The Potential of Social-aware Multimedia Prefetching on Mobile Devices

International Conference on Networked Systems (NetSys‘15), Cottbus, Germany


Stefan Wilk
Julius Rückert
Timo Thräm
Christian Koch
Wolfgang Effelsberg
David Hausheer
Mobile Video and its Challenges

Motivation

Mobile devices shape the trend for IP-based video delivery

- Video is forecasted to constitute up to 70% of mobile traffic (2019) [Ericsson14]
- A gap between network capacity demand and supply is forecasted
- Still mobile network based video delivery is challenging
  - Capped data volume contracts
  - Increased energy consumption
  - High latency and low bandwidth (3G)

If possible: Try to avoid video streaming over cellular network!
Prefetching of Video Content

Motivation

Content Prefetching

… store content on the mobile device prior to its consumption

… but prefetching requires:

- Accurate prediction on what users will watch in the future (what to prefetch)
- Knowledge on future network availability (when to prefetch)

Leveraging Online Social Network (OSN) information for prediction

- 59.4% of the worldwide video accesses are initiated on Facebook [Unruly2014]
- We see videos shared over an OSN as potential prefetching candidates

How can we leverage information from OSNs to predict future media consumption of a single mobile user?
Study Overview
Approach of our study

Study on the potential of OSNs for media prefetching

- Small-scale study with 14 participants using SonNet for 8 weeks
- Android App: SonNet as a proof of concept
- Analysis of media consumption behavior and their potential for prefetching on smartphones

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>male: 11</td>
<td>217.5</td>
<td>182.5</td>
</tr>
<tr>
<td># Friends</td>
<td></td>
<td>160</td>
<td></td>
</tr>
<tr>
<td># Posts</td>
<td>152.71</td>
<td>61.5</td>
<td>174.41</td>
</tr>
<tr>
<td># Group</td>
<td>31.29</td>
<td>13</td>
<td>46.71</td>
</tr>
<tr>
<td># Interests</td>
<td>10.78</td>
<td>6.5</td>
<td>10.45</td>
</tr>
</tbody>
</table>

Distribution of post types in our dataset:
- Link: 29%
- Pictures: 42%
- Status Update: 20%
- Video: 9%
Approach and Architecture of SonNet
Approach of our study

SonNet – OSN information collection
- SonNet is a mobile application mimicking the Facebook app
- Allows to intercept touch events and gather data such as:
  - Social ties
  - Newsfeed
  - Clicks and watching time
- Software allowed transmission of anonymized data to our server
Crucial Questions
Approach of our study

1. How much media – especially video – that is being distributed in the OSN Facebook is watched by the participants of the study?

2. Can the global popularity of a video be used to tell, whether one of the participants is going to consume media?

3. Is the origin of a post a suitable predictor for media consumption?

4. Is video consumption significantly different for individual user or user groups and can those groups be modeled?
Access rates - Multimedia content

Question 1

Results of interaction tracing - clicks on photo and videos posts

- In general – consumption rate of: 9.27% (Std. Dev.: 14.72; Median: 0)
- Videos show highest access rate with in average 19.15% (Standard Deviation: 28.5)
- Photos have an average of 7.76% (Standard Deviation: 14.53)

- Videos show the highest access rate in a dataset in which users tentatively skip most of the Facebook posts.

Smartphone video access rate in the OSN seems to be higher than access rates from stationary devices. [Li2014]
Can Popularity of a Post be a Predictor?

Question 2

Mapping of Caching approaches to Prefetching

- Popularity in OSNs can be measured by the number of annotations
  - Comments by other users
  - Likes of users
  - Shares by friends
- Most of the content is consumed shortly after „posting“ or solely shared in small groups

Popularity explains only a small fraction of the media consumption of an individual mobile user.
Origin of a Post

Question 3

How are posts shared – What is the origin?

- Most (77.55%) of the posts stem from direct friends (1-hop)
- Videos are much less shared by friends (55.45%)

Can we use the origin as a predictor?

