Adaptive News Access
Adaptive Information Systems
Peter Brusilovsky and Jae-wook Ahn

Technologies

1. Adaptive Search
2. News Filtering/Recommendation
3. Adaptive News Navigation
4. Contextual News Access
5. Adaptive Presentation of News
6. News Aggregation
7. Mobile News Access
Research Fields Involved

- Information retrieval
- Topic Detection and Tracking (TDT)
  - Topic tracking — keep track of stories similar to a set of examples
  - Topic detection — build clusters of stories that discuss the same topic
- Collaborative Filtering
- Natural Language Processing

Item Personalization: Search vs. Filtering

Find personally relevant news stories based on a user model
Item Personalization: A Direct Approach

Item Personalization: Complications

- Dynamic content
- Changing interests
- Multiple interests
- Novelty
- Avoiding tunnel vision
- Editorial Input
- Brittleness
- Availability of meta-tags
(1) Dynamic content

- Continuous update
- Possible to disappear soon
- Sparse matrix problem
  - Very limited ratings overlap
- Selecting content provider problem?

(1) Dynamic content

- GroupLens
- Collaborative usenet recommendation
(2) Changing interests

- User interest changes
- Model adjust to the interest change
- “Concept Drift” — algorithmically detect concept (interest) change
- Windowing
- kNN - requires small training data
(2) Changing interests

\[ P(t) = \frac{1}{1 + e^{-t}} \]

<table>
<thead>
<tr>
<th>Days</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.953</td>
</tr>
<tr>
<td>2</td>
<td>0.924</td>
</tr>
<tr>
<td>3</td>
<td>0.881</td>
</tr>
<tr>
<td>4</td>
<td>0.818</td>
</tr>
<tr>
<td>5</td>
<td>0.731</td>
</tr>
<tr>
<td>6</td>
<td>0.622</td>
</tr>
<tr>
<td>7</td>
<td>0.500</td>
</tr>
<tr>
<td>14</td>
<td>0.029</td>
</tr>
<tr>
<td>21</td>
<td>0.001</td>
</tr>
<tr>
<td>30</td>
<td>0.000</td>
</tr>
</tbody>
</table>
(3) Multiple interests

- Multiple user interests
- Multiple categories
- Selecting appropriate algorithm
  - e.g. kNN

(3) Multiple interests

kNN example

- Airplane crash
- Nuclear missile launch
- Space shuttle launch
(4) Novelty

- Goal — find similar but not identical to already seen articles

(4) Novelty

- First Story Detection (FSD) task at TDT
- “Detecting the first story that discusses a previously-unknown event”
(5) Avoiding Tunnel Vision

- Why it’s not a good idea to show only “relevant” items?
- User interest vs. diversity
- Editorial input
- Explicitly boosting the diversity of stories

(6) Editorial input

- News recommendation is automating news editing
- Commercial / political motivation
- News aggregation
  - Avoiding editorial bias
(7) Brittleness & (8) Meta-tags

• A single action should not have a drastic or unrecoverable effect on the user model

• Meta-tags — mostly impossible

II. Adaptive news navigation

• Sites with fixed structures
  • Navigation support through the site

• Sites with dynamic structures
  • Determining the position of menu items within a menu hierarchy

• Local navigation support
  • Item to item link annotation/generation
Findory: Link Annotation

Adaptive Site Structure

- Smyth and Cotter, 2003
- Menu-select and scroll -- reduced by over 50%
III. Contextual recommendations

- Just-in-time retrieval
- Training data
- Currently displayed information
- Current user interest
- Using web-proxies or toolbar

Blinkx: Contextual recommendations
IV: Adaptive Presentation

• Symbolic Adaptive Hypermedia approach
  • Adaptive fragment selection
• Subsymbolic IR approach
  • Sentence-level adaptation and news summarization

SeAN: Content Adaptation

• SeAN decomposed news items into fragments of different nature
• Adaptation strategies can select a specific subset of fragments given user interests/task

Adaptive News Summarization

- News summarization is common, but how can we make it adaptive?
- Díaz and Gervás: Summary composed from selected statements. Statements to select are chosen with several heuristics.
- Rating by position - first 5 sentences are highest rated.
- Representation of sense-bearing thematic words - words that are specific for this text.
- Presence of terms from user interest models.
- As in the selection part, these three ratings are composed with three coefficients.


V. News aggregation

- RSS (Really Simple Syndication)
V. News aggregation

Personalization for News Aggregation

- Feed Selection
- Learning feed relevance to user interests
- Collaborative feed recommendation
- Feed Mapping
- Feed relevance to a specific dimension of user interests
Case Study: Mobile News Access

- Billsus & Pazzani
- Mobile news recommendation
- Limited screen space

Case study short & long term

- Quickly adapt to interest change while avoiding brittleness
- Short term
  - More adapted to users’ changing interests
- Long term
  - General interests
Case study
short & long term

- Short term — shuttle landing event itself
- Long term — science, technology, etc.

Case study
ST model algorithm

- kNN
- Model size is limited to N most recent stories
- Novelty — reduce score of too similar ones (with the already seen articles)
- No neighbors — passed to the long term model
Case study
LT model algorithm

- Create informative terms set (feature selection)
- For each individual news category
- Use for all users
- Words that appear most frequently important for 10,000 news articles
- Sort by DF
- Apply Naive Bayes

drug, cancer, space, cells, patients, women, crops, gene, launched, disease, food, virus, rocket, city, mission, bacteria, infection, children, heart, hiv, satellite, eclipse, blood, genetic, suns, winds, trial, mice, orbit, antibiotics, vaccine, resistance, russian, human, aides, storm, percent, brain, fda, cdc, mosquitoes, energy, test, damage, hurricane, computer, baby, government, hospital, texas.
Case study evaluation

• Compare personalization & static list
• Personalization — ordering
• Measure — display rank of selected stories
  • Higher the rank, better the system

Case study evaluation

• Average display rank of selected stories — 6.7 (s) vs. 4.2 (a)
Case study evaluation

- Mix static & adaptive stories on the same screen → see if users select adaptive stories

Conclusions

- News story overload on the Web
- Techniques
  1. News Content Personalization
  3. Contextual News Access
  4. Adaptive presentation
  5. News Aggregation