Blendle: Personalized news selection

Participants: Hans Bos, Dror Guldin, Claudia Hauff, Daan Odijk (industry representative), Reza Aditya Permadi, Sihang Qiu, Emily Sullivan Mumm, Nava Tintarev (academic leader).

Context and problem

We looked into the challenges of filter bubbles and recommendation bias, aiming to provide a balanced, diverse, and relevant selection of news clearly that helps with discovering quality journalism.

A new generation is growing up that do not get their news from a newspaper. According to a study from 2016 indicates that social media has overtaken television as young people's main source of news. These readers have a hard time finding quality; for example, Twitter offers a noisy stream of information, Facebook does not optimise for quality, but for clicks and is increasingly flooded with fake news. In some cases this means that users are unaware of opposing points of view, so called `filter bubbles', leading to informational blindspots, and potentially more polarized opinions. Such filter bubbles are partially due to recommender systems which through personalization automatically select and filter content for users.

Recommender systems support people in making decisions about what to consume next, they propose and evaluate options while involving their human users in the decision-making process. Recommender systems also shape our opinions by automatically selecting and ranking information for us in online social networks such as Twitter and Facebook. These technologies have in some cases been found to decrease exposure to more diverse points of view, but they also have the potential to increase content diversity. One neglected aspect of the study is how recommender systems can help users manage their information diet .

In this hands-on workshop, we took on the challenges of filter bubbles and recommendation bias by developing methods for complete and personal news selection. With datasets of news articles and associated meta-data from Blendle, two recommender experts and the participants, we developed a better way of finding news.

Research approach

Our work is based on the assumption that recommender systems can select content that is both relevant and diverse, but that we need to improve our methods for selecting diverse content. The motivating use case is a user who is focused on certain topics in a given time frame, unaware of related topics which are perceived as important by other groups. The user goes to the system which automatically lists topics that are highly relevant that week (Figure 1).

Figure 1: Automated topic suggestion.

This use case relates to the idea of *agenda-setting* in political communication. Agenda-setting is the role of news media which relates to its power "to focus public attention on a few key public issues". Thus, "readers and viewers learn how much importance to attach to a topic on the basis of the emphasis placed on it in the news". When a user selects on of the current topics, the system can suggest recommendations that are representative of a diversity of view-points, resulting in a wider *framing*. The user also has some controls to modify the diversity in the list of recommendations.

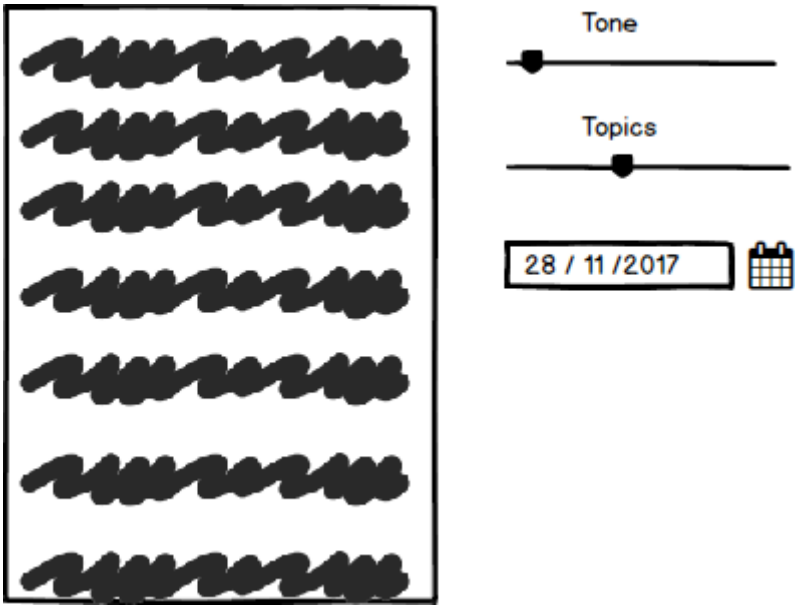

Figure 2: Diversified list of articles, with user controls.

To solve enable this use-case we needed to answer a number of research questions, Figure 3 operationalizes the workflow of the system in a more formal way.

- Which item features help us represent diversity in a way that reflect a number of viewpoints on the same topic?
- Which way should we combine these into a distance function?
- How do we translate the distance function to a ranked list, or which ranking function would be suitable to use?
- What is “optimal” diversity or coverage?

