

## Precisiones de la combinación de ránkings

Buscador	Pesos	Clueweb-1K		Clueweb-10K	
		P@5	P@10	P@5	P@10
Tf-Idf	1	0.16	0.12	0.0	0.05
Literal	1	0.28	0.26	0.08	0.08
Proximal	1	0.36	0.44	0.2	0.16
Tf-Idf + PageRank	0.1 + 0.9	0.0	0.0	0.0	0.0
	0.2 + 0.8	0.0	0.0	0.0	0.0
	0.3 + 0.7	0.0	0.0	0.0	0.0
	0.4 + 0.6	0.0	0.0	0.0	0.0
	0.5 + 0.5	0.0	0.04	0.0	0.0
	0.6 + 0.4	0.0	0.04	0.0	0.0
	0.7 + 0.3	0.0	0.04	0.0	0.0
	0.8 + 0.2	0.08	0.04	0.0	0.0
	0.9 + 0.1	0.08	0.04	0.0	0.0
Literal + PageRank	0.1 + 0.9	0.0	0.0	0.0	0.0
	0.2 + 0.8	0.0	0.0	0.0	0.0
	0.3 + 0.7	0.0	0.0	0.0	0.0
	0.4 + 0.6	0.0	0.0	0.0	0.0
	0.5 + 0.5	0.0	0.04	0.0	0.0
	0.6 + 0.4	0.0	0.04	0.0	0.0
	0.7 + 0.3	0.0	0.04	0.0	0.0
	0.8 + 0.2	0.08	0.04	0.0	0.0
	0.9 + 0.1	0.08	0.04	0.0	0.0
Tf-Idf + Proximal + PageRank	0.05 + 0.9 + 0.05	0.2	0.12	0.06	0.03

La agregación de ránkings tal como está pensada no mejora, si no que empeora drásticamente los resultados de los buscadores que no están agregados.

Esto se debe primariamente a como se aplica PageRank a los demás buscadores. El agregador implementado confecciona un ránking total haciendo uso de los ránkings pasados y aplicando unos pesos determinados a cada uno de estos. Este procedimiento es perfectamente válido cuando lo que se está combinando son resultados TOP5 y TOP10 de sendos buscadores, pero no es congruente con la definición de PageRank, ya que en la mayoría de los casos, el TOP10 de documentos de PageRank no serán relevantes para todas las consultas que se puedan realizar sobre la colección.

Para solucionar este problema, se debería diferenciar claramente en la implementación cuándo se está intentando agregar PageRank, en cuyo caso, además de realizar la suma pesada de cada uno de los documentos de los ránkings, debería aplicarse al final la puntuación PR de cada documento del TOP resultante.