# An Iterative Distance-Based Model for Unsupervised Weighted Rank Aggregation

Leonidas Akritidis[1], Athanasios Fevgas[1], Panayiotis Bozanis[2,*]

[1]Department of Electrical and Computer Engineering, University of Thessaly
[2]Department of Science and Technology, International Hellenic University
[*]Research conducted while being in [1]

The 2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2019)
October 14-17, 2019, Thessaloniki, Greece

# Rank Aggregation (RA)

- A set of $n$ voters $V$.
- Given a query or a subject, each voter $v$ submits a single ranked list $R^v$ of answers/suggestions $r_i^v$.
- A rank aggregation method $\mathcal{T}$ fuses all $R^v$ lists into one aggregated list $L$ with improved ranking of its elements.
- Applications: Voting systems, bioinformatics, Web metasearch, collaborative filtering, etc.

# Weighted vs. non weighted RA

- Each voter $v$ is assigned a weight $w_v$.
- Non weighted rank aggregation: $w_v = 1, \forall v \in V$
  - All voters are treated equally.
  - $L = \mathcal{T}\left(R^{v_1}, R^{v_2}, \ldots, R^{v_n}\right)$
- Weighted rank aggregation: Each voter is assigned a weight which reflects his/her importance and/or expertise on the subject.
  - $L = \mathcal{T}\left(w_{v_1}, w_{v_2}, \ldots, w_{v_n}, R^{v_1}, R^{v_2}, \ldots, R^{v_n}\right)$

# Motivation

| Player | Borda Score |
|---|---|
| Messi | 3+2+3+2=10 |
| Ronaldo | 2+3+0+0=5 |
| Neymar | 1+0+0+3=4 |
| Sallah | 0+1+2+0=3 |
| Mane | 0+0+0+2=2 |
| VanDijk | 0+0+1+0=1 |

- Who is the best football player?

| Voter 1 $v_1$ | Voter 2 $v_2$ | Voter 3 $v_3$ | Voter 4 $v_4$ |
|---|---|---|---|
| Messi | Ronaldo | Messi | Neymar |
| Ronaldo | Messi | Sallah | Mane |
| Neymar | Sallah | Van Dijk | Messi |

- What does this tell us about $v_1$?
  - All of his/her suggestions made it to the top-3 of $L$.
  - S/he is an expert on the subject.
- What about $v_3$?
  - Only one of his/her suggestions in the top-3 of $L$.
  - S/he is not so familiar with the subject.

# Idea

- Compare (find the distance of) each input list $R^v$ with the aggregate list $L$.
- The voter who submits a ranked list that has a *small* distance from $L$ *is considered as* an expert.
  - His/her choices have strong support by other voters.
  - S/he should be assigned a high weight.
- In contrast: The voter who submits a ranked list that has a *great* distance from $L$ is *not* an expert.
  - His/her weight should be smaller compared to the previous voter.

# Ranked List Distance Metrics

- Spearman Footrule distance: $d_F(R^v, L) = \sum_{j=1}^{k} |j - l_j|$
- Kendall's tau: $\tau = d_K(R^v, L) = \frac{2}{n(n-1)} \sum_{i<j} \text{sgn}(i-j)\text{sgn}(l_i - l_j)$
- Scaled Footrule Distance: $d_{F'}(R^v, L) = \sum_{j=1}^{k} \left| \frac{j}{k} - \frac{l_j}{|L|} \right|$
- Proposed: Locality-Sensitive Scaled Footrule distance takes into account not only the (scaled) difference in rankings, but also their location in the aggregate list. It "punishes" the discrepancies which occur in high positions of $L$.

$$d_{F'}(R^v, L) = \sum_{j=1}^{k} \left| \frac{j}{k} - \frac{l_j}{|L|} \right| \log \frac{|L|}{l_j}$$

# Distance-Based Model for URA

- Assign all voters an initial equal weight $w_{v,0}$.
- Apply the RA method $\mathcal{T}$ of choice in the traditional, non weighted fashion.
- Obtain an initial aggregate list $L_0$.
- For each voter $v$, compute the distance $d(R^v, L_0)$ between his/her list $R^v$ and $L_0$.
- Compute the new weight $w_{v,1}$ of $v$ according to:

$$w_{v,1} = w_{v,0} + f\left(d(R^v, L_0)\right)$$

- $f$ is a kernel function that depends on $d(R^v, L_0)$.
- Reapply $\mathcal{T}$ on its weighted form.