- No significant difference can be observed when solely observing the categories:
  - Friend
  - Non-Friend

The coarse-grained distinction of friends and non-friends is not sufficient to determine if a media post is going to be watched.

S. Wilk, J. Rückert, T. Thräm, C. Koch, W. Effelsberg, D. Hausheer
Question 3

Can we classify the origin of a post in a more fine granular manner?

- Data submitted to our server has been anonymized
- Standard friend lists categories could be refactored from anonymized data:
  - Close Friends
  - Family
  - Other lists
- Classifying posts based on this categorization allows a good prediction, but only for a small subset of the video!

<table>
<thead>
<tr>
<th></th>
<th>Photos</th>
<th>Videos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close Friends</td>
<td>70.1%</td>
<td>85.7%</td>
</tr>
<tr>
<td>Family</td>
<td>82.6</td>
<td>50.1%</td>
</tr>
<tr>
<td>Other lists</td>
<td>9.2%</td>
<td>8.3%</td>
</tr>
</tbody>
</table>
Classifying User’s Media Consumption

Question 4

Classification of users regarding their media consumption behavior

- Some of the users are unwilling to watch media on the smartphone!
- Focus on high consuming users promises to improve prediction accuracy

Determining what and when to prefetch is not enough. Decide if a user requires a prefetching algorithm.

<table>
<thead>
<tr>
<th>User</th>
<th>Photos</th>
<th>Videos</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Friend</td>
<td>Total</td>
</tr>
<tr>
<td>U2 (High)</td>
<td>20.8%</td>
<td>24</td>
</tr>
<tr>
<td>U10 (High)</td>
<td>22.3%</td>
<td>103</td>
</tr>
<tr>
<td>U6 (Low)</td>
<td>0%</td>
<td>35</td>
</tr>
</tbody>
</table>

S. Wilk, J. Rücker, T. Thräm, C. Koch, W. Effelsberg, D. Hausheer
Classification of users helps!

- Probability to watch a video from a certain origin increases!
- Similar findings can be applied to other areas as well:
  - Interests
  - (Popularity)
Conclusion and Outlook
What we have learnt so far…

Mobile Facebook users in our study …
- Have a large proportion of photos on their newsfeed but only some videos
- Consume multimedia differently from our expectations
- Caching approaches do not help in prefetching for a single device!

But, there is some potential to predict video consumption…
- By classifying friends using the social proximity to the uses
- Classifying users according their media consumption

We have to show that the results are significant…

1) Analyzing our databases for classical access to OSNs.

2) Extending our Mobile OSN Prefetching Study.

S. Wilk, J. Rückert, T. Thräm, C. Koch, W. Effelsberg, D. Hausheer
Conclusion and Outlook
Our next steps and results

Analysis of prefetching potential in our database [PPW+15]
- Analysis of a Facebook dataset crawled by a browser plugin (123 days):
  - 774 users
  - 618,165 newsfeed posts
  - Videos are more likely to be clicked

Similar findings could be made in this study as in our pre-study
- Popularity of content is only a small help for predicting video accesses of a single device
- Origin of post can help significantly in prediction
- Media consumption varies significantly between users

<table>
<thead>
<tr>
<th>User</th>
<th>Photos</th>
<th>Videos</th>
<th>Link</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>231,582</td>
<td>46,055</td>
<td>340,528</td>
<td>618,165</td>
</tr>
<tr>
<td>Clicked</td>
<td>13,342</td>
<td>5,259</td>
<td>21,636</td>
<td>40,237</td>
</tr>
<tr>
<td>% Clicked</td>
<td>5.62%</td>
<td>11.42%</td>
<td>6.35%</td>
<td>6.51%</td>
</tr>
</tbody>
</table>

S. Wilk, J. Rückert, T. Thräm, C. Koch, W. Effelsberg, D. Hausheer
Thank you for your attention! Questions?

S. Wilk, J. Rückert, T. Thräm, C. Koch, W. Effelsberg, D. Hausheer