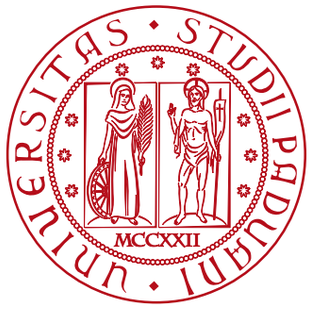


A graphical view of distance between rankings: the Point and Area measures

Giorgio Maria Di Nunzio and Gianmaria Silvello
Dept. of Information Engineering
University of Padua

IIR 2015 - 6th Italian Information Retrieval Workshop

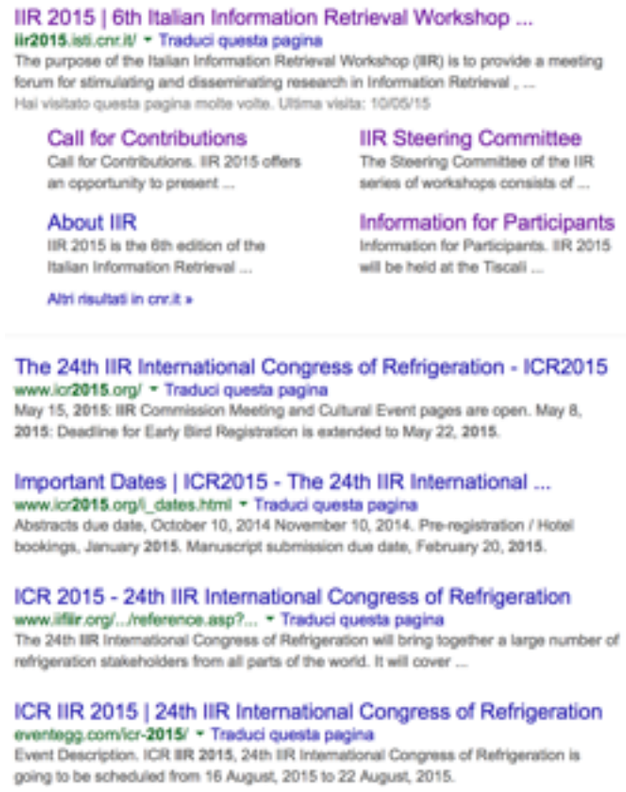
May 25th, 2015, Cagliari, Italy

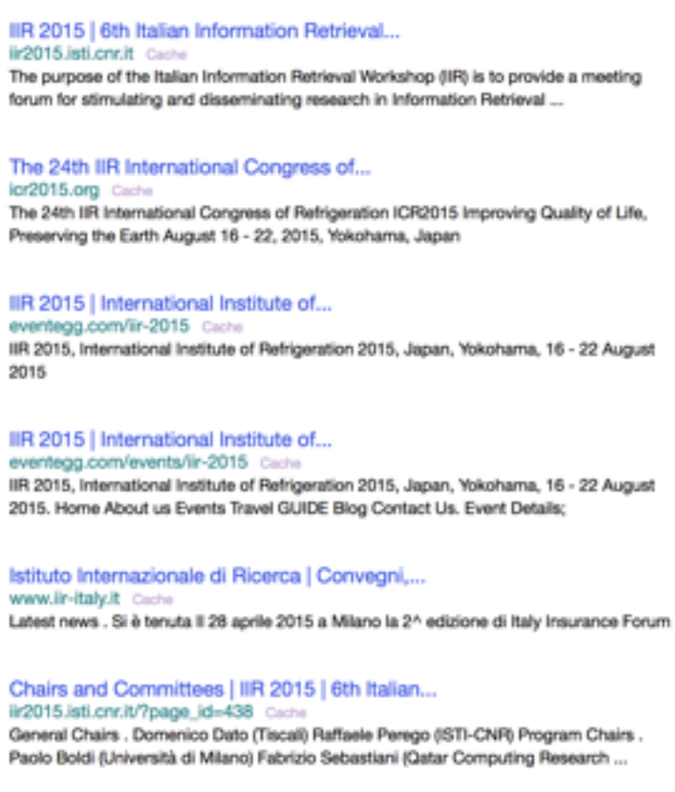
Outline

- Classification of rank similarity measures
- Spearman foot-rule and Kendall distance
- Point and Area measures
- Measures of effectiveness
- Conclusions

