

UB at the NTCIR-13 STC-2 Task: Exploring Syntactic Similarities and Sentiments

Jianqiang Wang

Department of Library and Information Studies, Graduate School of Education, University at Buffalo, U.S.A.
jw254@buffalo.edu, <http://www.buffalo.edu/~jw254>

TASK: Short Text Conversation (STC-2) in Chinese

For each of the 100 new posts, retrieving and ranking 10 most *appropriate* comments from a collection of 4,305,706 unique Sina WeiBo comments.

TECHNIQUES

Syntactic Similarity Analysis

Comments that are substrings of the post are eliminated

Sentiment Analysis

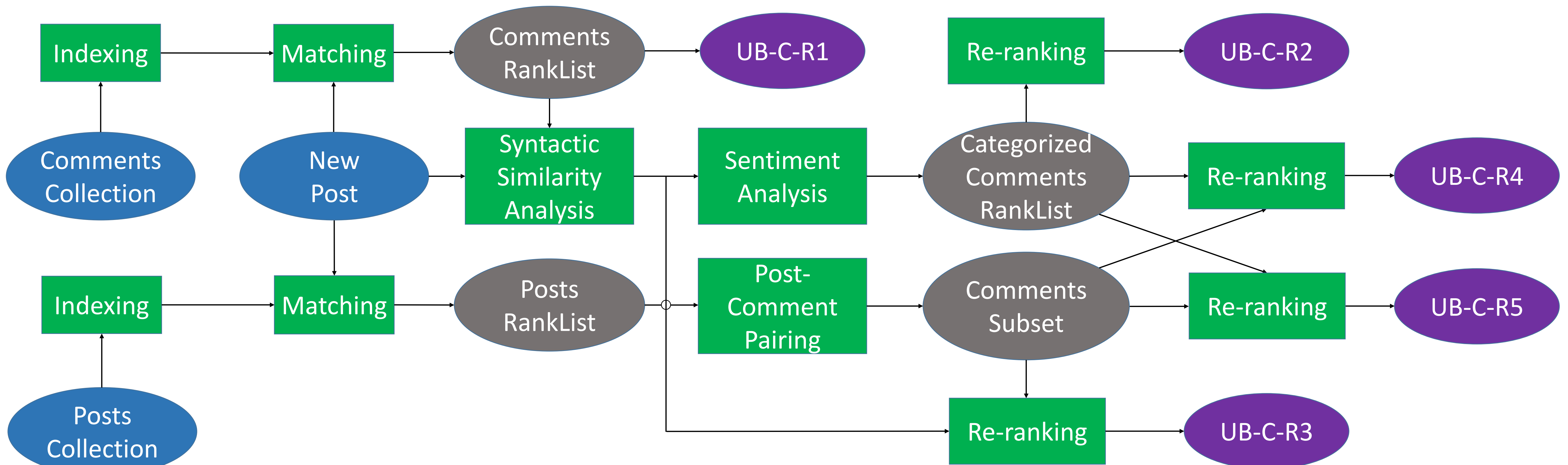
Comments of positive sentiments are given more weight if the post is also of a positive sentiment. A list of ~3k words of positive sentiments were used.

