

## Erratum to: Cluster-based polyrepresentation as science modelling approach for information retrieval

Muhammad Kamran Abbasi<sup>1</sup> · Ingo Frommholz<sup>1</sup>

Published online: 3 April 2015  
© Akadémiai Kiadó, Budapest, Hungary 2015

### Erratum to: Scientometrics DOI 10.1007/s11192-014-1478-1

The published version of the paper reports results based on simulated user methodology. There are some omissions in the algorithm and the statement of the approach. Section “Simulated User Methodology and Cluster Ranking” and the Algorithm 1 should read as:

It should be noted that by the way we simulate the user, a new ranking of documents is created based on the sequence of clusters examined by the user and the within-cluster ranking of the documents. This way we can compare the interactive ranking approach against a baseline ranking, which may not be based on any clustering or even polyrepresentation, in a controlled environment utilising standard IR evaluation measures. All the documents the simulated user visits form a ranking according to the procedure given in Algorithm 1 for fixed  $l$ . For each query, the  $l$  documents from each cluster are combined together and are re-ranked based on their actual retrieval weights—which are computed by applying combSum on the weights produced by Equation 2 and 3 for  $REP_{in}$  and  $REP_d$  respectively; to create the final ranking. Note that, for a fixed  $l$  the order of clusters (cluster ranking) has no effect (only the parameter  $l$ , which is fixed and the same for each cluster, influences the ranking), while for *Varireps* and *Variseq*, respectively, the cluster order matters because it determines the number of documents drawn from each cluster. In this

---

The online version of the original article can be found under doi:[10.1007/s11192-014-1478-1](https://doi.org/10.1007/s11192-014-1478-1).

---

✉ Muhammad Kamran Abbasi  
[abbasikamran@gmail.com](mailto:abbasikamran@gmail.com)

Ingo Frommholz  
[ingo.frommholz@beds.ac.uk](mailto:ingo.frommholz@beds.ac.uk)

<sup>1</sup> Institute for Research in Applicable Computing, University of Bedfordshire, Park Square, Luton LU1 3JU, UK

case  $l$  is variable and a different number of  $l$  documents are taken from each cluster and added to the ranking.

Note: These corrections do not affect the experimental results given in the paper.

**Require:** Clustering  $\mathcal{C}$ ,  $l$   
 $r \leftarrow ()$  {The ranking, initially an empty list}  
 $\mathcal{L}_{\mathcal{C}} \leftarrow$  ranked list of clusters in  $\mathcal{C}$  (using eF or SD)  
**for all** cluster  $C \in \mathcal{L}_{\mathcal{C}}$  **do**  
     $l_C \leftarrow$  ranked list of documents in  $C$  {process  $C$  in descending weight order}  
    **for**  $i = 1$  to  $l$  **do**  
         $r \leftarrow r + l_C[i]$  {append document at rank  $i$  to  $r$ }  
    **end for**  
**end for**  
 $r$  re-rank documents in  $r$  {in descending weight order}  
**return**  $r$

**Algorithm 1:** Cluster-based re-ranking for simulated user (fixed  $l$ )