

When Topic Models Disagree: Keyphrase Extraction with Multiple Topic Models

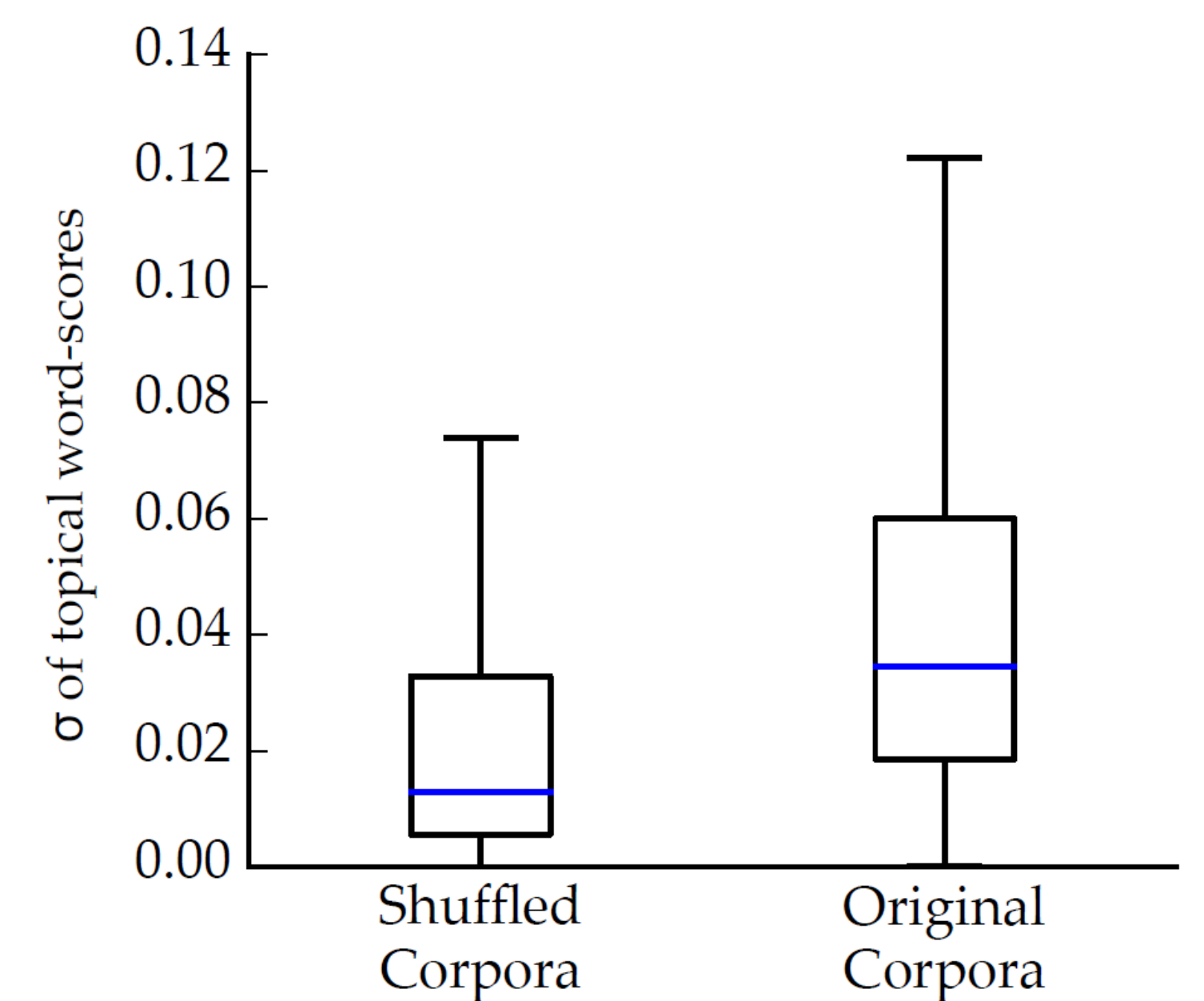
Lucas Sterckx, Thomas Demeester, Johannes Deleu, Chris Develder
Internet Based Communication Networks and Services, Ghent University - iMinds

