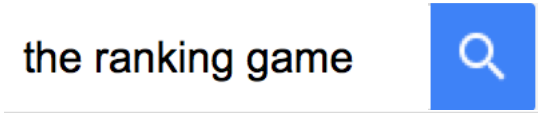
THE RANKING GAME

Ran Ben-Basat **Elad Kravi**

Computer Science faculty, Technion, Israel

A SEARCH ON THE WEB



GameRankings - Video Game Reviews from around the Internet
www.**gamerankings**.com/
Founded in 1999, GameRankings indexes over 240000 video **game** reviews from both online and offline sources, plus over 230000 other news articles.

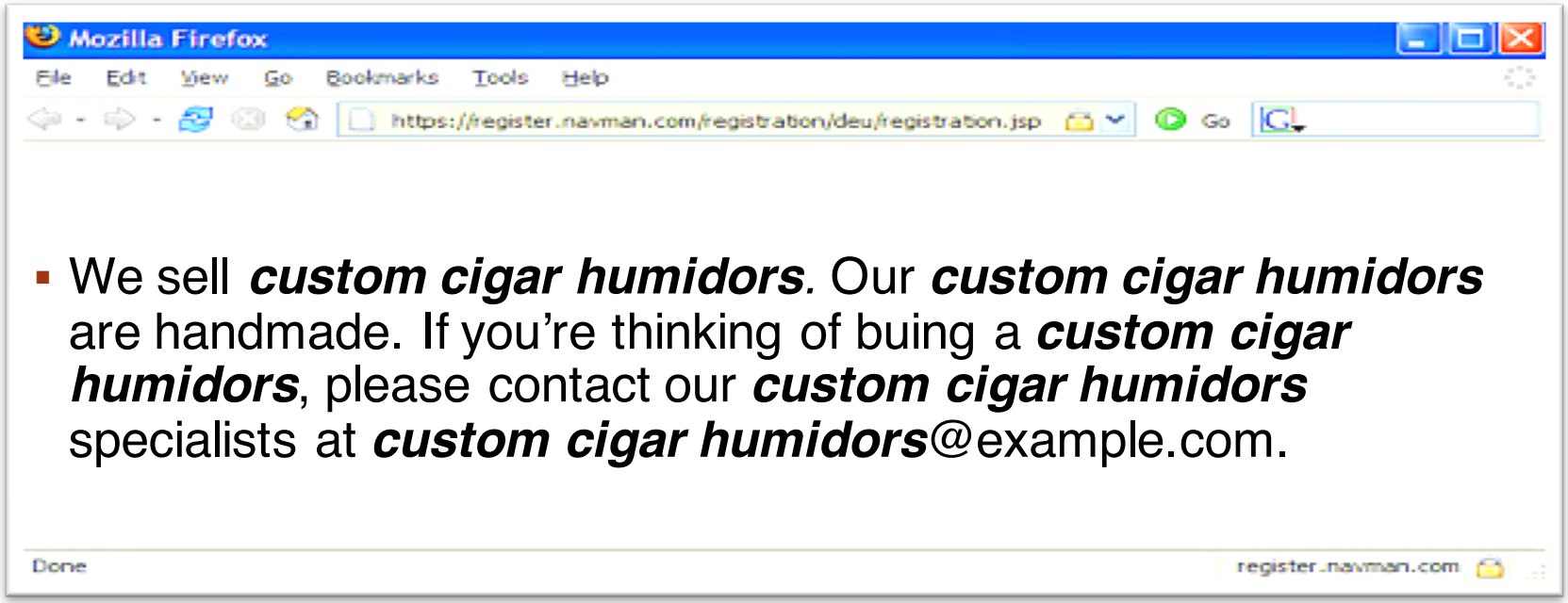
College Rankings - Reed College
https://www.reed.edu/apply/college-**rankings**.html
College Rankings. campus banner photo. Reed and **the Rankings Game**. Why doesn't Reed participate in U.S. News & World Report's college rankings?

[PDF] The University Rankings Game - Lehigh University
www.lehigh.edu/~incbeug/.../Dearden_Grewal_Lilien_**Rankings_Game**_May_2008.p...
by R Grewal - 2008 - Cited by 52 - Related articles
The University **Rankings Game**: Modeling the Competition among Universities for Ranking. Rajdeep Grewal. James A. Dearden. Gary L. Lilien. May 2008 ...