Exploring Existing Post-Comment Relationships

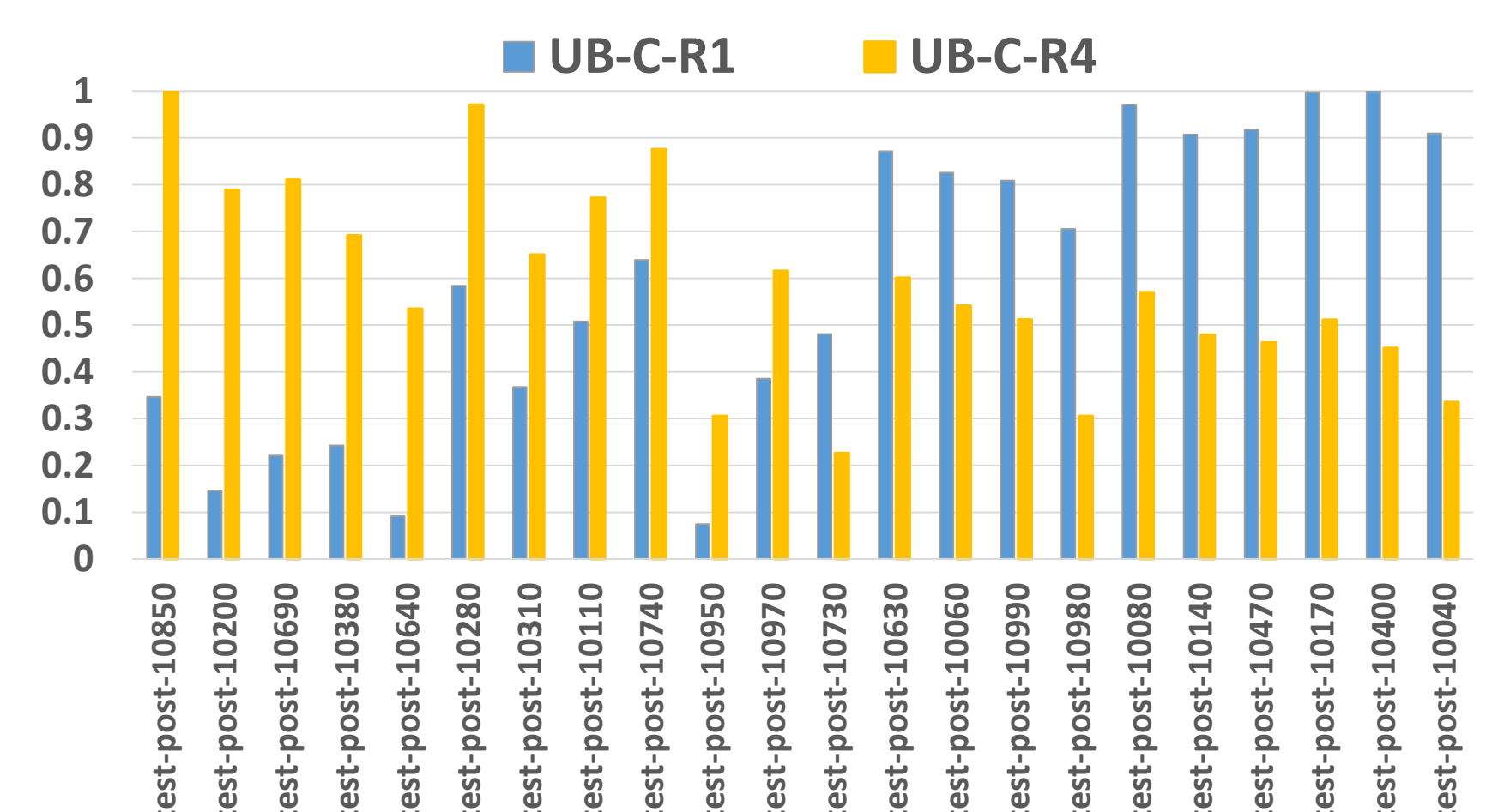