# Iterative Distance-Based Model

- The previous equation can be applied many times in an iterative fashion until the weights converge.

$$w_{v,i} = w_{v,i-1} + f\left(d(R^v, L_{i-1})\right), \quad i \in \mathbb{N}^*$$

- The kernel function f must:
  - be asymptotically upper bounded, and
  - Satisfy $f\left(d\left(R^v, L\right)\right) \le f\left(d\left(R^{v'}, L\right)\right)$ if $d\left(R^v, L\right) \ge d\left(R^{v'}, L\right)$
- Exp Kernel function: $f\left(d\left(R^v, L\right)\right) = \exp\left(i \cdot d\left(R^v, L\right)\right)$
- $i$: the $i^{th}$ iteration
- Finally: $w_{v,i} = w_{v,i-1} + \exp\left(i \cdot d(R^v, L_{i-1})\right), \quad i \in \mathbb{N}^*$

Initialize the weights of the voters

Apply non-weighted RA $\mathcal{T}$

Iteration Counter

Flag: Checks convergence of all weights

Iterate until convergence

Compute the distance between the input list $R^v$ of a voter $v$ and the aggregate list $L$. Update the weight $w_v$ of $v$ and check convergence.

Apply weighted RA $\mathcal{T}$

**Algorithm 1:** Distance-based iterative model for unsupervised weighted rank aggregation

initialize an empty aggregated list $L$;<br>
**for** *each voter* $v \in V$ **do**<br>
    $w'_v \leftarrow 1/|V|$;<br>
**end**<br>
$L \leftarrow \mathcal{T}(R^{v_1}, R^{v_2}, \ldots, R^{v_n})$ (Eq. 1);<br>
$i \leftarrow 0$;<br>
*allconverged* $\leftarrow$ false;<br>
**while** *not allconverged* **do**<br>
    $i \leftarrow i + 1$;<br>
    *allconverged* $\leftarrow$ true;<br>
    **for** *each voter* $v \in V$ **do**<br>
        Compute $d(R^v, L)$ (Eq. 3, 4, 5, or 6);<br>
        Compute normalized $d(R^v, L)$;<br>
        Set $w_v \leftarrow w'_v + \exp\left(i \cdot d(R^v, L)\right)$ (Eq. 10);<br>
        **if** $w_v - w'_v > prec$ **then**<br>
            *allconverged* $\leftarrow$ false;<br>
        **end**<br>
    **end**<br>
    $L \leftarrow \mathcal{T}(w_{v_1}, \ldots, w_{v_n}, R^{v_1}, \ldots, R^{v_n})$ (Eq. 2);<br>
    **for** *each voter* $v \in V$ **do**<br>
        $w'_v \leftarrow w_v$;<br>
    **end**<br>
**end**

# Experiments

- We utilized 6 datasets from the Web Tracks of TREC 2009-2014.
- In these tracks, the participant groups are given a set of 50 predefined queries.
- They submit their result lists which are evaluated by using relevance judgments from human judges.
- Ideal for rank aggregation.
  - Fair, unbiased, reputable.

**Table 1: Datasets Characteristics**

| Dataset | Queries | Runs | Avg. Input Size |
|---|---|---|---|
| TREC 2009 | 50 | 71 | 988.7 |
| TREC 2010 | 50 | 56 | 9180.7 |
| TREC 2011 | 50 | 62 | 8325.1 |
| TREC 2012 | 50 | 48 | 6719.5 |
| TREC 2013 | 50 | 61 | 7174.4 |
| TREC 2014 | 50 | 30 | 6514.9 |