Gaming the College Rankings - The New York Times
www.nytimes.com/2012/02/01/education/**gaming**-the-college-**rankings**.html
Feb 1, 2012 - Any love-hate relationship must have its share of pain, so the academic world, in its obsession with college **rankings**, is suitably dismayed by ...

Top Core PC Games | US & Europe | Newzoo
https://newzoo.com/insights/**rankings**/top-20-core-pc-**games**/
Every month Newzoo and Overwolf publish the Top 20 Core PC **Games** in the US and Europe. **The ranking** is based on the number of unique sessions during a ...

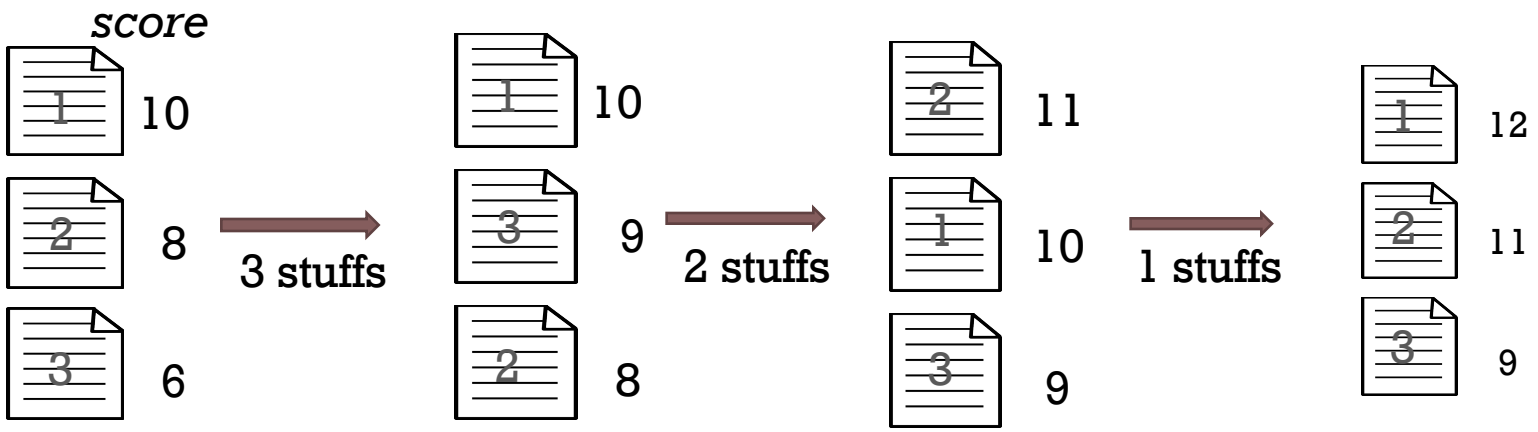
KEYWORD STUFFING



- We sell ***custom cigar humidors***. Our ***custom cigar humidors*** are handmade. If you're thinking of buing a ***custom cigar humidors***, please contact our ***custom cigar humidors*** specialists at ***custom cigar humidors***@example.com.

- One approach for *Search Engine Optimization (SEO)*

EXAMPLE


score
1
10
2
8
3
6
3 stuffs
1
10
3
9
2
8
2 stuffs
2
11
1
10
3
9
1 stuffs
1
12
2
11
3
9

PROBLEMS

- Search engine's performance worsens due to the competition
- Degraded coherency
- Hard to distinguish between white/black hat SEO

Goal:

- Can the search engine reduce the motivation for "bad" content manipulations?