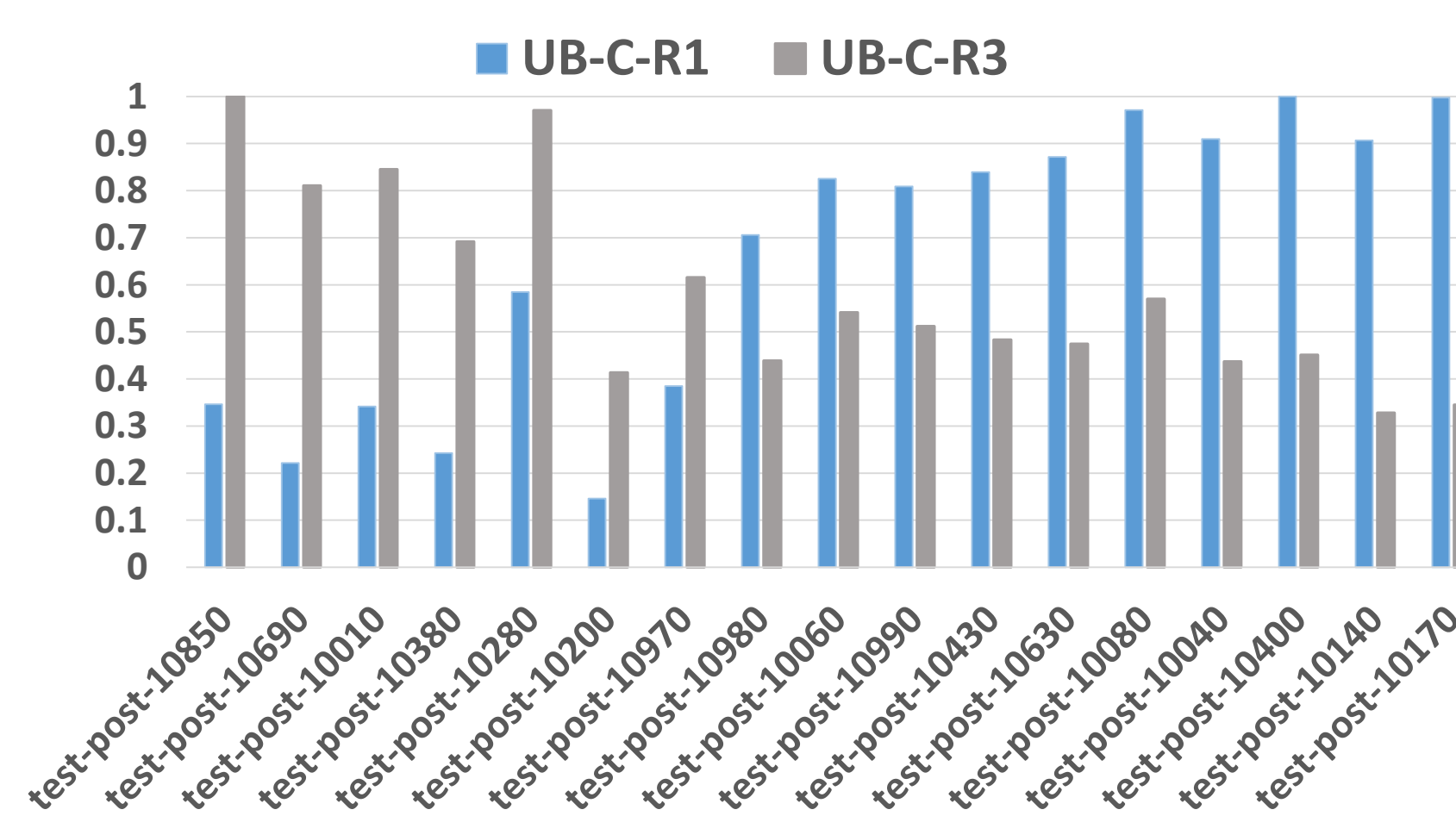
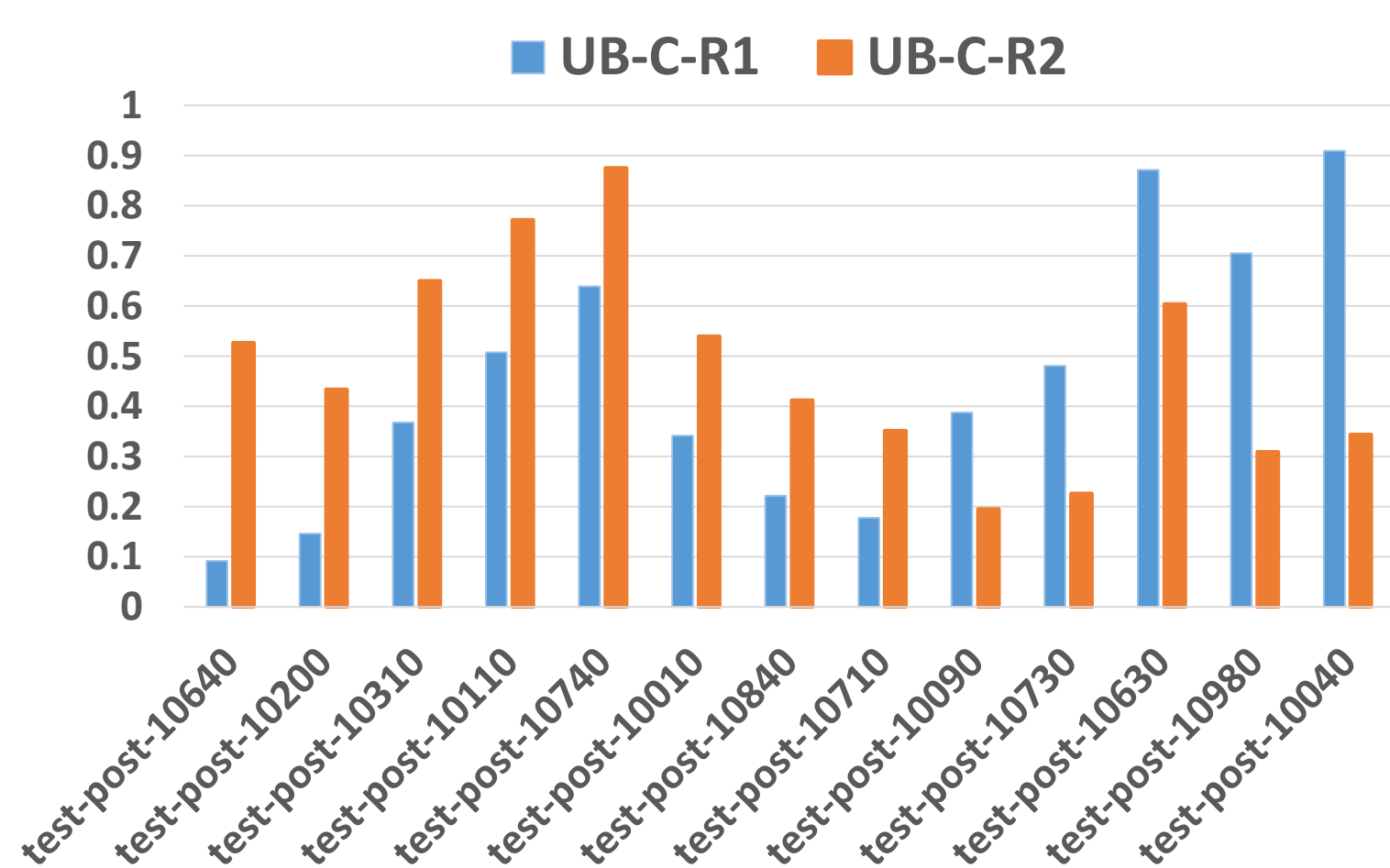
Comments are given more weight if they previously responded to the posts that are relevant to the new post