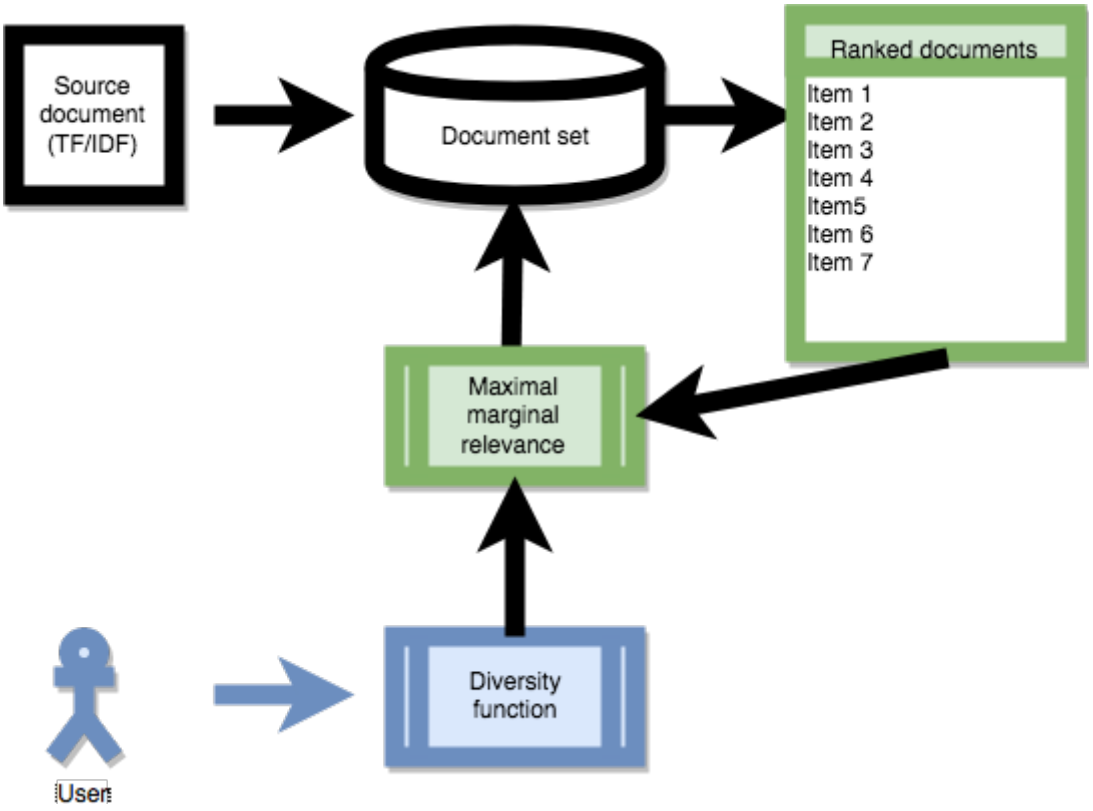

Figure 3: System workflow

Results

Building on the LIWC tool and Blendle's rich item metadata we identified a number of candidate features to aid diversification. We shortlisted them into a shorter list of features that are highly distributed and discriminating. These features were studied in several test corpora of different sizes and on different topics. We then combined the features as a weighted average into a single distance measure.

The combination was evaluated offline in a gridsearch in terms of their ability to diversity topics in terms of source and channel coverage. Our engineered weighting of features outperformed this automated maximum, as it was more coarse grained. However, the performance relative weightings are comparable.

This does not give us an indication of whether the "optimal parameters" were also optimal from a user perspective. To address this, we also designed several user studies to evaluate user acceptance of items in diversified lists.

The result of the workshop was BOA (Blendle Optimised A-list), an end-to-end system incorporating all the components above, including a user interface suitable for experimentation with end-users (see Figure 4). The LIWC tool and parameters were useful for Blendle and can be directly applied with impact in industry.

Future work

We plan to continue this work on a number of tracks including the fine tuning of the definition of some of the diversification features, and conducting more of the designed experiments. There are several mechanisms for collaboration that would interesting for us to pursue such as joint internships, MSc projects, and an application to NWO grant Industrial Doctorates.

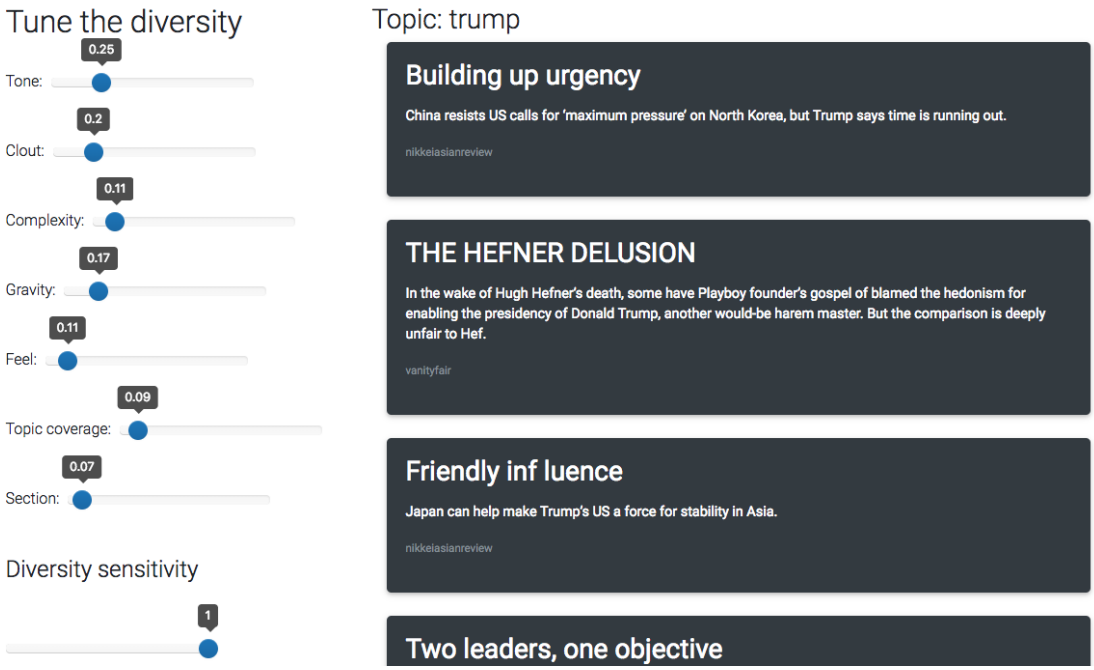

Figure 4: Screenshot of the complete system (BOA) and resulting headlines.