# Results (Mean Average Precision)

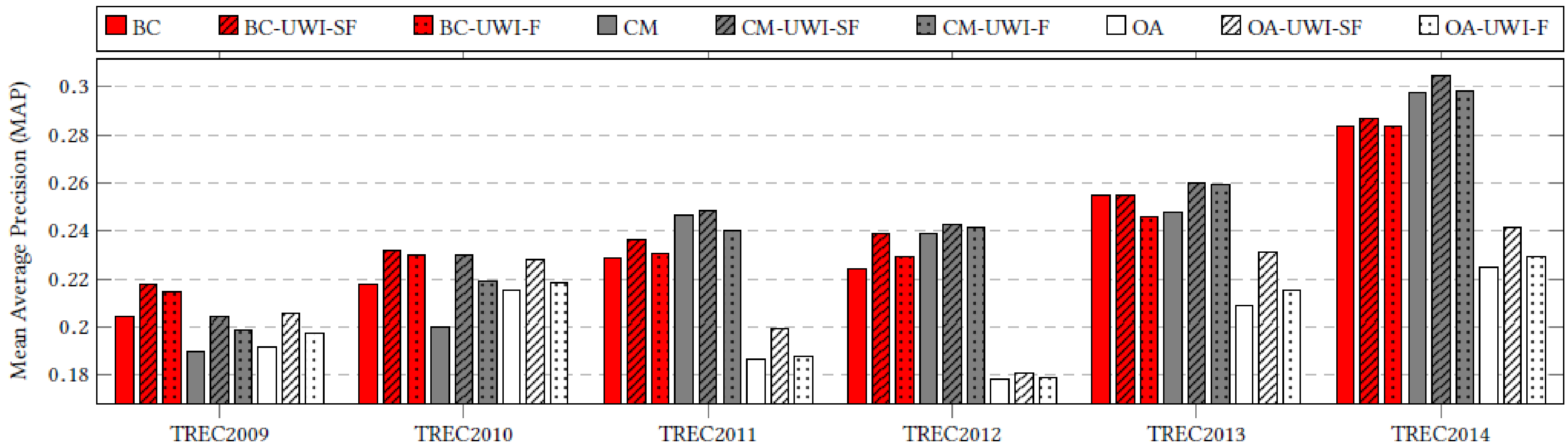


**Figure 1: Performance comparison of various rank aggregation methods on the Adhoc datasets of TREC 2009–2014.**

- **BC:** Borda Count, **CM:** Condorcet Method, **OA:** Outranking Approach (Vanderpooten et.al, ACM SIGIR 2011).
- **UWI:** Unsupervised, Weighted, Iterative.
  - **SF:** Locality-Sensitive Scaled Footrule Distance.
  - **F:** Standard Scaled  Footrule Distance.
- Our model achieves a consistent improvement of about 4-10%.

# Results (other evaluation metrics)

**Table 1: Performance comparison of various rank aggregation methods with different evaluation metrics on the Adhoc datasets of the Web Tracks of TREC 2009–2014.**