OUR CONTRIBUTION

- Developed formal methodology to measure the tolerance of a ranking scheme to keyword stuffing
- Showed that competition may degrade retrieval effectiveness
- *Offered a probabilistic ranking scheme* and *showed an indication for it to* reduce the incentive to manipulate the content

INSPIRED BY GAME THEORY

- Documents are *strategic players*
  - **Profit** derived by rank
  - **Cost** by SEO actions
  - **Utility** = Profit - Cost

- Competition is a game:
  - Each document “plays” in its turn – perform the ***Best Response***
  - A round – all the documents have a change to act

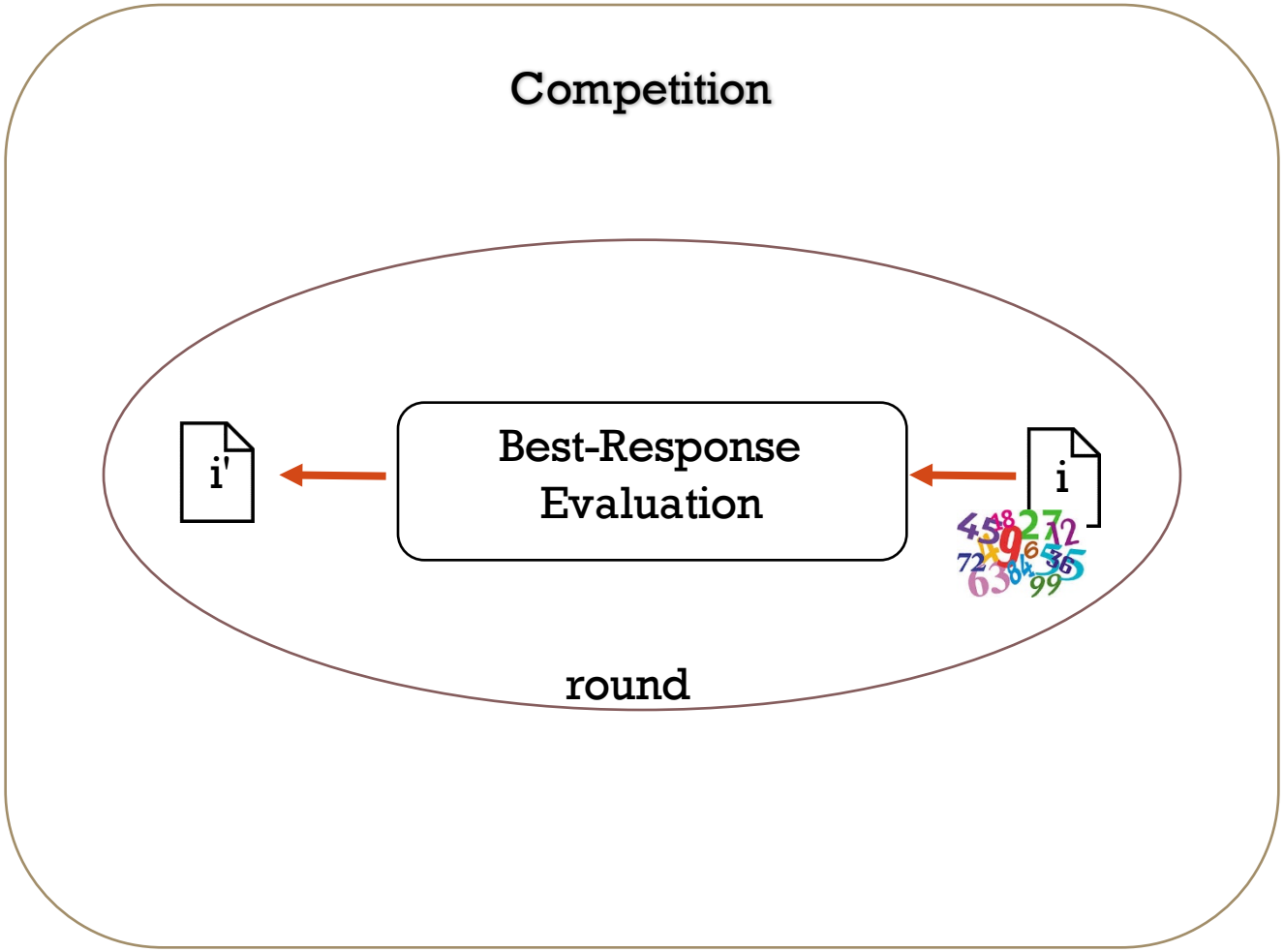
FORMAL MODEL

- Given a query $q \in \mathcal{Q}$, and a set of documents $\mathcal{D}$
- Ranking function: $\mathcal{D} \times \mathcal{Q} \longrightarrow [0,1]$
  - e.g., **BM25**
- Profit: $P: \{1, \ldots, n\} \longrightarrow \mathbb{R}^+$
  - e.g., **reciprocal ranking** $(p(i) = {}^1/_i)$
- Cost: $C: \mathcal{D} \times \mathcal{D} \longrightarrow \mathbb{R}^+$
  - representing cost of SEO
  - e.g., **fixed cost per stuffed term**
- Utility: $U_i(d_i') = E(P(r(d_i'))) - C(d_i', d_i)$
  - expectancy is explained later on