Combining Evidence for Re-Ranking

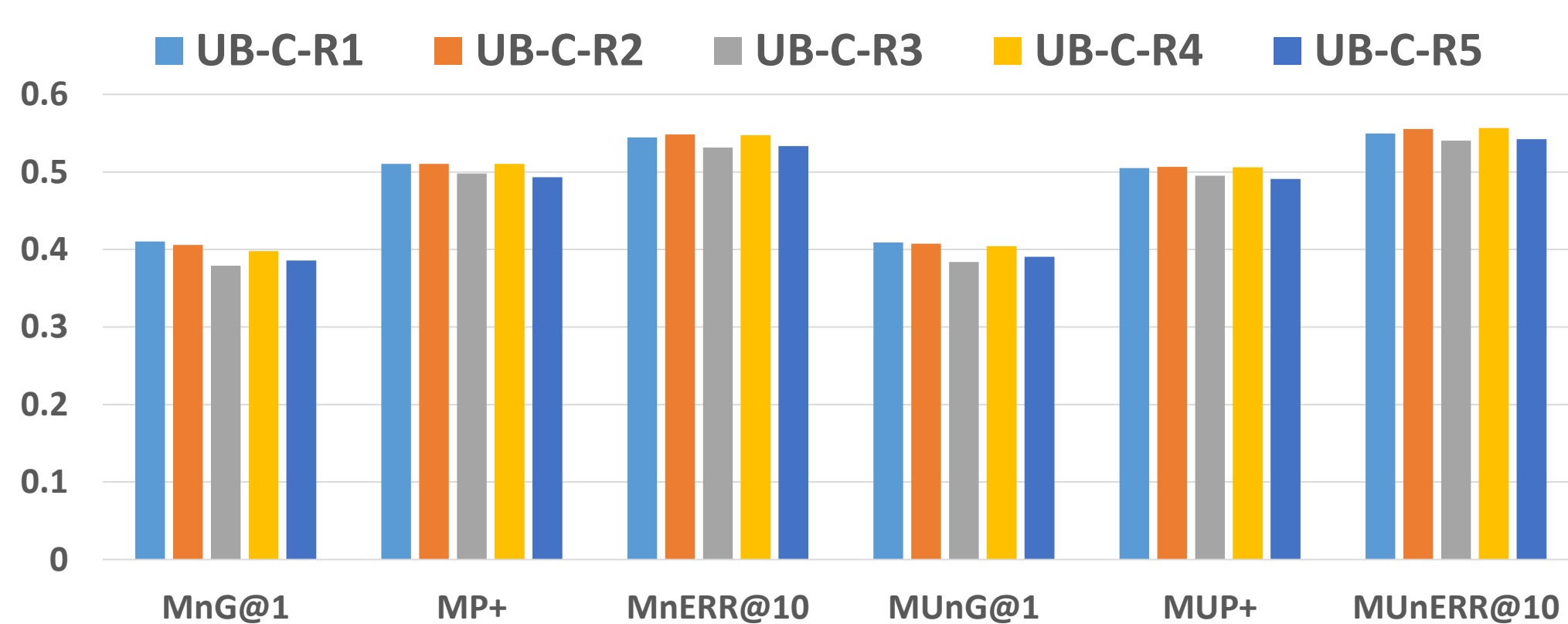
- Initial retrieval with Okapi BM25 weighting (UB-C-R1)
- Comments of positive sentiments are re-ranked higher (UB-C-R2)
- Comments re-ranked higher if they responded previously to posts relevant to the new post (UB-C-R3)
- Combine the above two: re-ranking based on sentiments first (UB-C-R4), or otherwise (UB-C-R5)



OFFICIAL RESULTS AND ANALYSIS



Pair-wise post-by-post comparisons of runs, showing only those posts whose absolute difference of nERR@10 between the two runs is at least 0.2.



Comparison of the five official runs

Summary

- Consistent relative effectiveness of runs across measures
- No statistically significant difference between runs
- Techniques helpful in some cases but not so in some other cases; causes including polysemy and cynical expressions
- Existing post-comment relationships not very helpful due to potential topic drifting

CONCLUSION AND FUTURE WORK

- Retrieval-based method is effective for STC
- Proposed techniques have limited effectiveness
- Need to refine the techniques by considering also negative sentiments
- Better modeling of different appropriateness aspects
- Exploring knowledge structure and knowledge base
- Incorporating reputations, active areas, posting frequencies, etc. of the poster/commenter in retrieval and ranking