

## Essential Readings: News Media Recommendation Methods

Silvan Hungersbühler & Haukur Jónsson & Grzegorz Lisowski  
& Max Rapp

01/06/2017

# Applying ComSoc in a News Media Setting

We want to devise a transparent, principled method to select a common core of issues which is broadly acceptable to a diverse readership.

# Motivation: A problem with the News Media

- People's biases affect their media content consumption.
- Media tend to feed their consumers' bias.
- This can entrench division and erode the commonly accepted factual base.
- Conversely, if the serious media do not take their readers' concerns seriously, they might lose some part of their readership to less credible sources.

# Scenario I: An Essential Readings News Recommendation Algorithm

Online social networks often recommend news articles to their users based on an algorithm that infers preferences from user's past behavior and demographic properties. Such a News Feed may have little overlap for people belonging to different social or political spheres.

Imagine a news aggregator that asks users for an amount of time they want to read every day, scrapes their social network accounts and based on this information creates a list of essential readings of the desired length.

## Scenario II: A Newspaper Page

A newspaper with a diverse readership has recently lost readers on one side of the political spectrum who feel that the medium does not report on the issues they care about. Responding to this criticism they would like to create a page that reflects preference data they collect from their readers.

How can they fill the page in a way that is a good compromise given their voters diverse preferences over issues?

# Outline

- 1 **Formal Framework:** where we adapt the usual ComSoc machinery to our setting.
- 2 **Desirable Properties of Recommendation Sets:** where we discuss several desiderata our recommendation sets should meet.
- 3 **Budgeted Voting Rules:** where we suggest several voting rules for our setting.
- 4 **Methods:** where we explain our methodology.
- 5 **Results:** where we present our findings and give an outlook on future work.

# Outline

- 1 Formal Framework
- 2 Desirable Properties of Recommendation Sets
- 3 Budgeted Voting Rules
- 4 Methods
- 5 Results

# Formal Framework

We work in a framework familiar from this course...

- A set of *news items*  $A = \{a_1, \dots, a_m\}$
- A set of *recommended items*  $W \subseteq A$
- A set of *consumers*  $N = \{n_1, \dots, n_n\}$
- A *profile of preferences* over the set of items  $\mathcal{R} \in \mathcal{L}^n$

...which we additionally enrich by...

- A utility function  $u_i : \mathcal{L}^n \times N \times A \rightarrow \mathbb{N}$  assigning each news item its Borda score for the preference order of reader  $i$ .
- A *cost function*  $C : A \rightarrow \mathbb{N}$



# Outline

- 1 Formal Framework
- 2 Desirable Properties of Recommendation Sets
- 3 Budgeted Voting Rules
- 4 Methods
- 5 Results

# Axioms

We devised budgeted variants of the following multiwinner axioms:

- Non-Imposition
- Consistency
- Homogeneity
- Monotonicity
- Committee-Monotonicity
- Unanimity

## Reference

Elkind, Edith and Faliszewski, Piotr and Skowron, Piotr and Slinko, Arkadii: "Properties of multiwinner voting rules", *Social Choice and Welfare*, Vol 48, n 3, 2017, pp.599-632.

# Budgeted Utility Maximization

Let  $\mathcal{W}_B$  be the set of all elements of  $\mathcal{P}(A)$  s.t.  $C(W) \leq B$ . Extend the utility function to sets by putting  $u(a) = \sum_{i=1}^n u_i(a)$  and  $u(W) = \sum_{a \in W} u(a)$  for  $W \in \mathcal{P}(A)$ . Utility is maximized if:

$$W = \arg \max_{W' \in \mathcal{W}_B} (u(W'))$$

## $\theta$ -Minority Consistency

A recommendation set has a least general threshold  $\theta$  if

$$a \in W \text{ whenever } \frac{\mathcal{N}_{a \succ b}}{N} \geq \theta \text{ for all } b \in A \setminus \{a\}$$

# Gini coefficient

For each recommendation set we compute the Gini-coefficient  $G(W) \in [0, 1]$  in order to assess the inequality of the utility distribution of  $W$

# Outline

- 1 Formal Framework
- 2 Desirable Properties of Recommendation Sets
- 3 Budgeted Voting Rules**
- 4 Methods
- 5 Results

# Budgeted Utility Maximization (BUM)

This rule takes the voters preference orders and maximizes utility given the budget:

$$\max_{W \in \mathcal{W}_B} u(W)$$

# Least General Budget-compatible Threshold (LGBT)

We define a news item's  $\theta$ -score as follows:

$$\theta(a) = \min_{b \in A \setminus \{a\}} \frac{\mathcal{N}_{a \succ b}}{N}$$

The General Budget-compatible Threshold committee is then chosen as follows:

$$W_1 = \{\arg \max_{a \in A} \theta(a)\} \cap \mathcal{W}_B$$

$$W_n = W_{n-1} \cup \left( \{\arg \max_{a \in A \setminus W_{n-1}} \theta(a)\} \cap \mathcal{W}_{B-C(W_{n-1})} \right)$$

$$W = W_{|A|}$$



## Further Rules

Further we tested budgetized versions of the k-Plurality (BP), k-Borda (BB) and k-Copeland (BC) rules.

