Total of 14 systems submitted
5 of our systems were selected in top 30 systems to be evaluated in the top 30 systems.

Main Challenges
- Queries are typically short and ambiguous in nature and might not capture the user’s intention effectively.
- For example for Japanese query: “喫煙”, English translation: “smoking”
  - Possible Query Intention-1: “dangers of smoking”
  - Possible Query Intention-2: “mechanism to quit smoking”
- Complex problem to re-rank the questions without understanding the user’s intent and focus of the query.
- Aim: How to model the aspects of textual relevance and information gained through user click data, to retrieve and present the information effectively to a user.

Systems Submission & Results
- Total of 14 systems submitted
- 5 of our systems were selected in top 30 systems to be evaluated in the final phase, out of a total of 65 participant submissions.

Overall Results

<table>
<thead>
<tr>
<th>ID</th>
<th>NDCG@10</th>
<th>ERR@10</th>
<th>Q-Measure</th>
<th>Credit-Phase-1</th>
<th>Credit-Phase-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>106</td>
<td>32</td>
<td>24</td>
<td>26</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>112</td>
<td>36</td>
<td>35</td>
<td>64</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>118</td>
<td>45</td>
<td>38</td>
<td>65</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>126</td>
<td>34</td>
<td>34</td>
<td>32</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>147</td>
<td>21</td>
<td>23</td>
<td>20</td>
<td>29</td>
<td>23</td>
</tr>
</tbody>
</table>

Our top systems’ ranking based on different evaluation metrics

Findings & Future Work
- Ranking of systems based on the online evaluation metric differed from that for the offline evaluation metrics.
- Need for more research to understand the factors behind contrary ranking results arising from the use of online and offline evaluation metrics.
- Our best systems in the online phase focused on modelling users click logs. Thus in future work we would like to explore more effective techniques for the exploitation of user logs and click distributions for ranking questions.
- Need for further investigation to find online and offline evaluation metrics that correlate well in order to address the task of ranking questions.

Dataset
- Yahoo Queries and respective Question-Answers

Methodology
- Learning to Rank algorithms: Explored L2R algorithms including Coordinate Ascent and MART
- Feature Selection & Combination: Explored alternative combinations of diverse feature sets capturing relevance of the user query and retrieved ranked list of questions

Analysis
- Coordinate Ascent algorithm performs relatively better than the Mart algorithm.
- Our best system (ID-130) based on NDCG@10 and ERR@10 was ranked “2” and “3” respectively.
- Based on Q-scores our best system (ID-123) was ranked “6”.
- Based on the cumulative credit our best system (ID-118) was ranked “4” and “6” for online phase-1 and final phase evaluation.
- Most of our submissions were heavily tuned to focus on relevance-based features, such as BM25 and LM scores, measuring the similarity of queries with the set of questions to be re-ranked.

Findings & Future Work
- Our best systems in the online phase focused on modelling users click logs.
- Thus in future work we would like to explore more effective techniques for the exploitation of user logs and click distributions for ranking questions.
- Need for further investigation to find online and offline evaluation metrics that correlate well in order to address the task of ranking questions.

Acknowledgement:

Piyush Arora and Gareth J. F. Jones
ADAPT Centre, School of Computing, Dublin City University, Dublin 9, Ireland
{piyush.arora,gareth.jones}@dcu.ie

The ADAPT Centre for Digital Content Technology is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund. This research is supported by Science Foundation Ireland through the CNGL Programme (Grant 12/CE/I2267) in the ADAPT Centre for Digital Content Technology.

Original Japanese page translated using the Google translation.