accessible from both computers.

What are the key factors that influence reading depth?

Why do some articles perform better than others?
MEASURE ATTENTION: BENEFITS

understand engagement
rank content
fight clickbait
MEASURE ATTENTION: BENEFITS

reason about cognition, language

predict and understand content success
DATASET

2.3M read events from Chartbeat
9 popular news/media sites

For each event:
  Page metadata
  User ID
  Reading depth
What does it look like?
MEASURING: MAXIMUM DEPTH (PROPORTION)

A : 44.3%  B : 31.3%  C : 24.3%
Device differences?
Figure 2: Distribution of read proportion for short (red), medium (green) and long (blue) articles on 4 exemplary sites of Financial, Sports, Magazine and Premium News.

Does the type of device matter for the reading depth? Figure 3 offers a similar breakdown of read events by device type. In this case, we include for each page in our dataset exactly three read events \((r, p)\), one event (chosen at random) for each device type. For example, the blue curve on the Magazine row shows the distribution of reading depths for readers using a desktop computer when accessing the page. Again here, the X-axis is the portion of a page that was read by reader \(r\), and the Y-axis represents the density of people that read to that point. A two-sample Kolmogorov-Smirnov test found all three distributions in each site significantly different \((p < 0.001)\), except the Tablet and Desktop distributions in our Financial site \((p=0.065)\).

Across all sites, mobile users (red) drop earlier and in almost-uniform manner compared to tablet and desktop readers for the same articles. The larger-form devices, tablet and desktop, the read depths track very similarly, with higher likelihood of reading to completion than in mobile. The "flat" reading distribution for mobile users is particularly interesting in the Magazine site's case – it suggests a gap between form and user expectations. Mobile browsing is much more likely to occur when people are moving, often with limited time at their hands. One possible explanation is that the long form content available on the Magazine site does not match the time allotment of mobile readers and therefore gets dropped earlier.

Does reading depth change depending on where the readers come from? Figure 4 provides some insight into reading depth based on the user's source of referral to the article page. Again in this case, for each page in our dataset we include exactly four read events \((r, p)\), one event (chosen at random) for each referral type. The four different referral sources we consider are search, social, news, or internal. These sources capture whether the reader came to the article page from a search engine (e.g., Google), from a social referral (e.g. clicking on a link on Twitter or Facebook), or from other news source or news aggregation site (e.g. BBC, Google News). The internal category is for read events when the reader navigated to the page from another page on the same site. Note that, technically, we identify internal referrals as read events with no referral data. In most cases, this suggests intra-site traffic.

Figure 4 shows that only the Social referral source significantly differs from the rest of sources, and only in the 25%-50% region \((p < 0.05\) according to two-sample Kolmogorov-Smirnov test). For example, magazine readers that come from Social sources (dashed green line) are more likely to drop earlier than readers from other sources. The pattern of early drop for readers from Social was consistent for most sites we examined (including those not in the figure). There are of course a number of possible explanations for this phenomena. One is that "social" visitors are not as invested in the publication and content as other readers, who explicitly searched for the content or were referred to it by another news page (both strong signals of interest). We showed significant impact on reading depth of several...
Referral source?
Figure 2: Distribution of read proportion for short (red), medium (green) and long (blue) articles on 4 exemplary sites of Financial, Sports, Magazine and Premium News.

Does the type of device matter for the reading depth? Figure 3 offers a similar breakdown of read events by device type. In this case, we include for each page in our dataset exactly three read events \( (r, p) \), one event (chosen at random) for each device type. For example, the blue curve on the Magazine row shows the distribution of reading depths for readers using a desktop computer when accessing the page. Again here, the X-axis is the portion of a page that was read by reader \( r \), and the Y-axis represents the density of people that read to that point. A two-sample Kolmogorov-Smirnov test found all three distributions in each site significantly different (\( p < 0.001 \)), except the Tablet and Desktop distributions in our Financial site (\( p=0.065 \)).

Across all sites, mobile users (red) drop earlier and in almost-uniform manner compared to tablet and desktop readers for the same articles. The larger-form devices, tablet and desktop, the read depths track very similarly, with higher likelihood of reading to completion than in mobile. The "flat" reading distribution for mobile users is particularly interesting in the Magazine site’s case – it suggests a gap between form and user expectations. Mobile browsing is much more likely to occur when people are moving, often with limited time at their hands. One possible explanation is that the long form content available on the Magazine site does not match the time allotment of mobile readers and therefore gets dropped earlier.

Does reading depth change depending on where the readers come from? Figure 4 provides some insight into reading depth based on the user’s source of referral to the article page. Again in this case, for each page in our dataset we include exactly four read events \( (r, p) \), one event (chosen at random) for each referral type. The four different referral sources we consider are search, social, news, or internal. These sources capture whether the reader came to the article page from a search engine (e.g., Google), from a social referral (e.g. clicking on a link on Twitter or Facebook), or from other news source or news aggregation site (e.g. BBC, Google News). The internal category is for read events when the reader navigated to the page from another page on the same site. Note that, technically, we identify internal referrals as read events with no referral data. In most cases, this suggests intra-site traffic.

Figure 4 shows that only the Social referral source significantly differs from the rest of sources, and only in the 25%-50% region (\( p < 0.05 \) according to two-sample Kolmogorov-Smirnov test). For example, magazine readers that come from Social sources (dashed green line) are more likely to drop earlier than readers from other sources. The pattern of early drop for readers from Social was consistent for most sites we examined (including those not in the figure). There are of course a number of possible explanations for this phenomena. One is that "social" visitors are not as invested in the publication and content as other readers, who explicitly searched for the content or were referred to it by another news page (both strong signals of interest). We showed significant impact on reading depth of several
Can we predict?
Can we predict?

