The NTCIR-11 IMine Task Kickoff Meeting

Yiqun Liu, Ruihua Song, Min Zhang, Zhicheng Dou, Takehiro Yamamoto, Makoto P. Kato, Hiroaki Ohshima, Ke Zhou

http://www.thuir.org/IMine/

September, 2013@NII, Tokyo
Diversified search

• Given an ambiguous/underspecified query, produce a single Search Engine Result Page that satisfies different user intents!

• Challenge: balancing relevance and diversity
IMine

Understanding user intents in Web search

Intent Mining

曖昧 (ambiguous)
The IMine task

Three subtasks

• **Subtopic Mining (SM):** Chinese, English, Japanese
  INPUT : query (e.g. “harry potter”)
  OUTPUT: ranked list of subtopic string
  (e.g. “harry potter book, harry potter film, harry potter the character…”)

• **Document Ranking (DR):** Chinese, English
  INPUT: query (e.g. “harry potter”)
  OUTPUT: diversified ranked list of web pages

• **Search Task Mining (TASKMINE):** Japanese
  INPUT : query (e.g. “pollen allergy treatment ”)
  OUTPUT: ranked list of task string which satisfies the given query
  (e.g: “Laser surgery”, “antiallergic drug”, “allergy mask”)
More User Behavior Data

• THUIR@INTENT/INTENT2: More user behavior data led to better performance.
• More raw click-through data: SogouQ has doubled its size to include click-through data collected from Sogou.com in 2011
  • 1.85GB => 3.85GB, over 40M user clicks
• More subtopic candidates generated from more recent user behavior data.
  • Search engine data provided by two major Chinese search engines will be adopted
• Find out whether more logs help improve SM/DR performance
Intent Annotation Using Logs

• SM results in pools will be clustered with click-through/pseudo RF data at first to generate preliminary candidate intent groups.

• Query frequency information will be taken into consideration during subtopic importance voting process.

• Data source: recently collected data from Sogou for Chinese SM, Bing for English/Japanese SM

• More credible SM qrels with less annotation efforts

• Perhaps the reusability of results could be increased
Crowd-Sourcing based Evaluation

• A search-engine-like annotation interface to collect feedback information from a relatively large number of unprofessional users (e.g. 50+ undergraduate students)

• Data collected from the interface: query, click, examination

\[ P(C_i = 1) = P(E_i = 1)P(R_i = 1) \]

• High correlation with examination
  • Preliminary results on 10 users.
  • KAPPA: 0.65, Accuracy: 0.83.

• Find out whether D# measures accords with user satisfaction
INTENT with/for Knowledge Graph

- Fuji in Wikipedia:
  - Fuji the actress
  - Fuji the Mountain
  - Fuji the apple

- Ambiguous queries: NTCIR INTENT #0205: 功夫 (kung fu), #0206: 生日快樂 (happy birthday)

- Mining and evaluating hierarchical user intents
  - SM task: return a two-level hierarchical list of subtopics
  - First level: at most 10 major categories of user intent ("Microsoft windows" for "windows")
  - Second level: other minor subtopics under the categories ("windows update", "windows 8.1 installation")
Task Mining (TASKMINE) subtask

Find tasks that satisfy the given query

Query: move to a new house

OUTPUT:

- Stop gas
- Stop electricity
- Submit document of move out
- Submit document of move in

Organisers: Takehiro Yamamoto, Makoto Kato, Hiroaki Ohshima

http://www.dl.kuis.kyoto-u.ac.jp/ntcir-11/taskmine/
# Summary of IMine

<table>
<thead>
<tr>
<th></th>
<th>INTENTZ2</th>
<th>IMINE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Topics</strong></td>
<td>• Chinese: 100</td>
<td>• Chinese: 50</td>
</tr>
<tr>
<td></td>
<td>• Japanese: 100</td>
<td>• Japanese: 50</td>
</tr>
<tr>
<td></td>
<td>• English: 50</td>
<td>• English: 50</td>
</tr>
<tr>
<td><strong>DR task setting</strong></td>
<td>• Chinese: SogouT (Ver.2008)</td>
<td>• Chinese: SogouT (Ver.2008)</td>
</tr>
<tr>
<td></td>
<td>• Japanese: ClueWeb JA</td>
<td>• English: ClueWeb12-B13</td>
</tr>
<tr>
<td><strong>Manual annotation</strong></td>
<td>• SM: 100 Chinese topics, 50 English topics, 100 Japanese topics</td>
<td>• SM: 50 Chinese topics, 50 English topics, 50 Japanese topics</td>
</tr>
<tr>
<td></td>
<td>• DR: 100 Chinese topics, 100 Japanese topics, pool depth=20</td>
<td>• DR: 50 Chinese topics, 50 English topics, 50 Japanese topics, pool depth=20</td>
</tr>
<tr>
<td><strong>Support from log analysis for annotation</strong></td>
<td>No</td>
<td>Support from log analysis for SM/DR annotation</td>
</tr>
<tr>
<td><strong>Crowd sourcing</strong></td>
<td>No</td>
<td>Crowd sourcing for Chinese DR</td>
</tr>
<tr>
<td><strong>Subtopic candidate</strong></td>
<td>Query suggestions from Bing, Google, Sogou and Baidu</td>
<td>• Query suggestions from Bing, Google, Sogou, Yahoo! and Baidu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Query facets generated by MSR from search engine results</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Query facets generated by THU from Sogou log data</td>
</tr>
<tr>
<td><strong>User behavior data</strong></td>
<td>SogouQ (data collected in 2008): 2GB approximately</td>
<td>SogouQ (data collected in 2008 and 2011): 4GB approximately</td>
</tr>
<tr>
<td><strong>DR Baseline</strong></td>
<td>• Chinese DR baseline</td>
<td>• ClueWeb12-B13 retrieval service is provided by CMU</td>
</tr>
<tr>
<td></td>
<td>• Japanese DR baseline</td>
<td>• SogouT retrieval service is provided by Tsinghua</td>
</tr>
</tbody>
</table>
Thank you

http://www.thuir.org/IMine/
http://www.dl.kuis.kyoto-u.ac.jp/ntcir-11/taskmine/