UP NEXT

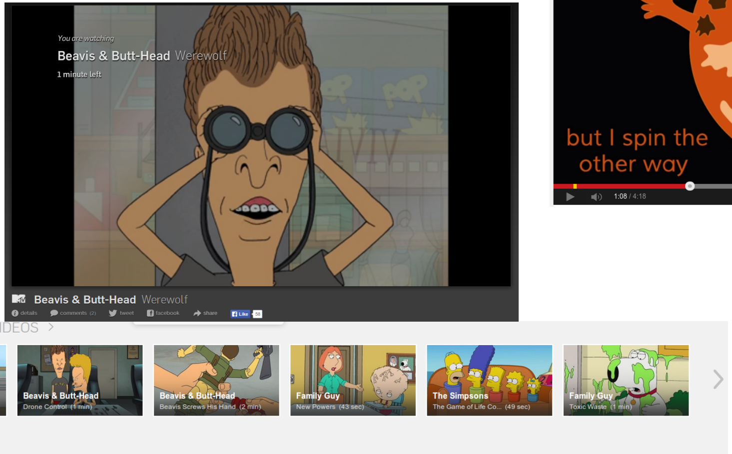
Retrieval Methods for Large Scale Related Video Suggestion

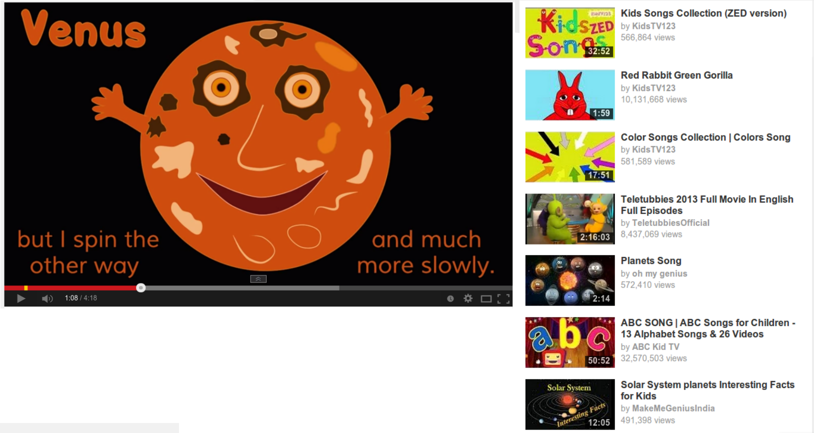
M. Bendersky, L. Garcia-Pueyo, J. Harmsen, V. Josifovski, D. Lepikhin

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
...

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:

Watch Video

$$sc(V_W, V_R) = q(V_R) \sum_{\tau \in V_W \cap V_R} \mathcal{I}_s(\tau) \frac{c(\tau, V_W)}{\log(1 + df(\tau))} c(\tau, V_R)$$

Related Video

Approach I IR weights

Rank the related videos by:

"Stopword" removal

Topic weight

$$sc(V_W, V_R) = q(V_R) \sum_{\tau \in V_W \cap V_R} \mathcal{I}_s(\tau) \frac{c(\tau, V_W)}{\log(1 + df(\tau))} c(\tau, V_R)$$

Prior on the related video quality

Inverse document frequency

Approach II Learning topic transitions from user feedback

Consider a pair of potentially related videos

Viewed by the user

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

Ignored by the user

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

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

Topic is associated with $V^{(+)}$

Topic is associated with $V^{(-)}$

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**

Algorithm 1: The *parallel-update* optimization algorithm [12] for learning topic transition weights.

