UP NEXT
Retrieval Methods for Large Scale Related Video Suggestion

M. Bendersky, L. Garcia-Pueyo, J. Harmsen, V. Josifovski, D. Lepikhin

Thanks to: Ajith Ramanathan, Jianming He, Su-Lin Wu, Tomas Lloret Llinares, Tamas Sarlos

Google, Inc.
Related video suggestion on the web
Current popular approaches

- **Hand-crafted playlists**
  - Does not scale for large video collections

- **Metadata based playlists**
  - More of the same/non-diverse

- **Co-view counts**
  - Works well for popular videos with many views
  - Tail/fresh content will have very sparse and noisy co-view data
Content-based video suggestion

- Model videos using weighted topic vectors
- Use an information retrieval approach to find related videos
  - Inverted index - topic → video index
  - Query - watch video
  - Documents - ranked videos using the index
  - Topic weights - tf-idf / learn from user feedback
Main contributions

● Effective deployment of content-based video suggestion on a very large scale

● Using implicit user feedback to improve content-based video suggestion

● Evaluation using a large live experiment
Video representation

metadata
uploader keywords
common search queries
playlist names
Freebase entities
...

World War Z TRAILER 2 (2013) - Brad Pitt Movie HD

Trailer (0.335)
World_War_Z (0.894)
Horror_Movie (0.112)
Brad_Pitt (0.995)
Approach I IR weights

Rank the related videos by:

\[ sc(V_W, V_R) = q(V_R) \sum_{\tau \in V_W \cap V_R} I_s(\tau) \frac{c(\tau, V_W)}{\log(1 + df(\tau))} c(\tau, V_R) \]
Approach I \textbf{IR weights}

Rank the related videos by:

\[ sc(V_W, V_R) = q(V_R) \sum_{\tau \in V_W \cap V_R} I_s(\tau) \frac{c(\tau, V_W)}{\log(1 + df(\tau))} c(\tau, V_R) \]

- Prior on the related video quality
- Inverse document frequency
- "Stopword" removal
- Topic weight
Consider a pair of potentially related videos

\[ P_R = \langle V_R^{(+)}, V_R^{(-)} \rangle \]

Learn a pairwise classification model to prefer \( V^{(+)} \) to \( V^{(-)} \)
Approach II Learning topic transitions from user feedback

Represent $P_R$ using a feature vector

$$X_{P_R} = \left[ I_{V^+}(\tau) - I_{V^-}(\tau) : \tau \in T \right]$$

Pairwise classification will assign higher weight to topics that "transition" to relevant videos
Parallel-update optimization

Large-scale optimization problem with
- huge feature space
- continuously updated training set


\[
\begin{align*}
t &= 1; \\
&\text{repeat} \\
&\quad \text{for instance } i = 1, \ldots, |S_P| \text{ do} \\
&\quad \quad q^i(i) = \mathcal{L}(w, \{x_i\}); \\
&\quad \quad \text{for transition } j = 1, \ldots, |T| \text{ do} \\
&\quad \quad \quad \mu_j^+ = \sum_{i: \text{sign}(x_{ij})=+1} q^i(i)x_{ij}; \\
&\quad \quad \quad \mu_j^- = \sum_{i: \text{sign}(x_{ij})=-1} q^i(i)x_{ij}; \\
&\quad \quad \quad \Delta_j^t = \frac{1}{2} \log \frac{\mu_j^+}{\mu_j^-}; \\
&\quad \quad \quad w^{t+1} = w^t + \Delta_j^t; \\
&\quad \quad t = t + 1; \\
&\quad \text{end} \\
&\text{until convergence or max # of iterations reached;} \\
&\text{return } w^{t+1}
\end{align*}
\]

Based on Collins et al. "Logistic regression, AdaBoost and Bregman Distances", 2012
Live experiment

- One month long live experiment on a sample of YouTube user traffic.
- Integrating co-view video suggestion with content-based video suggestion.
- Metrics:
  - Watch Time
  - Completion Rate
  - Abandonment Rate
Live experiment metrics summary

- Significant improvements compared to the baseline (co-view) system.
- Learning topic transitions outperforms pre-defined topic weighting.
## Live experiment by video type

<table>
<thead>
<tr>
<th>Video Category</th>
<th>IRTopics</th>
<th>TransTopics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Music</strong></td>
<td>$-0.64% \ (\pm 0.09%)$</td>
<td>$+0.28% \ (\pm 0.09%)$</td>
</tr>
<tr>
<td><strong>Gaming</strong></td>
<td>$+0.86% \ (\pm 0.68%)$</td>
<td>$+1.14% \ (\pm 0.66%)$</td>
</tr>
<tr>
<td><strong>News</strong></td>
<td>$+1.61% \ (\pm 0.41%)$</td>
<td>$+3.53% \ (\pm 0.41%)$</td>
</tr>
<tr>
<td><strong>Science and Technology</strong></td>
<td>$+2.43% \ (\pm 0.5%)$</td>
<td>$+3.79% \ (\pm 0.51%)$</td>
</tr>
<tr>
<td><strong>Pets and Animals</strong></td>
<td>$+3.70% \ (\pm 0.68%)$</td>
<td>$+4.16% \ (\pm 0.66%)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Video Age</th>
<th>IRTopics</th>
<th>TransTopics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>&lt; 1 month</strong></td>
<td>$+0.99% \ (\pm 0.26%)$</td>
<td>$+3.34% \ (\pm 0.26%)$</td>
</tr>
<tr>
<td><strong>1 month – 1 year</strong></td>
<td>$+0.50% \ (\pm 0.11%)$</td>
<td>$+2.24% \ (\pm 0.11%)$</td>
</tr>
<tr>
<td><strong>&gt; 1 year</strong></td>
<td>$+0.87% \ (\pm 0.07%)$</td>
<td>$+1.06% \ (\pm 0.08%)$</td>
</tr>
</tbody>
</table>
Live experiment by video type

<table>
<thead>
<tr>
<th>Video Category</th>
<th>IRTopics</th>
<th>TransTopics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.64% (±0.09%)</td>
<td>+0.28% (±0.09%)</td>
</tr>
<tr>
<td></td>
<td>+0.86% (±0.68%)</td>
<td>+1.14% (±0.66%)</td>
</tr>
<tr>
<td>Music</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaming</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>+1.61% (±0.41%)</td>
<td>+3.53% (±0.41%)</td>
</tr>
<tr>
<td>News</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science and Technology</td>
<td>+2.43% (±0.5%)</td>
<td>+3.79% (±0.51%)</td>
</tr>
<tr>
<td>Pets and Animals</td>
<td>+3.70% (±0.68%)</td>
<td>+4.16% (±0.66%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Video Age</th>
<th>IRTopics</th>
<th>TransTopics</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1 month</td>
<td>+0.99% (±0.26%)</td>
<td>+3.34% (±0.26%)</td>
</tr>
<tr>
<td>1 month – 1 year</td>
<td>+0.50% (±0.11%)</td>
<td>+2.24% (±0.11%)</td>
</tr>
<tr>
<td>&gt; 1 year</td>
<td>+0.87% (±0.07%)</td>
<td>+1.06% (±0.08%)</td>
</tr>
</tbody>
</table>
Summary

- Adapting the information retrieval paradigm for video suggestion

- Content analysis can significantly improve performance of current web video services

- Implicit user feedback can further improve content-based video suggestion