Can Readability Enhance Recommendations on Community Question Answering Sites?

GOAL

Examine the impact text complexity has when incorporated into the recommendation process in community question answering (CQA) sites

BACKGROUND

- Readability information has improved recommendations in Twitter hashtag and K-12 book domains [1, 2]
- We argue that the recommendation process within CQA sites should go beyond content matching and answer-feature analysis, as users do not have similar reading capabilities

RESEARCH QUESTION

- Does the quality of answer recommendations improve when readability is incorporated into the process?

DATA & INITIAL ANALYSIS

Dataset

YAHOO! **L16 Dataset**

1527 tuples of the form < id, q, Q, A, R_{qA} >

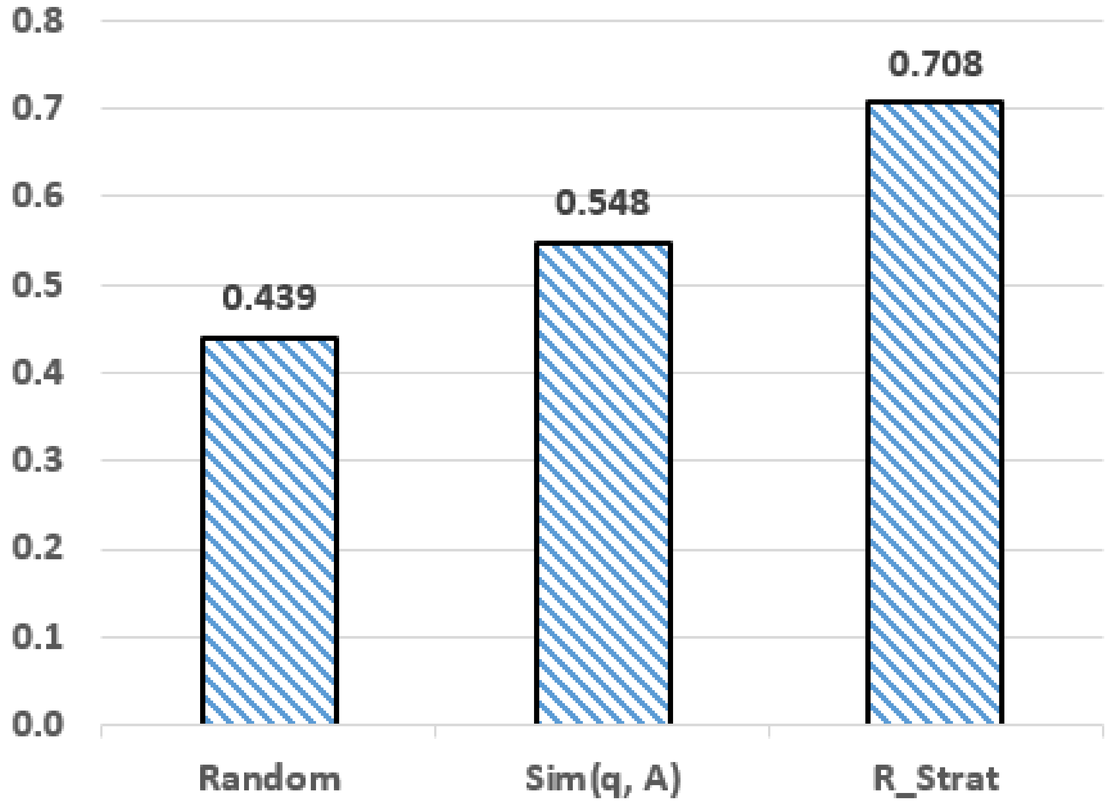
Query_id — Query — Question — Answer to Q — User Satisfaction Rating of A given q

% of Queries	16	8	2	1	1	72
Answers(s)	1	2	3	4	5	> 5

Answer distribution on Y! L16

Initial Analysis

- **Metric:** Mean Reciprocal Rank (MRR)
- **Readability Enhanced Strategy (R_Strat):**
  - ✓ Sim(q, A): WordNet based similarity between q and A
  - ✓ Sim(q, Q): WordNet based similarity between q and Q
  - ✓ RSim(q, A):
    - ➢ Euclidean distance between readability scores of q and A
    - ➢ Readability estimated using *Read2Vec*, a deep neural network based strategy suitable for short texts
  - ✓ Aggregate Sim(q, A), Sim(q, Q), Rsim(q, A) using Linear Regression
- **Baselines**:
  - Random: Answer recommendations generated arbitrarily
  - Sim(q,A)



Performance assessment based on MRR

FINDINGS

In the CQA domain, reading level information is an influential factor in terms of enhancing answer recommendations and can be used to improve user satisfaction on a recommendation

FUTURE WORK

- Conduct a deeper study using other CQA sites, such as Quora or StackExchange
- Analyze queries for additional factors, such as relative content-area expertise

REFERENCES

[1] M. S. Pera and Y.-K. Ng. Automating readers' advisory to make book recommendations for k-12 readers. In ACM RecSyS, pages 9–16, 2014.

[2] I. M. Azpiazu and M. S. Pera. Is readability a valuable signal for hashtag recommendations? Poster at ACM RecSys, 2016.

BOISE STATE UNIVERSITY
COLLEGE OF ENGINEERING
Department of Computer Science

Oghenemaro Anuyah, Ion Madrazo Azpiazu, David McNeill and Sole Pera
People and Information Research Team
Department of Computer Science
Boise State University, Boise, Idaho, USA

People and Information Research Team