ASSUMPTIONS

- Ranking function is known to all the documents
- No content deletions
- Only query terms are stuffed
- Documents optimize for a single query

SIMULATION



query

documents

modified
documents

BEST RESPONSE

- A document should iterate over all possible modifications
  - however, this number is large
  - when limiting the number of stuffed terms to k
    - $CC_n^k = \binom{n+k+1}{n}$ where there are n query terms
    - still large…

BEST RESPONSE

- **Greedy algorithm:**

  iteratively add the term which locally maximizes the Ranking score

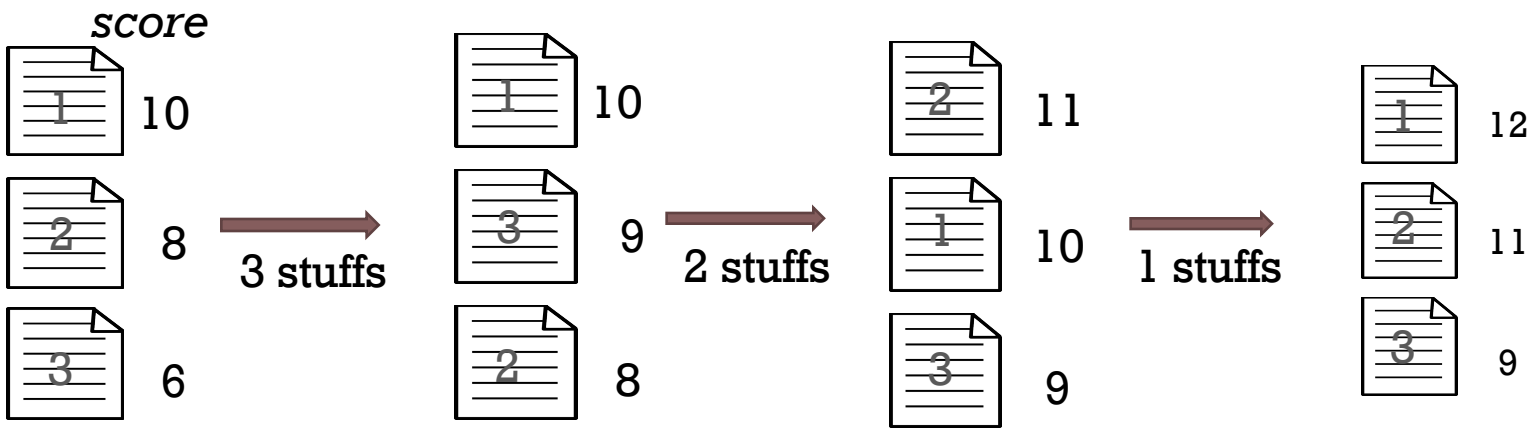
  - *eventually* return the revised document or the original one if no improvement was found