| Dataset | Method | MAP | P@5 | P@10 | P@20 | P@100 | nDCG@10 | nDCG@20 |
|---|---|---|---|---|---|---|---|---|
| TREC 2009 | Borda Count | 0.2044 | 0.5040 | 0.4980 | 0.4620 | 0.2964 | 0.3972 | 0.3761 |
| | Borda Count-UWI-SF | **0.2162** | 0.5440 | 0.5020 | **0.4750** | **0.3056** | **0.3996** | **0.3773** |
| | Condorcet method | 0.1900 | 0.4240 | 0.4380 | 0.4410 | 0.2874 | 0.3436 | 0.3450 |
| | Condorcet method-UWI-SF | 0.1987 | 0.4680 | 0.4760 | 0.4450 | 0.2976 | 0.3752 | 0.3543 |
| | Outranking approach | 0.1916 | 0.5440 | 0.5060 | 0.4450 | 0.2888 | 0.3691 | 0.3423 |
| | Outranking approach-UWI-SF | 0.2057 | **0.5640** | **0.5160** | 0.4510 | 0.2918 | 0.3725 | 0.3543 |
| TREC 2010 | Borda Count | 0.2181 | 0.4960 | **0.4580** | 0.4080 | 0.2766 | 0.2596 | 0.2572 |
| | Borda Count-UWI-SF | **0.2316** | **0.5000** | **0.4580** | 0.4130 | 0.2902 | **0.2599** | **0.2578** |
| | Condorcet method | 0.2001 | 0.4320 | 0.3980 | 0.3900 | 0.2814 | 0.2116 | 0.2290 |
| | Condorcet method-UWI-SF | 0.2193 | 0.4160 | 0.4060 | 0.3920 | 0.2932 | 0.2206 | 0.2383 |
| | Outranking approach | 0.2154 | 0.4760 | 0.4560 | 0.4180 | 0.2818 | 0.2514 | 0.2546 |
| | Outranking approach-UWI-SF | 0.2280 | 0.4930 | 0.4560 | **0.4240** | **0.2972** | 0.2564 | 0.2560 |
| TREC 2011 | Borda Count | 0.2287 | 0.4080 | 0.3600 | 0.3360 | 0.2326 | 0.2417 | 0.2392 |
| | Borda Count-UWI-SF | 0.2363 | 0.4400 | 0.3620 | 0.3300 | 0.2354 | 0.2373 | 0.2387 |
| | Condorcet method | 0.2463 | **0.4520** | **0.3760** | **0.3610** | 0.2490 | 0.2713 | 0.2697 |
| | Condorcet method-UWI-SF | **0.2484** | **0.4520** | **0.3760** | 0.3520 | **0.2524** | **0.2751** | **0.2730** |
| | Outranking approach | 0.1863 | 0.4120 | 0.3460 | 0.3010 | 0.1932 | 0.2380 | 0.2176 |
| | Outranking approach-UWI-SF | 0.1994 | 0.4240 | 0.3460 | 0.3010 | 0.2074 | 0.2457 | 0.2260 |
| TREC 2012 | Borda Count | 0.2241 | 0.4000 | 0.3800 | 0.3420 | 0.2366 | 0.1605 | 0.1632 |
| | Borda Count-UWI-SF | 0.2392 | 0.4360 | 0.3860 | 0.3490 | 0.2398 | 0.1688 | 0.1683 |
| | Condorcet procedure | 0.2391 | **0.4720** | 0.4520 | 0.3960 | 0.2446 | 0.2066 | 0.2046 |
| | Condorcet method-UWI-SF | **0.2430** | 0.4680 | **0.4560** | **0.4030** | **0.2458** | **0.2112** | **0.2063** |
| | Outranking approach | 0.1784 | 0.3120 | 0.2960 | 0.2880 | 0.2034 | 0.1329 | 0.1484 |
| | Outranking approach-UWI-SF | 0.1811 | 0.3240 | 0.3120 | 0.2980 | 0.2056 | 0.1345 | 0.1501 |
| TREC 2013 | Borda Count | 0.2551 | 0.4880 | 0.4620 | 0.4190 | 0.2724 | 0.2509 | 0.2677 |
| | Borda Count-UWI-SF | 0.2551 | 0.5000 | **0.4680** | **0.4260** | 0.2730 | **0.2559** | **0.2733** |
| | Condorcet method | 0.2480 | 0.4600 | 0.4220 | 0.4090 | 0.2684 | 0.2385 | 0.2637 |
| | Condorcet method-UWI-SF | **0.2599** | 0.4720 | 0.4560 | 0.4200 | **0.2816** | 0.2506 | 0.2691 |
| | Outranking approach | 0.2091 | 0.4960 | 0.4420 | 0.3870 | 0.2438 | 0.2310 | 0.2425 |
| | Outranking approach-UWI-SF | 0.2311 | **0.5120** | 0.4560 | 0.3900 | 0.2604 | 0.2409 | 0.2517 |
| TREC 2014 | Borda Count | 0.2838 | 0.5960 | 0.5820 | 0.5460 | 0.3382 | 0.2815 | 0.3020 |
| | Borda Count-UWI-SF | 0.2866 | 0.5940 | 0.5800 | 0.5520 | 0.3456 | 0.2912 | 0.3159 |
| | Condorcet method | 0.2976 | **0.6000** | **0.5800** | **0.5720** | 0.3712 | 0.2880 | 0.3177 |
| | Condorcet method-UWI-SF | **0.3048** | **0.6000** | **0.5880** | **0.5720** | **0.3728** | **0.2928** | **0.3207** |
| | Outranking approach | 0.2246 | 0.5160 | 0.5080 | 0.4530 | 0.2978 | 0.2385 | 0.2487 |
| | Outranking approach-UWI-SF | 0.2416 | 0.5160 | 0.5080 | 0.4550 | 0.3112 | 0.2512 | 0.2619 |

# Preliminary Convergence Study

- The proposed distance metric needs on average 1-2 more iterations compared to the original scaled footrule distance.
- The weights converge more slowly for Borda Count. Borda Count is much faster than Condorcet Method and Outranking Approach.

**Table 2: Average number of iterations per query before weights convergence for various UWI extensions**

| Dataset | BC-UWI | | CM-UWI | | OA-UWI | |
|---|---|---|---|---|---|---|
| | -SF | -F | -SF | -F | -SF | -F |
| TREC 2009 | 14.10 | 12.48 | 7.26 | 5.62 | 7.32 | 6.14 |
| TREC 2010 | 14.70 | 12.96 | 6.98 | 5.56 | 6.90 | 5.72 |
| TREC 2011 | 12.22 | 10.24 | 7.26 | 5.72 | 6.50 | 5.62 |
| TREC 2012 | 9.50 | 8.78 | 6.48 | 5.52 | 5.80 | 5.22 |
| TREC 2013 | 15.12 | 11.90 | 8.42 | 6.32 | 7.26 | 5.82 |
| TREC 2014 | 11.70 | 10.38 | 7.28 | 5.50 | 6.26 | 5.30 |

# Conclusions

- We presented a new unsupervised model for weighted rank aggregation.
- The model considers that the voters whom lists are close to the aggregate list as experts.
- It iteratively computes the distance of each input list with the aggregate list and modifies the voter weights accordingly.
- We introduced a metric for computing the distance between two ranked lists.
  - Sensitive to the “locality” of disagreements.
- The model was evaluated by using datasets from TREC and was found to improve rankings by 4-10%.

# Thank you

Any questions?