Comparing ranked list

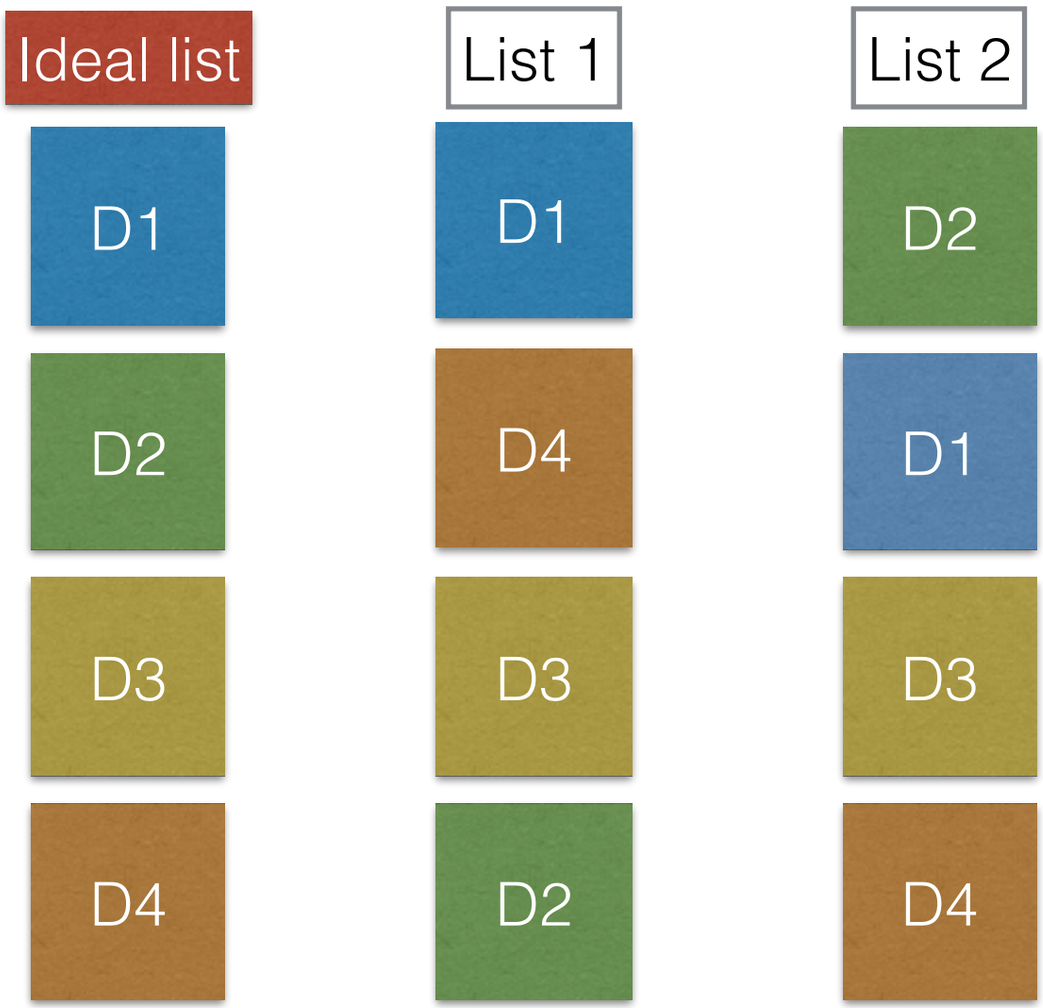
- Search engines effectiveness can be measured by analyzing their visible outcomes
  - lists of documents ranked in descending order of relevance to a given topic





Comparing ranked list

- Correlation among rankings can be used to assess the search engines effectiveness