AN EXAMPLE OF COMPETITION

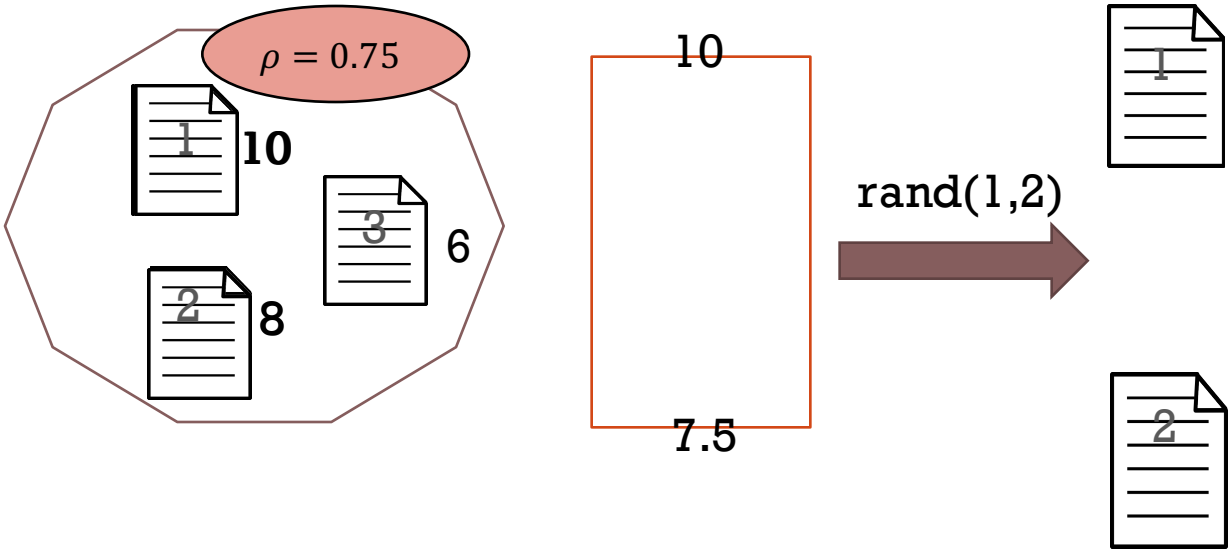
- $\Sigma = \{a, b, c, d\},\ q = \{a, b\},\ d_1 = \{a\}, d_2 = \{b, c\}$
- RSV – *Query Likelihood* with *Laplace smoothing*
- P(1)=1, P(2)=0, C:0.75 per term
- $\mathrm{RSV}(d_1, q) = \frac{2}{5} * \frac{1}{5} = \frac{2}{25}$, $\mathrm{RSV}(d_2, q) = \frac{1}{6} * \frac{2}{6} = \frac{2}{36}$
- $d_2' = \{\mathrm{a}, \mathrm{b}, \mathrm{c}\} \Longrightarrow \mathrm{RSV}(d_2', q) = \frac{2}{7} * \frac{2}{7} = \frac{4}{49} > \frac{2}{25}$
- $d_1' = \{\mathrm{a}, \mathrm{b}\} \Longrightarrow \mathrm{RSV}(d_1', q) = \frac{2}{6} * \frac{2}{6} = \frac{4}{36} > \frac{4}{49}$