# Outline

- 1 Formal Framework
- 2 Desirable Properties of Recommendation Sets
- 3 Budgeted Voting Rules
- 4 Methods**
- 5 Results

## Proofs & Data-Generation

We proved satisfaction or failure of satisfaction of the axioms. In addition we simulated the behaviour of the voting rules with respect to the desirable properties and the failed axioms. For this purpose, since unfortunately good real world data was hard to find, we generated datasets randomly to our needs. We generated one dataset of small profiles with 5000 readers and 10 news items and one dataset of large profiles with 5000 readers and 100 news items. In addition, for each dataset we used various procedures to generate fully random profiles as well as profiles of several degrees of "fragmentization" from 1 to 100 clusters of preference orders.

# Outline

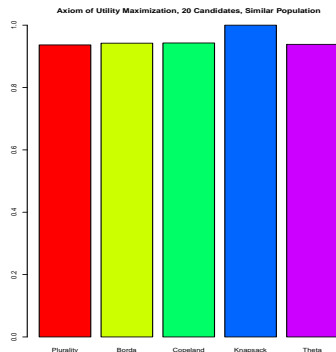
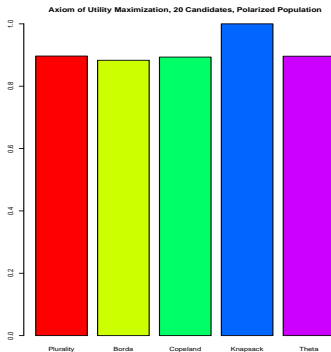
- 1 Formal Framework
- 2 Desirable Properties of Recommendation Sets
- 3 Budgeted Voting Rules
- 4 Methods
- 5 Results

# Axioms

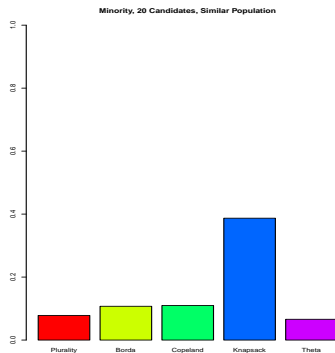
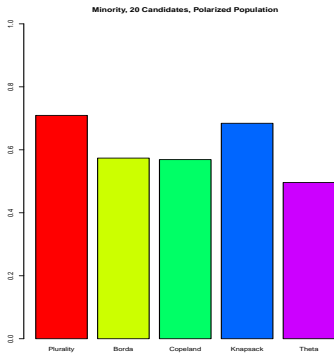
Table 1: Axiom Satisfaction

Axiom \ Rule	BP	BB	BC	BUM	LGBT
Non-Imposition	×	×	×	×	×
Consistency	?	?	?	?	?
Homogeneity	✓	✓	✓	✓	✓
Monotonicity	✓	✓	✓	✓	✓
Committee-Monotonicity	×	×	×	×	×
Unanimity	×	✓	✓	×	✓

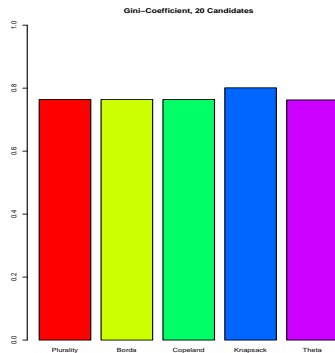
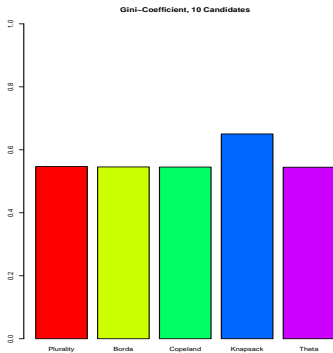
# Budgeted Utility Maximization



# $\theta$ -Minority Consistency



# $\delta$ -Equality





## Summary of experimental data

- No rule can be treated as optimal with respect to all desirable properties.
- The BUM rule deals the best in terms of maximizing utilities on the global level.
- LGBT rule deals much better in terms of *theta*-Minority Consistency and  $\delta$ -Equality which concern the equality of users.

# Future Work

- The cost and utility functions could be used to express dependencies (some items form series of articles, cost might vary depending on reader)
- Implement slack: in many application leftover budget also has a utility
- Find, collect or mine real data (webscraping...).
- Build a ready-to-use implementation!

# Conclusions

- ➊ **Overview:** We presented a formal Social Choice framework to give a recommendation of news items.
- ➋ **Desirable Properties of Recommendation Sets:** We adjusted multiwinner axioms and devised further desirable criteria.
- ➌ **Budgeted Voting Rules:** We devised new rules and extended three multiwinner voting rules.
- ➍ **Methods:** We proved axiomatic results and ran simulations.
- ➎ **Results:**
  - There is a trade-off between maximizing utility and equal utility distribution.
  - The LGBT-Rule favours minority groups, while BUM leads to maximal total utility.
  - In polarized populations, minority groups tend to fare worse and maximal total utility is sub-optimal.
  - More news-items tend to increase utility inequality.