How Many Experts?
A New Task for
Enterprise Search Evaluation

Gianluca Demartini
demartini@L3S.de
L3S Research Center
Leibniz Universität Hannover
Outline

– Enterprise Search Evaluation
  • Past approaches
  • Tasks and metrics

– How Many Experts Are There?
  • Motivation
  • Evaluation
  • Possible approaches
– A TREC track for ES started in 2005 (TRECent)
– TRECent measures
  • Mean AP over expert
  • and reports bpref
– Different concepts for relevance have been proposed
– Privacy is the preeminent issue in PIM and Enterprise KM
TASKS AND METRICS

– Navigational tasks
  • Known Item Search
    – Generalized Success@10 (GS10)
  • Home Page Finding
    – GS10 or Mean Reciprocal Rank (MRR)
– **Informational tasks**
  - Document Search: MAP
  - Email Search: MAP
  - Entity Search: MAP, MRR, P@10

– **Transactional tasks**
  - Success@N

– **People Search tasks**
  - Expert Search task: R-Prec $\approx$ MAP
  - Number of experts
Outline

– Enterprise Search Evaluation
  • Past approaches
  • Tasks and metrics

– How Many Experts Are There?
  • Motivation
  • Evaluation
  • Possible approaches
Task: Find the *Number of experts* on a given topic, in the Enterprise

Motivation:
- HR managers can better understand the knowledge power available
- well-conceived selection of new project types
- move the core business of the enterprise
- identification of need for new expert employees
HME? - Approach

– General topic: larger number of people with some expertise
– Very specific topic: only few people
– An IRS should:
  • understand the specificity of the topic
  • retrieve a reasonable number of experts
- Look outside the Enterprise!
- Consider not only experts within the Enterprise but overall experts in the topic
- Compare the most knowledgeable people working for the enterprise with the current state-of-the-art knowledge in the world on the given topic
HME? - Evaluation of the task

– Measure the quality of the estimation made by the IRS
– Possible errors are overestimation / underestimation

\[ H := 1 - \frac{|SNE - UNE|}{|C|} \]
## HME? - Experiments

- TRECent 2006: computed H for 91 runs
- Computed correlation if H with standard metrics

<table>
<thead>
<tr>
<th>Collection</th>
<th>Avg Number of experts per topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRECent 2005</td>
<td>30.18</td>
</tr>
<tr>
<td>TRECent 2006</td>
<td>28.4</td>
</tr>
<tr>
<td>TRECent 2007</td>
<td>3.04</td>
</tr>
</tbody>
</table>
HME? - H vs Gold Standard

- Correlation values of the IRSs rankings done according standard IR metrics and H, for the TRECent 2006 collection.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Kendall $\tau$</th>
<th>Spearman $\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAP</td>
<td>-0.25</td>
<td>0.005</td>
</tr>
<tr>
<td>R-Prec</td>
<td>-0.22</td>
<td>0.008</td>
</tr>
<tr>
<td>GMAP</td>
<td>-0.27</td>
<td>0.0006</td>
</tr>
<tr>
<td>P5</td>
<td>-0.1</td>
<td>0.30</td>
</tr>
<tr>
<td>P10</td>
<td>-0.15</td>
<td>0.09</td>
</tr>
</tbody>
</table>
IDENTIFYING HIGHLY EXPERT EMPLOYEES

– Thresholding the number of experts
– Thresholding on the average score
– Top N% thresholding on the average score

<table>
<thead>
<tr>
<th></th>
<th>Ex1</th>
<th>Ex2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>3</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>4</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>

– Thresholding the expertise score
CONCLUSIONS

– We have motivated and defined a new Enterprise Search task (Number of Experts)
– The evaluation focuses on the quality of the estimation made by IRS
– We listed some possible approaches for solving the task
Investigate, evaluate, and compare possible ways of identifying highly expert employees.

Study the distributions of expertise over candidates and topics:
- Empirically finding the distribution of the expertise scores (from the RSVs)
- Approximating and normalizing RSVs.
Thanks