- and the second player cannot improve its utility from stuffing

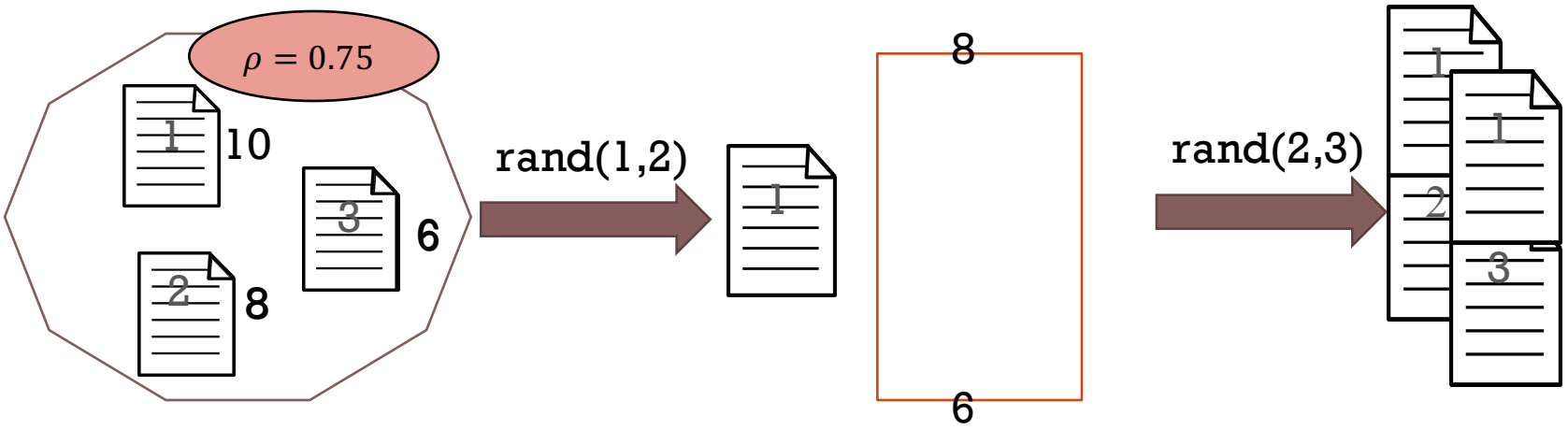
DETERMINISTIC RANKING



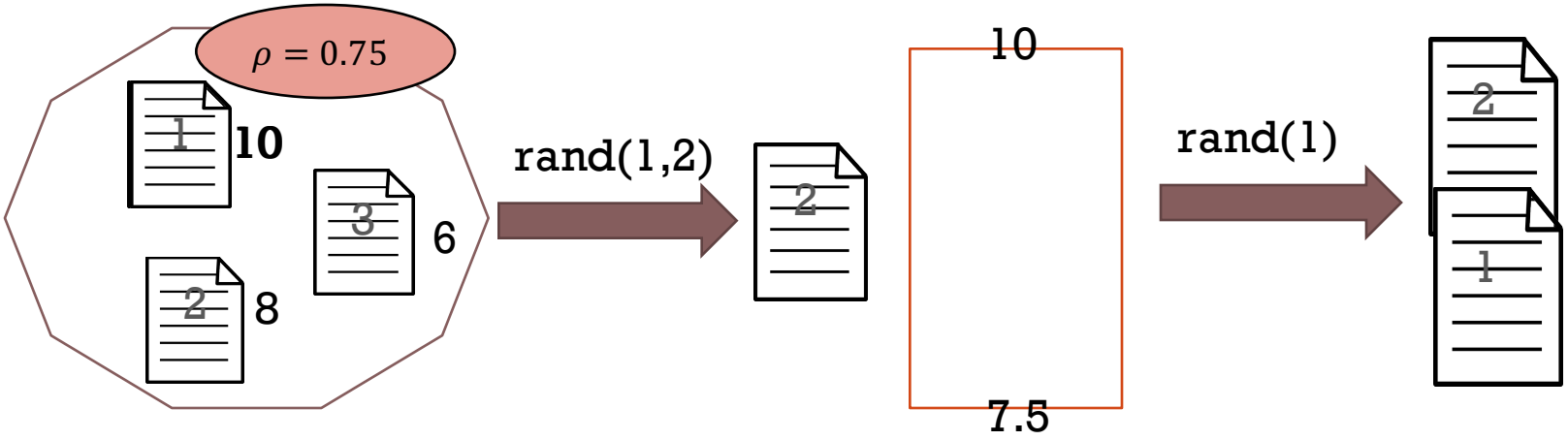
PROBABILISTIC RANKING



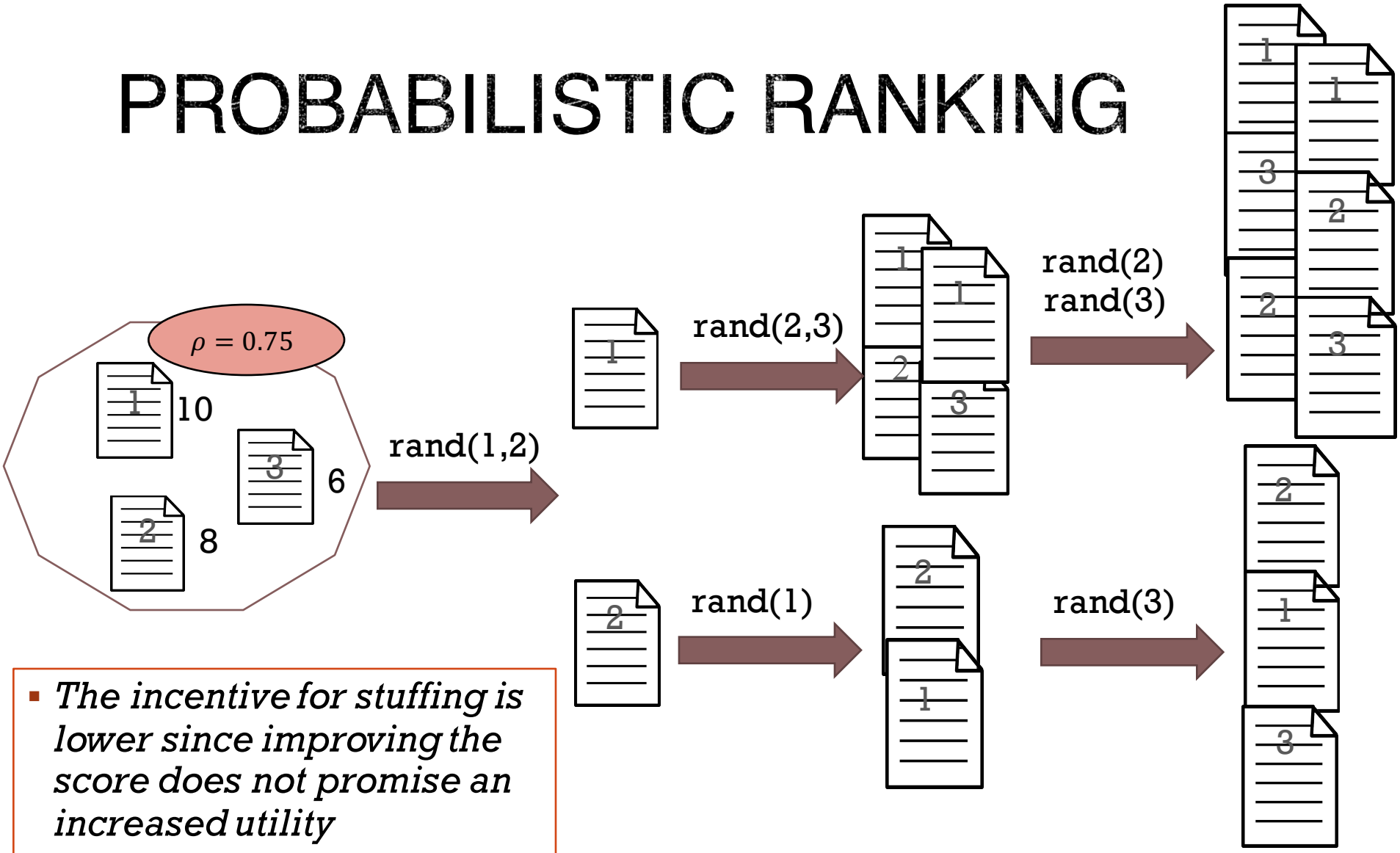