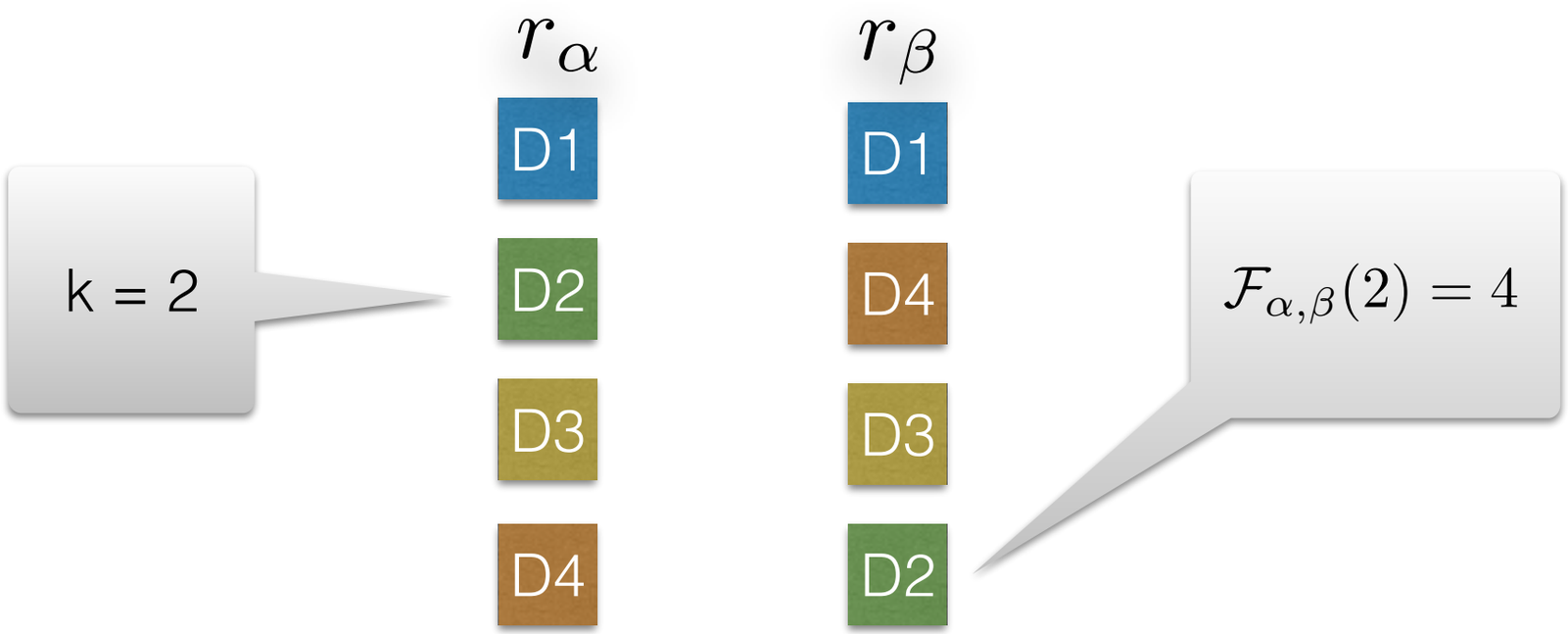
Classification of rank similarity measures

- Weighted / non-weighted
  - Exchanges in the ordering at the top of the ranking are more significant than those at the bottom
  - Any perturbation has the same importance
- Conjoint / non-conjoint
  - Two rankings have the same elements
  - Some elements in one list do not appear in the other

Find elements in a ranked list

- Map index of a document from one list to the other
- Given k-th element in list r_α, return index of that element in list r_β

