

1. Background

- RBO $RBO_{S,L,p} = \frac{1-p}{p} \sum_{d=1}^{\infty} A_{S,L,d} \cdot p^d$
- P = Persistence - [0,1]
- Higher p => less top-weighted
- Lower p => more top-weighted
- Seen ranking / Unseen ranking
- Ties
- W-variant - gives each item in a tied group the top rank of that group
- A-variant - computes the average contribution of an item in a tied group for every permutation of this group
- B-variant - "accounts for the amount of information actually available to measure overlap" as explained in [1, Corsi and Urbano p.4].
- RBO_ext = uses the seen part of a ranking to estimate what RBO would be if we also saw the unseen part
- RBO_max = assumes every item in the unseen part will be matched to an item in the other ranking
- RBO_min = assumes every item in the unseen part will not be matched to an item in the other ranking

x	b	(c d)	e	f	g	h	i	j	k	l	m	n	(o p)
z	y	x	w	(v u t)	s	e	(g i a)						

Table 1. Example of two rankings where the parentheses represent tied groups and the longer ranking will be referred to by L and the shorter by S

x	b	(c d)	e	f	g	h	i	j	k	l	m	n	o	p
z	y	x	w	(v u t)	s	e	g	i	a					

Table 2. Rankings shown in table 1 where ties are removed after depth 10 for S ranking and depth 11 for L ranking

x	b	(c d)	e	f	g	h	i	j	k					
z	y	x	w	(v u t)	s	e	g							

Table 3. Rankings shown in table 1 but truncated to depth 10 for S ranking and depth 11 for L rankings

Table	W-variant		A-variant		B-variant	
	1	2	1	2	1	2
Depth 10	2/11	1/10	1/15	1/10	69/1000	1/10
Depth 11	4/23	2/11	4/33	2/11	1/8	2/11
Depth 12	2/12	2/12	2/12	2/12	2/12	2/12

Table 1: Shows agreement of rankings in table 1 and 2 at depth 10, 11, and 12 for each variant (only taking item g, i, and a into account)

2. Ties in unseen part

- Using rankings shown in Tables 1 and 2 on the left
- Simulating seen and unseen part
- Using agreement and contribution functions shown below for their respective variants made by Weber et al. [2] and Urbano and Corsi [1]
- X = size of the union of the two rankings
- Table 4 shows the difference in agreement at each depth (Using only items g, i, and a as other items are all seen at those depths so aren't affected by the assumption)
- For each variant, table 2 gives the same agreement showing that when no ties are involved all variants reduce to the agreement function of bare RBO
- At depth 12 all agreements are the same as from this depth the whole tied group is seen.

$$c_{e|d} = \begin{cases} 0 & d < t_e \text{ (inactive)} \\ 1 & b_e \leq d \text{ (active)} \\ \frac{d-t_e+1}{b_e-t_e+1} & \text{otherwise (crossing)} \end{cases}$$

$$A_{S,L,d}^w = \frac{2 * X_{S,L,d}}{|S_{,d}| + |L_{,d}|}$$

$$A_{S,L,d}^a = \frac{1}{d} \sum_{e \in \Omega} c_{e,S|d} * c_{e,L|d}$$

$$A_{S,L,d}^b = \frac{\sum_{e \in \Omega} c_{e,S|d} * c_{e,L|d}}{\sqrt{\sum_{e \in \Omega} c_{e,S|d}^2} * \sqrt{\sum_{e \in \Omega} c_{e,L|d}^2}}$$

3. Methodology

- Simulate rankings using code provided by Urbano and Corsi (<https://github.com/julian-urbano/sigir2024-rbo>)
- P experimentations
 - Length of the ranking between 15 and 20
 - Domain = 1000 (reproduce infinity)
 - Random conjointness
 - Truncated depth between 25% and 50% of the length
 - Enforce ties after this depth
 - Get full ranking (1) (table 1)
 - Get full ranking with no ties after truncation depth (2) (table 2)
 - Get truncated ranking (3) (table 3)
- Extrapolation experimentations
 - Almost the same methodology as p experimentations
 - Different length of 50
 - 65 files where each full ranking is the same
 - For rankings referring to number 2 above where the percentage of unseen items differ per file ranging from 10 to 75
 - The same is used for the truncated rankings

4. Results P

- Shown below, in a scatter plot, is how RBO is affected by the assumption
- Shown is RBO reality (referring to ranking 1 of methodology) and RBO under assumption (referring to ranking 2)
- On average most are around the regression line
- Variant w and b have more outliers than a-variant
- When p tends more to 1 the cloud around the regression is much more dense meaning difference in RBO is much smaller

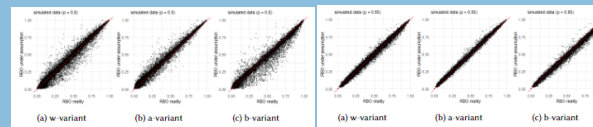
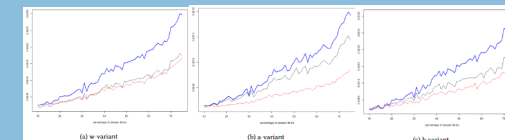


Fig. 1. RBO with p = 0.8 of full ranking and full ranking with no ties after truncation depth. Fig. 2. RBO with p = 0.95 of full ranking and full ranking with no ties after truncation depth.

4. Results Extrapolation

- Shown below are the average difference between RBO's
- The blue line (difference in RBO between table 1 and 3) represents how the current extrapolation performs.
- The red line (difference in RBO between table 1 and 2) represents how with a perfect extrapolation, it would perfectly estimate if an item at a certain rank would match an item in the other ranking, only the assumption affects RBO
- The black line (difference in RBO between table 2 and 3) represents how the loss of information, not knowing which items are in the unseen part, affects the RBO, not factoring in ties but only individual items.
- Comparing the blue and black line gives us a nice overview of how the assumption affects the extrapolation



5. Conclusion

- On average all variants perform well
- Extrapolation also on average but outliers are significant

6. Future Work

- Focus on changing extrapolation
- Not only to better estimate the agreement used for the unseen part
- Try to incorporate possible ties at truncation depth in the formula

7. References

- [1] Matteo Corsi and Julian Urbano. "The Treatment of Ties in Rank-Biased Overlap". In: Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '24), July 14–18, 2024, Washington, DC, USA. ACM, New York, NY, USA, (2024). url: <https://doi.org/10.1145/3626772.3657700>.
- [2] William Webber Alistair Moffat and Justin Zobel. "A Similarity Measure for Indefinite Rankings". In: ACM Transactions on Information Systems 28.4 (2010)