PROBABILISTIC RANKING



PROBABILISTIC RANKING



PROBABILISTIC RANKING



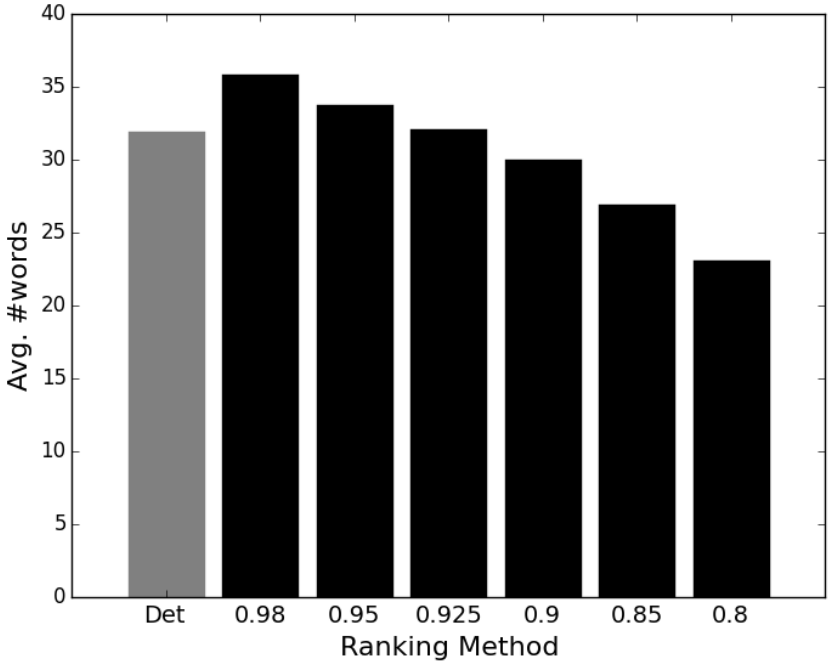
- *The incentive for stuffing is lower since improving the score does not promise an increased utility*