Topical Keyphrase Extraction

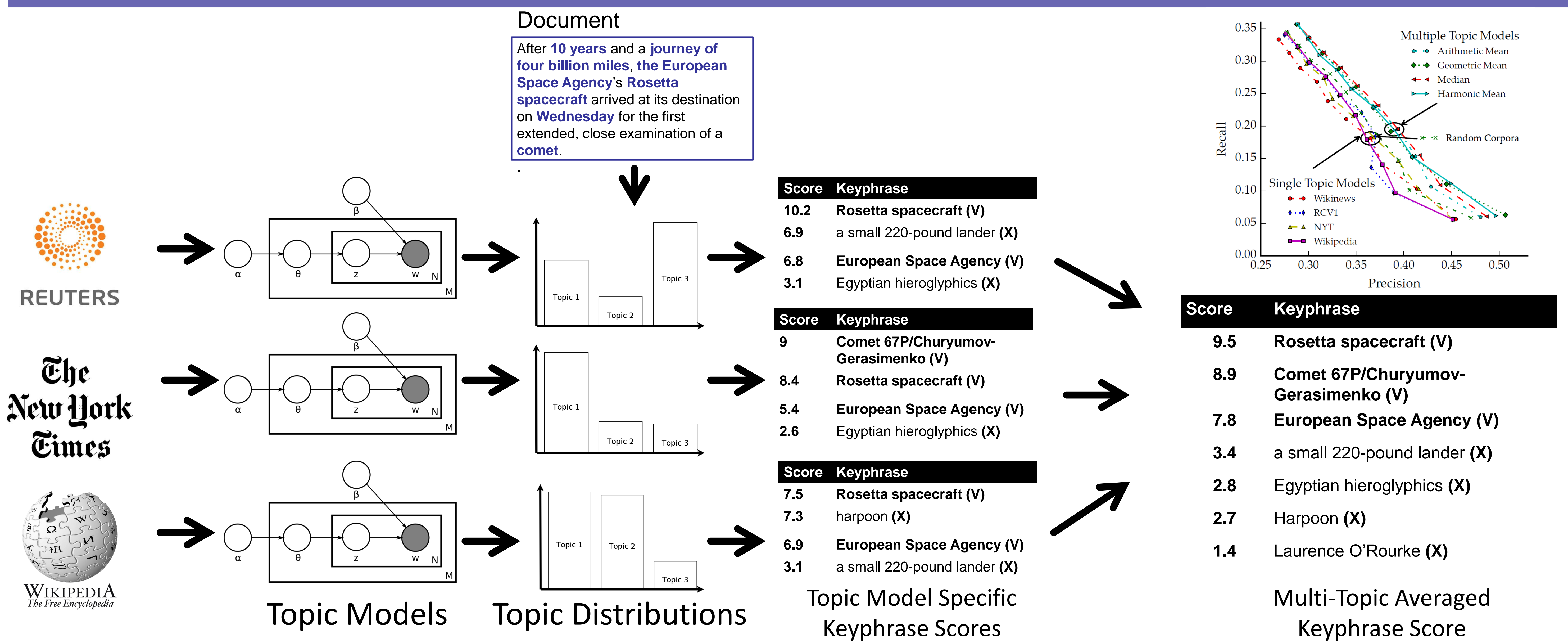
- Automatic selection of important and topical phrases
- Useful for Summarization, Contextual Web Advertisement, classification
- **We demonstrate how we can improve the accuracy of Topical Keyphrases by combining information from multiple topic models.**

Disagreement by Topic Models

- Standard deviations of different weights are shown for topic models from shuffled corpora and for original corpora, for each word in the documents of a test-corpus.
- Substantially higher variance in the importance of the words between models when trained on the specific contexts of documents from the original collections
- Different topic models trained on corpora with distinct contexts, used in a Topic Keyphrase Extraction Algorithm, will produce very different word scores and thus keyphrases, whereas topic models trained on more uniform contexts lead to similar keyphrase rankings