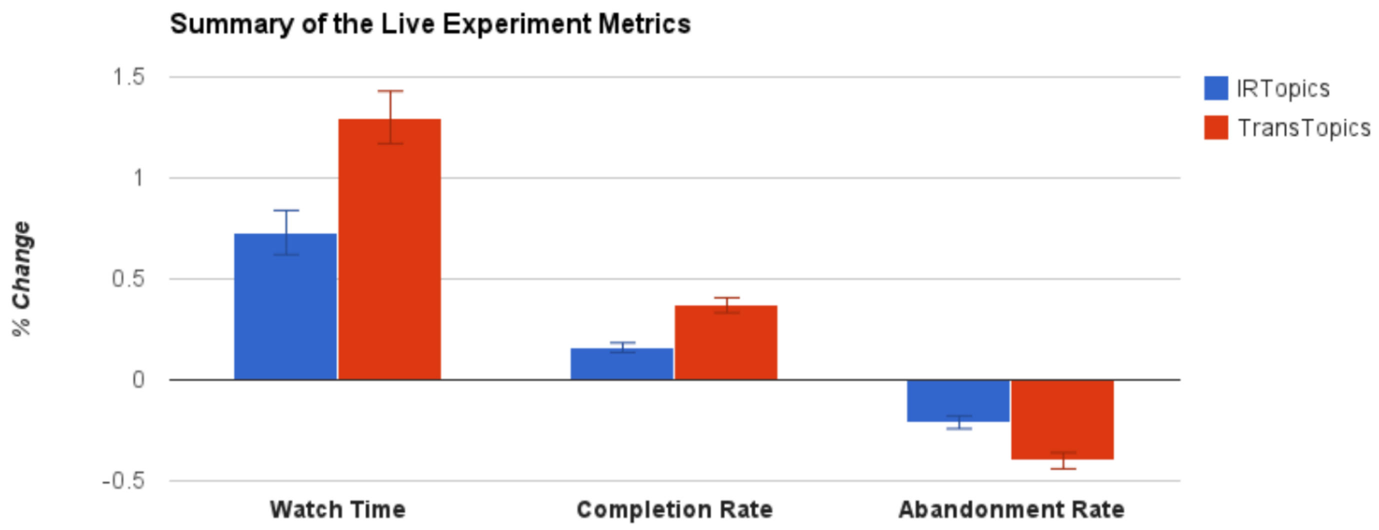
$t = 1$;
repeat
  for *instance* $i = 1, \ldots, |\mathbf{S}_P|$ **do**
    $q^t(i) = \mathcal{L}(\mathbf{w}, \{\mathbf{x}_i\})$;
    for *transition* $j = 1, \ldots, |\mathbf{T}|$ **do**
      $\mu_j^+ = \sum_{i:\, sign(x_{ij})=+1} q^t(i)|x_{ij}|$;
      $\mu_j^- = \sum_{i:\, sign(x_{ij})=-1} q^t(i)|x_{ij}|$;
      $\Delta_j^t = \frac{1}{2} \log \frac{\mu_j^+}{\mu_j^-}$;
      $\mathbf{w}^{t+1} = \mathbf{w}^t + \mathbf{\Delta}^t$;
      $t = t + 1$;
    end
  end
until *convergence* **or** *max # of iterations reached*;
return $\mathbf{w}^{t+1}$

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

	IRTopics	TransTopics
Video Category		
Music	-0.64% $(\pm 0.09\%)$	$+0.28\%$ $(\pm 0.09\%)$
Gaming	$+0.86\%$ $(\pm 0.68\%)$	$+1.14\%$ $(\pm 0.66\%)$
News	$+1.61\%$ $(\pm 0.41\%)$	$+3.53\%$ $(\pm 0.41\%)$
Science and Technology	$+2.43\%$ $(\pm 0.5\%)$	$+3.79\%$ $(\pm 0.51\%)$
Pets and Animals	$+3.70\%$ $(\pm 0.68\%)$	$+4.16\%$ $(\pm 0.66\%)$
Video Age		
< 1 month	$+0.99\%$ $(\pm 0.26\%)$	$+3.34\%$ $(\pm 0.26\%)$
1 month – 1 year	$+0.50\%$ $(\pm 0.11\%)$	$+2.24\%$ $(\pm 0.11\%)$
> 1 year	$+0.87\%$ $(\pm 0.07\%)$	$+1.06\%$ $(\pm 0.08\%)$

Live experiment by video type

	IRTopics	TransTopics
Video Category		
Music	-0.64% ($\pm$0.09%)	$+0.28\%$ ($\pm$0.09%)
Gaming	$+0.86\%$ ($\pm$0.68%)	$+1.14\%$ ($\pm$0.66%)
News	$+1.61\%$ ($\pm$0.41%)	$+3.53\%$ ($\pm$0.41%)
Science and Technology	$+2.43\%$ ($\pm$0.5%)	$+3.79\%$ ($\pm$0.51%)
Pets and Animals	$+3.70\%$ ($\pm$0.68%)	$+4.16\%$ ($\pm$0.66%)
Video Age		
$<$ 1 month	$+0.99\%$ ($\pm$0.26%)	$+3.34\%$ ($\pm$0.26%)
1 month – 1 year	$+0.50\%$ ($\pm$0.11%)	$+2.24\%$ ($\pm$0.11%)
$>$ 1 year	$+0.87\%$ ($\pm$0.07%)	$+1.06\%$ ($\pm$0.08%)

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