DATASET

- AP news & Wall Street Journal
- Including TREC's relevance judgments (79 queries)

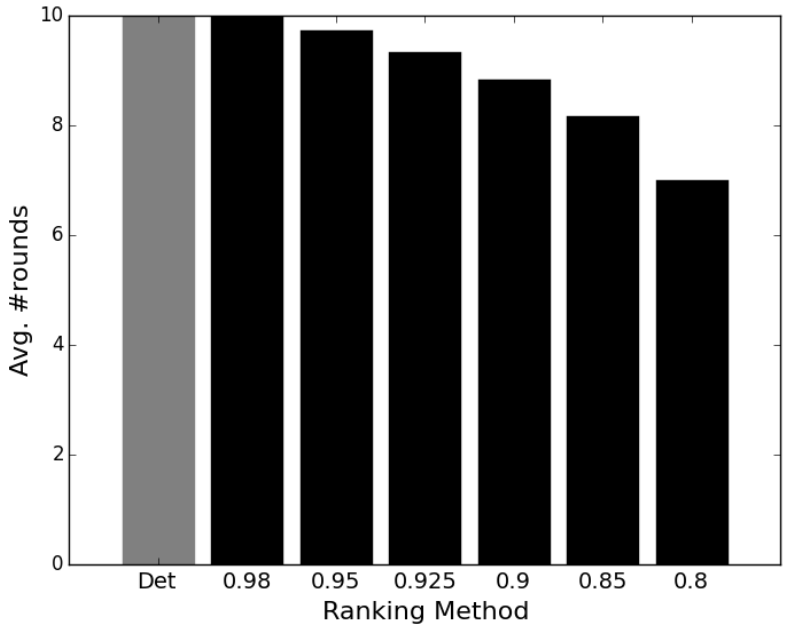
- Settings:
  - *profit* – **reciprocal ranking** ($p(i) = {}^{1}/_{i}$)
  - ***fixed cost*** *– 0.05 per term*
  - ***20 players*** *(documents)*
  - *up to* ***10 rounds***
  - ***BM25*** *as a deterministic ranking function*
  - *Probabilistic ranking – using* ***10k Monte-Carlo simulations***

KEYWORD STUFFING EXPERIMENT



- ρ is decreasing $\longrightarrow$ stuffing is less profitable
- In high values of ρ highly-ranked pages stuff terms

CONVERGENCE EXPERIMENT



- Randomness in the ranking process limits the competition between web pages

RELEVANCE EXPERIMENT

No Competition	P@3	P@5	P@10	MAP[1]
Initial Set	**0.5105**	**0.4835**	**0.4734**	**0.0655**
Det. Ranking	0.4557	0.4481	0.4341	0.0591

- Decrease in relevance across all metrics due to the competition
  - The ranks of irrelevant pages improve throughout the competition
- Due to small corpus size - not significant statistically

[1]values are low since the MAP was calculate over 20 documents

RELEVANCE EXPERIMENT

	P@3	P@5	P@10	MAP[1]
Initial Set	**0.5105**	**0.4835**	**0.4734**	**0.0655**
Det.	0.4557	0.4481	0.4341	0.0591
Prob. (0.9)	0.4492	0.4490	0.4449	0.0590

- Decrease in relevance across all metrics due to the competition
  - Irrelevant pages ranks improve throughout the competition
- Due to small corpus size - not significant statistically

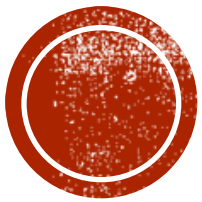
[1]values are low since the MAP was calculate over 20 documents

CONCLUSIONS

- We built a methodology for estimating techniques for handling web spam
- Inspired by game theory pages are strategic players with profit, cost and utility
- Illustrated the system on probabilistic ranking, and show it reduces spam with limited implication on search quality

FUTURE WORK

- Additional and bigger datasets
  - Including web pages
  - Optimized for SEO
- Comparison to other spam-aware methods
- Extend the framework to other SEO operations e.g., link farming
  - Can the probabilistic approach reduce manipulations?



THANK YOU!