$$\mathcal{F}_{\alpha,\beta}(k) = idx_\beta(r_\alpha(k))$$

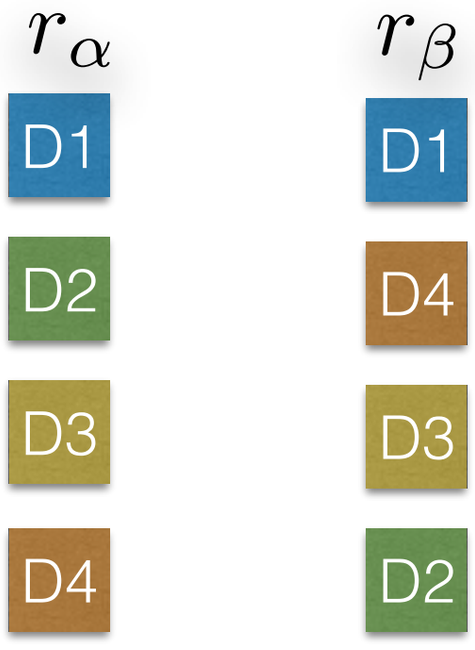


Spearman foot-rule

- Compute the total element-wise misplacements between two ranked lists

- Non-weighted, conjoint

$$S_{\alpha,\beta}(i) = \sum_{k=1}^{i} |\mathcal{F}_{\alpha,\beta}(k) - k|$$



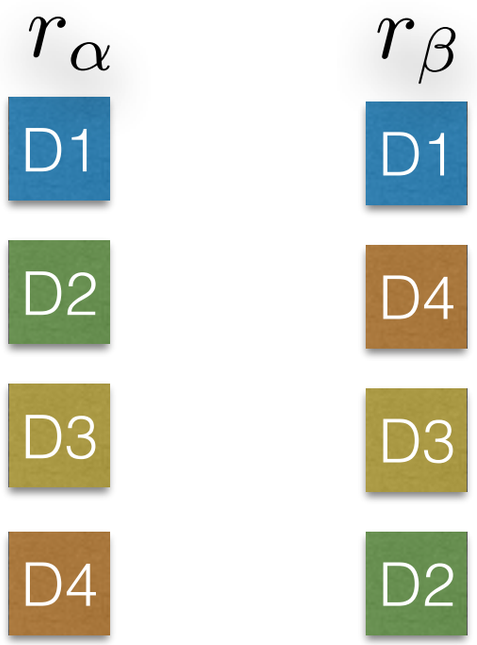
$$S_{\alpha,\beta}(4) = 0 + 2 + 0 + 2$$

Kendall distance

- Number of adjacent swaps that are necessary to reorder one list as the other

- Non-weighted, conjoint

$$K_{\alpha,\beta} = \sum_{(i,j):i<j} \widetilde{K}_{i,j}(r_\alpha, r_\beta)$$



$$K_{\alpha,\beta} = 0 + 2 + 1$$

Spearman and Kendall

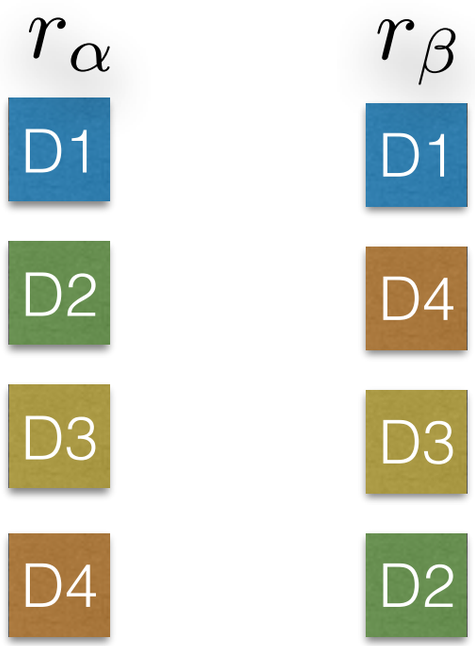
$$S_{\alpha,\beta}(i) = \sum_{k=1}^{i} |\mathcal{F}_{\alpha,\beta}(k) - k|$$

$$K_{\alpha,\beta}(i) = \sum_{k=1}^{i} (\mathcal{F}_{\alpha,\beta}(k) - k) + \underbrace{(r_\alpha[1:k] \cap r_\beta[(\mathcal{F}_{\alpha,\beta}(k)+1):n])}_{X}$$

