

# An Empirical Analysis of Pruning Techniques

Performance, Retrievability and Bias

Ruey-Cheng Chen, Leif Azzopardi, Falk Scholer



## What is the Relationship between Retrieval Effectiveness, Efficiency and Bias?

### Research Questions

When an inverted index is optimized...

- How does the retrieval bias of an IR system change?
- What is the relationship between performance and bias?

### What is Retrieval Bias?

Retrievability  $r$  of a document  $d$  with respect to the configuration of an IR system is defined as (Azzopardi and Vinay, 2008):

$$r(d) \propto \sum_{q \in Q} \mathbb{I}(k_{dq} \leq c) \times 1/k_{dq}^{\beta}$$

where  $k_{dq}$  is the rank at which  $d$  is retrieved given  $q$ ,  $c$  is a predefined threshold, and  $\beta$  is a hyperparameter. The Gini Coefficient is used to measure the retrieval bias of the system on the population of documents.

### Experiment Setup

#### Retrieval System

- Indri index over GOV2 (25M docs/39M unique terms)
- All documents Krovetz-stemmed and stopwords removed
- Retrieval using optimized BM25:  $k_1 = 0.9$  and  $b = 0.4$
- Effectiveness measured over TREC Topics 701–850

#### Static Index Pruning Methods

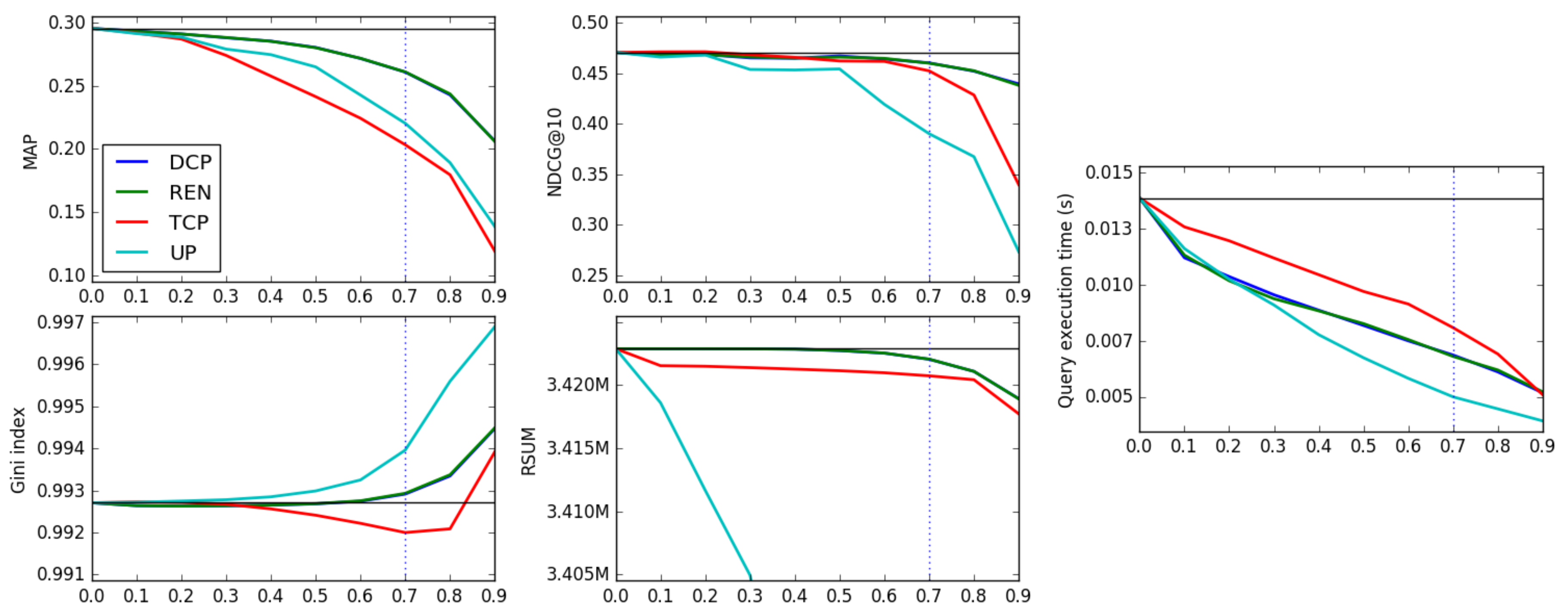
- TCP: Term-based pruning (Carmel et al., 2001)
- DCP: Document-centric pruning (Büttcher and Clarke, 2006)
- UP: Uniform pruning (Carmel et al., 2001; Chen and Lee, 2013)
- REN: Rényi divergence of order infinity (Chen et al, 2015)

Comparisons are made between prune ratios 0.1, 0.2, ..., 0.9.

## Main Results

### Relationship between Performance, Efficiency and Retrieval Bias

- A lower Gini score indicates less bias, and a lower RSUM score (i.e.  $\sum_d r(d)$ ) indicates that less documents are retrievable.
- Bias remains stable until a turning point, between prune ratio 0.3–0.7 depending on the pruning algorithm, after which bias increases.
- DCP and REN appear to give a better trade-off than TCP.
- Selecting prune ratio based on the Gini score would result in good pruning performance without a sizable loss in early precision.



### Interaction between Performance and Retrieval Bias

- The star indicates the “starting point” (un-pruned index), and each subsequent point corresponds to an 0.1 increase in the prune ratio.
- For UP, DCP and REN, performance tends to improve as bias decreases (i.e. less bias  $\Rightarrow$  better performance), but for TCP the relationship appears more complex.