- article's average reading depth
- depth of individual read event
MODEL 1 (PRE PUBLICATION)

- length
- readability
- Sentiment
- site
- author
- topic
- read depth
SOME CONTRIBUTING FEATURES...

Length ↓

Author’s historical average ↑

Reading score ↓

Topics: Sports ↑  Science ↓
TAKEAWAY: READING DEPTH/ATTENTION IS (SOMEWHA) PREDICTABLE FROM SOURCE AND CONTEXTUAL CUES
FOLLOW-UP QUESTIONS

Better attention model (reading? skimming? skipping??)

Language and attention

Narrative and attention

Engagement with other media (YouTube)
STUDY 2: HOW DO PEOPLE WATCH (PAY ATTENTION TO) ONLINE VIDEO?
Monty Python's Ministry of Silly Walks (Full Sketch)
TWO DIFFERENT DATASETS

Random YouTube videos, average view duration

Tracking users, view time for each video they looked at
FOR EACH VIDEO...

actual duration
views
# comments
like ratio
language (sentiment) of comments
like/views, comments/views, share/view
...

50% viewed

Monty Python's Ministry of Silly Walks (Full Sketch)
MIXED-EFFECTS LINEAR REGRESSION

Control for category (humor vs. education)

Control for user
SOME EARLY RESULTS

- Length matters (obv)

- Positive association with view count (thank god)

- Positive association with like/view ratio (good)

- Negative attitudes matter too:
  - Negative language in comments pos. correlated
  - No impact of like/dislike ratio
<table>
<thead>
<tr>
<th></th>
<th>Random Videos</th>
<th></th>
<th>Individual Logs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
<td>Estimate</td>
<td>SE</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>0.490</td>
<td>0.011</td>
<td>0.658</td>
<td>0.021</td>
</tr>
<tr>
<td>video_length</td>
<td>-0.146**</td>
<td>0.005</td>
<td>-0.123***</td>
<td>0.009</td>
</tr>
<tr>
<td>elapsed_time</td>
<td>-0.034***</td>
<td>0.005</td>
<td>-0.006</td>
<td>0.013</td>
</tr>
<tr>
<td>view_count</td>
<td>0.026**</td>
<td>0.009</td>
<td>0.045**</td>
<td>0.014</td>
</tr>
<tr>
<td>comment_percentage</td>
<td>-0.029**</td>
<td>0.011</td>
<td>0.010</td>
<td>0.014</td>
</tr>
<tr>
<td>discussion_density</td>
<td>0.015**</td>
<td>0.005</td>
<td>0.023</td>
<td>0.014</td>
</tr>
<tr>
<td>like_percentage</td>
<td>0.094***</td>
<td>0.012</td>
<td>0.048**</td>
<td>0.016</td>
</tr>
<tr>
<td>dislike_percentage</td>
<td>0.004</td>
<td>0.006</td>
<td>-0.042**</td>
<td>0.015</td>
</tr>
<tr>
<td>like_dislike_ratio</td>
<td>0.014*</td>
<td>0.006</td>
<td>0.004</td>
<td>0.012</td>
</tr>
<tr>
<td>share_percentage</td>
<td>0.015**</td>
<td>0.005</td>
<td>-0.010</td>
<td>0.009</td>
</tr>
<tr>
<td>pos</td>
<td>-0.014**</td>
<td>0.005</td>
<td>0.009</td>
<td>0.011</td>
</tr>
<tr>
<td>neg</td>
<td>0.019***</td>
<td>0.005</td>
<td>0.022*</td>
<td>0.011</td>
</tr>
<tr>
<td>like_percentage:pos</td>
<td>0.011</td>
<td>0.011</td>
<td>0.014</td>
<td>0.012</td>
</tr>
<tr>
<td>like_percentage:neg</td>
<td>0.020*</td>
<td>0.009</td>
<td>0.004</td>
<td>0.012</td>
</tr>
<tr>
<td>dislike_percentage:pos</td>
<td>0.005</td>
<td>0.004</td>
<td>-0.012</td>
<td>0.008</td>
</tr>
<tr>
<td>dislike_percentage:neg</td>
<td>0.005</td>
<td>0.006</td>
<td>-0.028*</td>
<td>0.013</td>
</tr>
<tr>
<td>share_percentage:pos</td>
<td>-0.004</td>
<td>0.005</td>
<td>-0.012</td>
<td>0.011</td>
</tr>
<tr>
<td>share_percentage:neg</td>
<td>0.005</td>
<td>0.005</td>
<td>-0.011</td>
<td>0.010</td>
</tr>
<tr>
<td>comment_percentage:pos</td>
<td>0.010</td>
<td>0.008</td>
<td>-0.025*</td>
<td>0.011</td>
</tr>
<tr>
<td>comment_percentage:neg</td>
<td>0.004</td>
<td>0.007</td>
<td>-0.012</td>
<td>0.010</td>
</tr>
<tr>
<td>discussion_density:pos</td>
<td>0.009</td>
<td>0.005</td>
<td>0.000</td>
<td>0.010</td>
</tr>
<tr>
<td>discussion_density:neg</td>
<td>0.000</td>
<td>0.005</td>
<td>0.002</td>
<td>0.011</td>
</tr>
<tr>
<td>N</td>
<td>1,125</td>
<td></td>
<td>1,814</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.58</td>
<td></td>
<td>0.28</td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001
TAKEAWAY: PROPORTION WATCHED IS ASSOCIATED WITH CONTEXTUAL CUES
Paying Attention to Attention