Point-wise distance

- Spearman without absolute value (!)
- Non-weighted, conjoint

$$P_{\alpha,\beta}(i) = \sum_{k=1}^{i} (\mathcal{F}_{\alpha,\beta}(k) - k)$$



$$P_{\alpha,\beta}(4) = 0 + 2 + 0 - 2$$

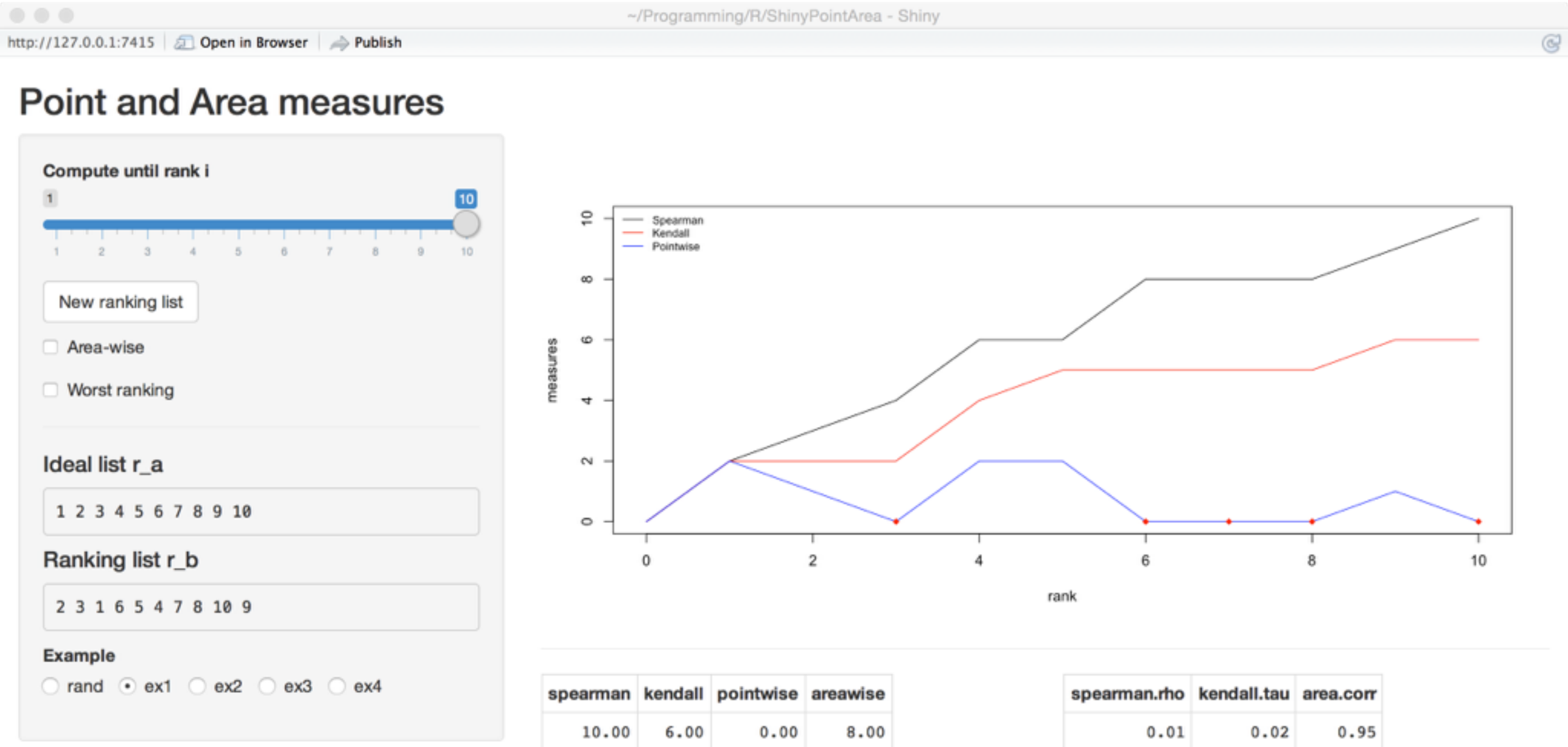
Spearman, Kendall, Point-wise

$$S_{\alpha,\beta}(i) = \sum_{k=1}^{i} |\mathcal{F}_{\alpha,\beta}(k) - k|$$

$$K_{\alpha,\beta}(i) = \sum_{k=1}^{i} (\mathcal{F}_{\alpha,\beta}(k) - k) + \underbrace{(r_\alpha[1:k] \cap r_\beta[(\mathcal{F}_{\alpha,\beta}(k)+1):n])}_{X}$$

