Overcoming the Imbalance Between Tag Recommendation Approaches and Real-World Folksonomy Structures with Cognitive-Inspired Algorithms

Dominik Kowald & Elisabeth Lex

Know-Center & Graz University of Technology

ESCSS, London, 15.11. – 17.11.2017

www.afel-project.eu  www.know-center.at
Social Tagging

- Social tagging is the process of **collaboratively annotating content** with keywords (i.e., *tags*).
- Essential instrument of Web 2.0 to **structure and search** Web content.

(a) Narrow folksonomy

(b) Broad folksonomy

[Zubiaga, 2009]
Tag Recommendations

BibSonomy
The blue social bookmark and publication sharing system.

edit your bookmark post

general information

URL
https://github.com/learning-layers/TagRec
This field is required.

title
TagRec framework
This field is required.

Description
Open-source tag recommendation evaluation framework

tags - describe the post

tags
learning-layers recommender tagrec
space separated

recommendation
recommender tagrec eval google learning-layers

post visibility

visibility settings
- public
- private
- other

save save and rate

Dominik Kowald & Elisabeth Lex
Know-Center & Graz University of Technology

[BibSonomy, 2017]
Imbalance

- Current tag recommendation algorithms are designed in a purely data-driven way
  - Tag popularity, user similarities, topic modeling, factorization of resource features, etc.
  - Rely on dense / broad folksonomy structures
- Most real-world folksonomies are sparse / narrow

| Dataset   | |U|  |R|  |T|  |Y|   |B|   |B/|U|  |B/|R|   |
|-----------|-----------------|-----------------|-----------------|----------|-----------------|----------|-----------------|-----------------|-----------------|----------|-----------------|----------|-----------------|
| Flickr    | 9,590           | 856,755         | 125,119         | 3,328,590| 856,755         | 89.338   | 1.000           |                 |
| CiteULike | 18,474          | 811,175         | 273,883         | 3,446,650| 900,794         | 48.760   | 1.110           |                 |
| BibSonomy | 10,179          | 683,478         | 201,254         | 2,986,396| 772,108         | 75.853   | 1.129           |                 |
| Delicious | 15,980          | 963,741         | 184,012         | 4,266,206| 1,447,267       | 90.567   | 1.501           |                 |
| LastFM    | 1,892           | 12,522          | 9,748           | 186,474  | 71,062          | 37.559   | 5.674           |                 |
| MovieLens | 4,009           | 7,601           | 15,238          | 95,580   | 55,484          | 13.839   | 7.299           |                 |
Approach

• The way users choose tags for their resources strongly corresponds to **processes in human memory** and its cognitive structures [Fu, 2008; Seitlinger & Ley, 2012]

• **Activation processes** in human memory → **ACT-R** [Anderson et al., 2004]

• **Activation equation** → usefulness of memory unit depends on **general usefulness** (i.e., frequency and recency) and usefulness in **current semantic context**

\[
A_i = B_i + \sum_j (W_j \cdot S_{j,i})
\]

\[
B_i = \ln(\sum_{j=1}^{n} t_{j,i}^{-d})
\]
How are **activation processes in human memory** influencing the **tag reuse behavior** of users in social tagging systems?


RQ1 Results

- The **more frequently** a tag was used in the past \((k > 0)\), the higher its reuse probability is.
- The **more recently** a tag was used in the past \((k < 0)\), the higher its reuse probability is.
- The **more similar** a tag is to tags of the **current sem. context** \((k > 0)\), the higher its reuse probability is.

→ The **activation equation of ACT-R** models these factors

[citeULike, 2016]
Can the activation equation of the cognitive architecture ACT-R be exploited to develop a tag recommendation algorithm, which is capable of overcoming the imbalance current approaches and real-world folksonomy structures?


RQ2 Results (nDCG@10)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>MP_r</th>
<th>MP_u,r</th>
<th>CF</th>
<th>LDA</th>
<th>PITF</th>
<th>FR</th>
<th>GIRPTM</th>
<th>ACT-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flickr</td>
<td>-</td>
<td>.569</td>
<td>.666</td>
<td>.280</td>
<td>.535</td>
<td>.561</td>
<td>.686</td>
<td>.711</td>
</tr>
<tr>
<td>CiteULike</td>
<td>.063</td>
<td>.392</td>
<td>.359</td>
<td>.138</td>
<td>.294</td>
<td>.392</td>
<td>.422</td>
<td>.438</td>
</tr>
<tr>
<td>BibSonomy</td>
<td>.091</td>
<td>.407</td>
<td>.369</td>
<td>.219</td>
<td>.327</td>
<td>.408</td>
<td>.409</td>
<td>.434</td>
</tr>
<tr>
<td>Delicious</td>
<td>.187</td>
<td>.358</td>
<td>.356</td>
<td>.271</td>
<td>.302</td>
<td>.292</td>
<td>.393</td>
<td>.431</td>
</tr>
<tr>
<td>LastFM</td>
<td>.283</td>
<td>.386</td>
<td>.317</td>
<td>.388</td>
<td>.414</td>
<td>.399</td>
<td>.397</td>
<td>.425</td>
</tr>
<tr>
<td>MovieLens</td>
<td>.271</td>
<td>.328</td>
<td>.254</td>
<td>.296</td>
<td>.324</td>
<td>.319</td>
<td>.326</td>
<td>.338</td>
</tr>
</tbody>
</table>

- **ACT-R** outperforms related tag recommendations methods in **narrow and broad** folksonomy settings

→ Cognitive-inspired approaches can **overcome the imbalance** between tag recommendations and folksonomies
Given that activation processes in human memory can be modeled to improve tag recommendations, can they also be utilized for hashtag recommendations in Twitter?


Scenario 1: Hashtag recommendations w/o current tweet

Scenario 2: Hashtag recommendations w/ current tweet

Activation processes in human memory can be utilized for hashtag recommendations in Twitter.
Conclusion

**RQ1** Activation processes in human memory (i.e., frequency, recency and semantic context) have an influence on tag usage practices.

**RQ2** The activation equation of ACT-R can be used to design a tag recommendation algorithm that **overcomes the imbalance** between current algorithms and the structure of real-world folksonomies.

**RQ3** This approach can also be generalized for **hashtag recommendations in Twitter**.

- **Future Work**
  - **Adapt** approach for other types of cognitive-inspired recommender systems (e.g., resource recommendation)
  - **Validate** offline results with **online studies**
Thank you for listening!
Questions / suggestions? → Poster

Dominik Kowald
Elisabeth Lex

Know-Center & Graz University of Technology

→ All evaluations have been conducted using the open-source TagRec tag recommendation benchmarking framework
   • https://github.com/learning-layers/TagRec
References (i)


References (ii)