Xiangsheng Li, Yiqun Liu, Jiaxin Mao, Zexue He, Min Zhang, Shaoping Ma
Information Retrieval Group, Department of Computer Sci. & Tech., Tsinghua University

Understanding user reading attention during relevance judgement

Background & Motivations

- Reading behavior plays an essential role in multiple areas of Information Retrieval (IR) researches
- Reading in cognitive psychology
  - Fixation and saccade
  - Cognitive models such as EZ-Reader (Reichle et al. 1998), SWIFT (Engbert et al. 2002) and Bayesian reading model (Bicknell et al. 2010).
- Reading in information retrieval
  - Reading behavior is different from general reading because of the existence of information needs
  - Users’ reading attention has position bias
  - Influenced by search task types and query terms
- How ranking models estimate users’ reading attention
  - Query centric assumption (HoChung et al. 2007)
  - Relevant information only exists in the context around the query terms
- Attention mechanism in neural models (Dzmitry et al. 2015)
  - Weighting each word(sentence) automatically

Research framework

- Reading behavior plays an essential role in multiple areas of Information Retrieval (IR) researches
- Reading in cognitive psychology
  - Fixation and saccade
  - Cognitive models such as EZ-Reader (Reichle et al. 1998), SWIFT (Engbert et al. 2002) and Bayesian reading model (Bicknell et al. 2010).
- Reading in information retrieval
  - Reading behavior is different from general reading because of the existence of information needs
  - Users’ reading attention has position bias
  - Influenced by search task types and query terms
- How ranking models estimate users’ reading attention
  - Query centric assumption (HoChung et al. 2007)
  - Relevant information only exists in the context around the query terms
- Attention mechanism in neural models (Dzmitry et al. 2015)
  - Weighting each word(sentence) automatically

Research questions

- RQ1: How does user make the relevance judgment while reading a document?
- RQ2: How does user allocate his/her attention during the relevance judgment process?
- RQ3: What are the factors that affect user’s attention allocation in relevance judgment?

Experiment Setup

- 15 Queries from NTCIR13 WWW
- 4 documents in a topic, 60 in total
- Only read 15 documents for a single user
- Relevance agreement (Fleiss Kappa):
  - 4-level: 0.326
  - 2-level: 0.757
- Annotation agreement:
  - Kappa: 0.364
  - Average Precision: 0.734

Two-stage Reading Model

- Average Fixation Rate
  - Position: 0–20% 20–40% 40–60% 60–80% 80–100%
  - Relevant: 0.233 0.233 0.230 0.223 0.169
  - Irrelevant: 0.242 0.242 0.242 0.242 0.242
  - Diff: -0.806% 4.55% 7.48%* 13.71%* 15.11%*

- Average Reading Time Per Word (msec)
  - Position: 0–20% 20–40% 40–60% 60–80% 80–100%
  - Relevant: 106.3 112.61 92.81 82.93 52.47
  - Irrelevant: 116.33 99.57 81.71 67.07 45.65
  - Diff: -8.62% 13.10%* 13.58%* 23.65%* 14.94%*

Table 1: Reading behaviors on different vertical position in relevant and irrelevant documents.

Reading Behavior Bias

- Position bias is common during reading
- Users’ reading attention correlates with word surprisal
- Search task type affects users’ reading attention
- Query centric bias is not significant in the first reading step

- Table 2: Influence of query terms to reading attention at different vertical positions when window size is 5

Attention Prediction

- Structure, content, query features can be used to model reading attention
- Relevant text relied on different features on different granularity prediction

Table 3: PCC results of different level predictions