$$P_{\alpha,\beta}(i) = \sum_{k=1}^{i} (\mathcal{F}_{\alpha,\beta}(k) - k)$$

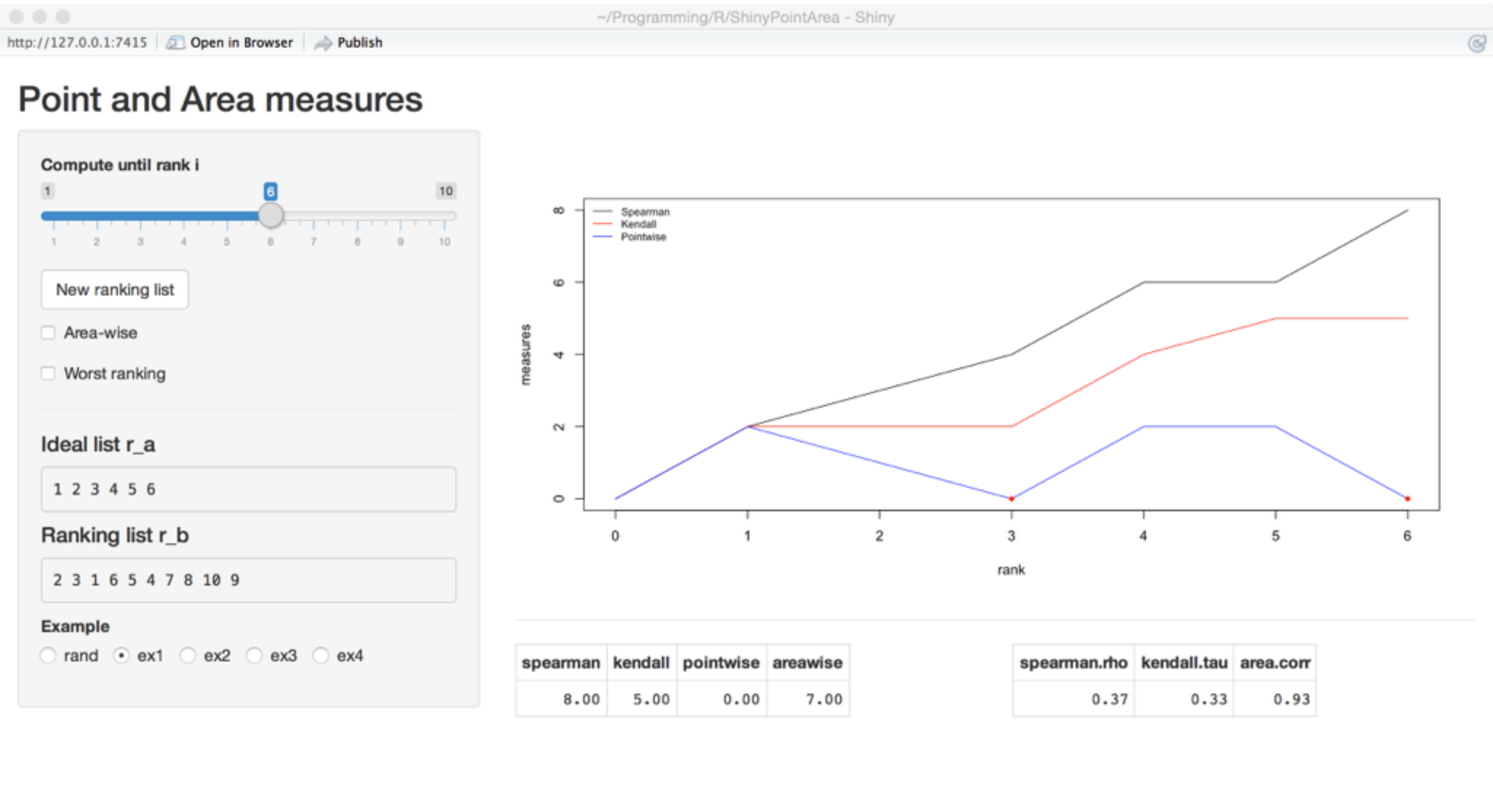
Visualization analysis



spearman	kendall	pointwise	areawise
10.00	6.00	0.00	8.00

spearman.rho	kendall.tau	area.corr
0.01	0.02	0.95

Visualization analysis



Area-wise distance

- The area-wise measure considers the area formed by the segments between two adjacent points (point distance) and the x-axis.

- h is the height of each trapezoid. It can be tuned to weight misplacements that occur in different part of the ranking list.

$$A_{\alpha,\beta}(i) = \sum_{k=1}^{i} \frac{\left(P_{\alpha,\beta}(k-1) + P_{\alpha,\beta}(k)\right)}{2} h$$

Area-wise normalization

- Divide the area of a relevance list at rank i by the largest obtainable area given by the worst possible ranking.

- It is in the [0,1] range, where 0 indicates the ideal case and 1 the worst case

$$\mathrm{nA}_{\alpha,\beta} = \frac{A_{\alpha,\beta}}{A^*_{\alpha,\beta}}$$

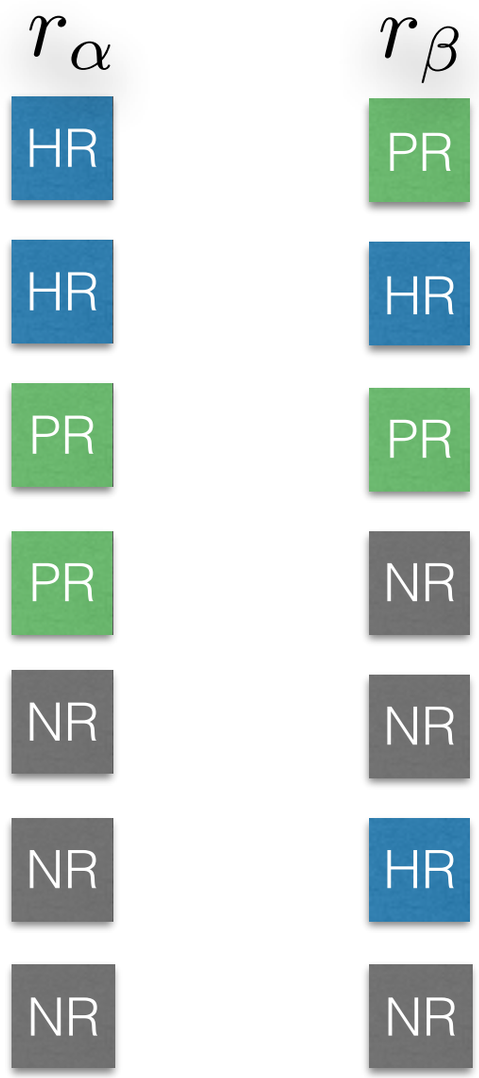
Area Correlation (A-corr)