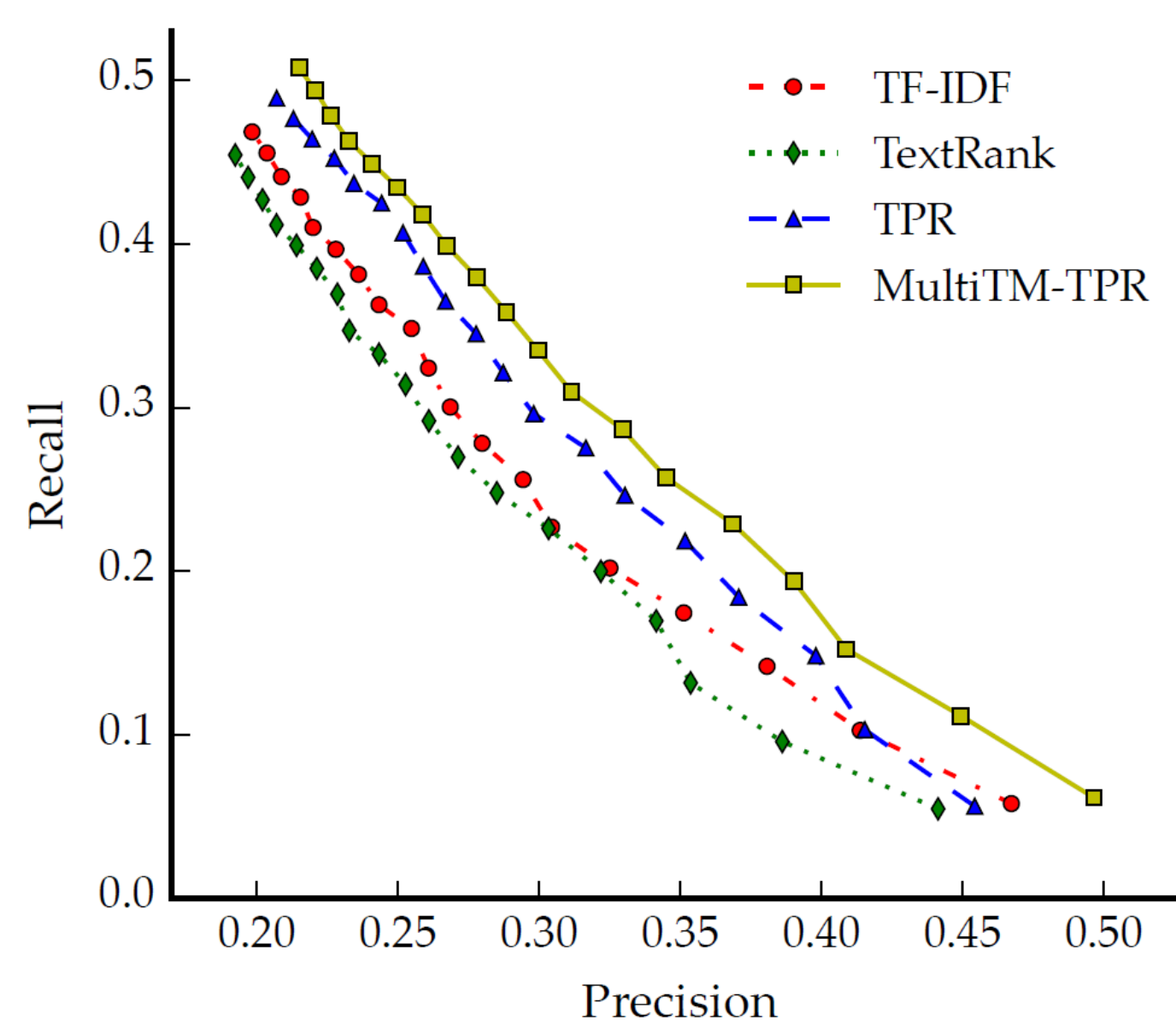


Bagging Topical Word Importance



Leveraging Disagreement

We leverage topic model disagreement by composing **word scores that reflect a more realistic importance of the words**. For this purpose we apply several metrics which combine all weights into a single weight to be used in a Word Graph algorithm biased according to Topical Word importance. When averaging scores generated from topic models from these original corpora, a change in accuracy is noticed. For each separate combination between different topic models some increase is obtained. This increase in accuracy is not observed when randomizing the contexts of the different topic models.



Conclusion

- **Ongoing work** demonstrating the benefit of combining multiple topic models
- Study of the influence of the corpus the topic model is trained on, show **disagreement between models** that are trained on different corpora
- **Averaging weights** from several topic models leads to an increase in precision of extracted phrases
- The difference in contexts between the corpora leads to different topic models and thus disagreement about word importance. We use that disagreement by computing a combined topical word importance value. Better performance is attained only when the **models differ substantially**
- **For future work**, we intend to research whether more sophisticated methods for combining or selection of specific models can be applied