- A-corr is an indicator of the correlation between two ranked lists
- It is in the range [0,1], where 0 indicates that two lists are not correlated and 1 that they are the same

$$\text{A-corr}_{\alpha,\beta} = 1 - \text{nA}_{\alpha,\beta}$$

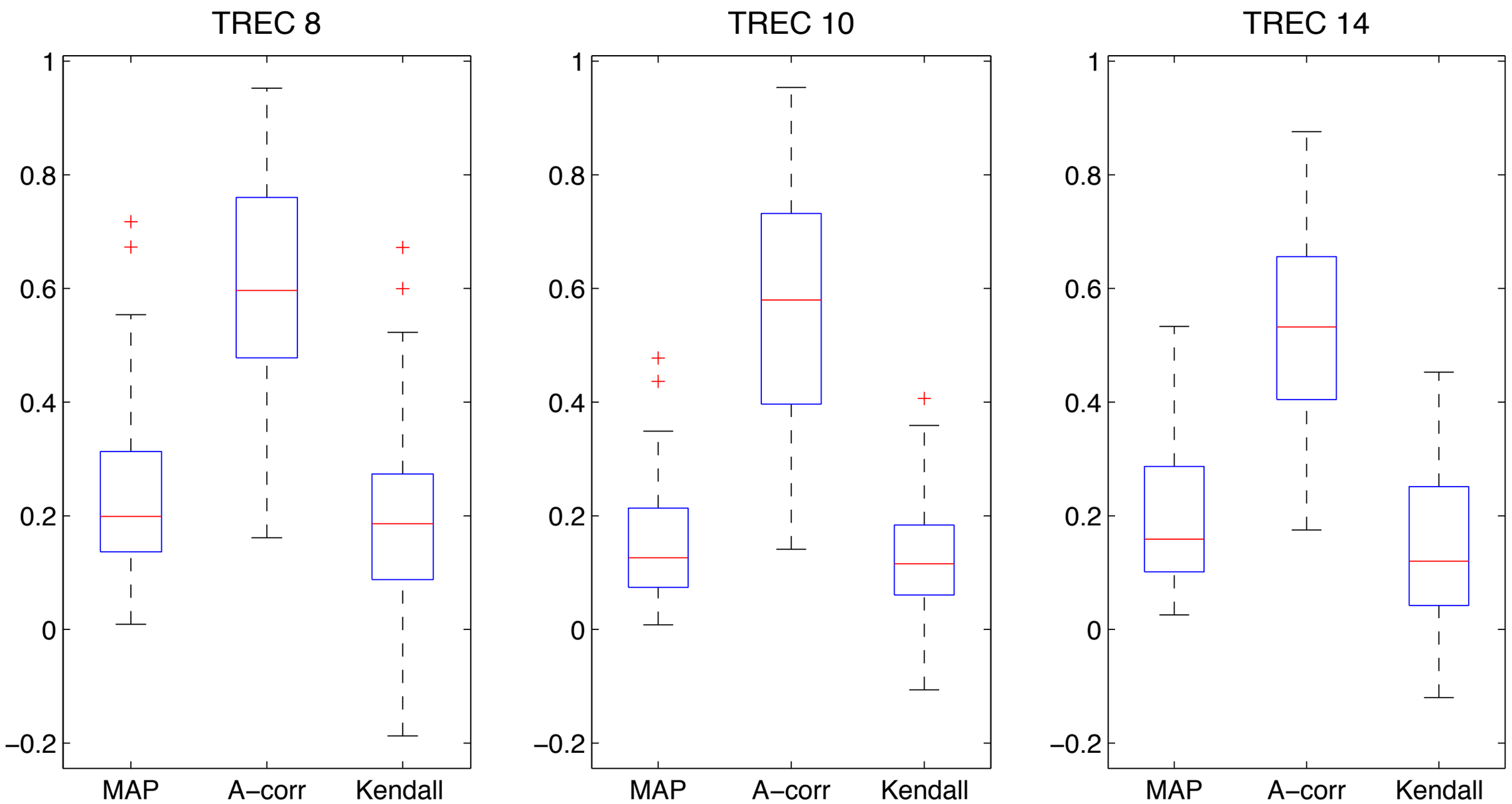
A-corr as an effectiveness measure (quantitative)

- We can calculate the point-wise measure by considering the relevance of documents

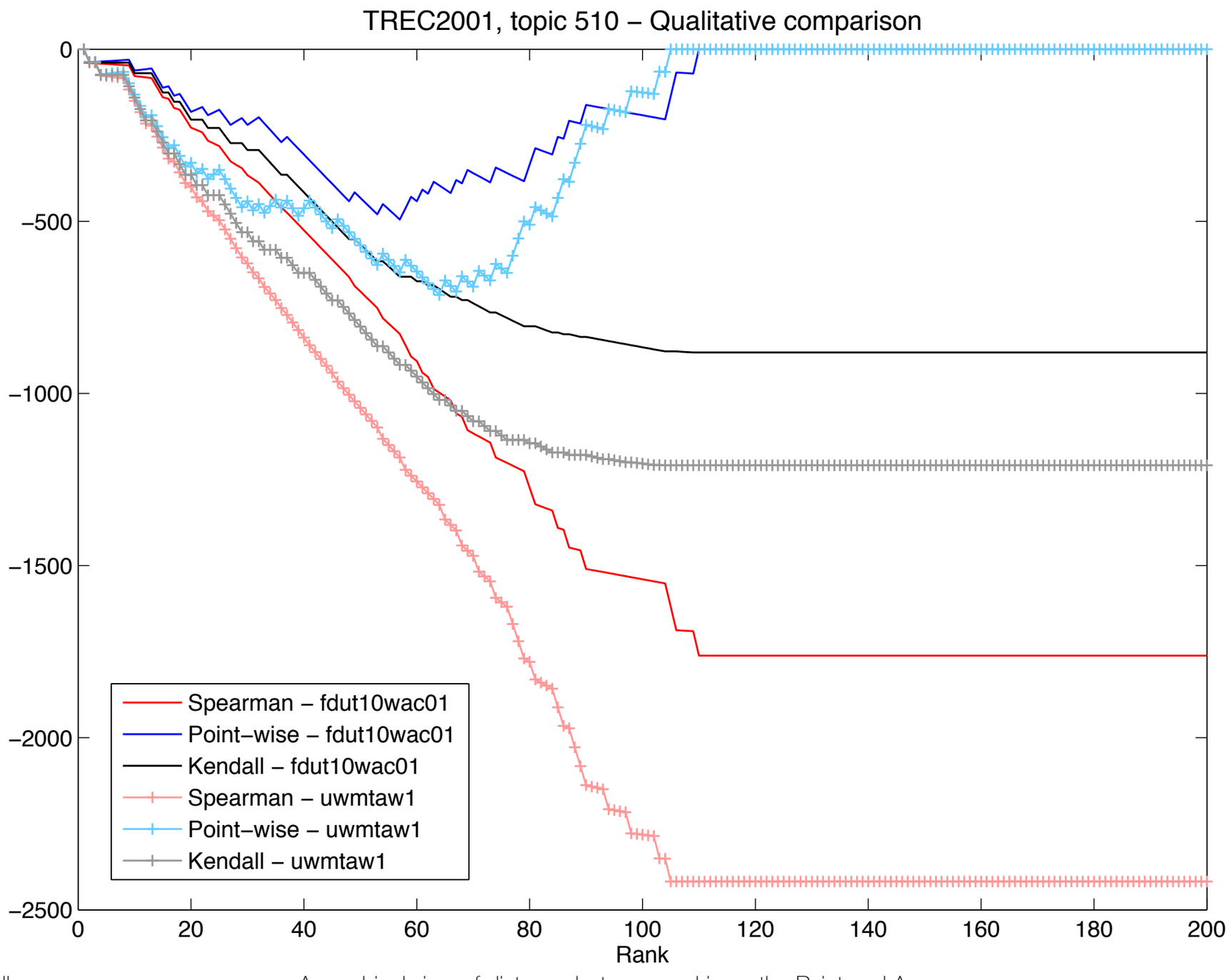


A graphical view of distance between rankings: the Point and Area measure

A-corr on TREC test collections



Point-wise as an effectiveness measure (qualitative)



Conclusions

- A correlation and an effectiveness measure for qualitative and quantitative evaluation

- We plan to:
  - compare A-corr with the Twist measure (Cumulative Relative Position)*
  - analyze its stability, sensitivity and correlation with other measures
  - define a weighted measure to model user behavior

* N. Ferro, G. Silvello, H. Keskustalo, A. Pirkola and K. Järvelin (2015), The Twist Measure for IR Evaluation: Taking User's Effort into Account *Journal of the Association for Information Science and Technology (JASIST)* in print (http://onlinelibrary.wiley.com/doi/10.1002/asi.23416).