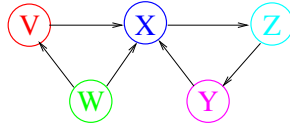
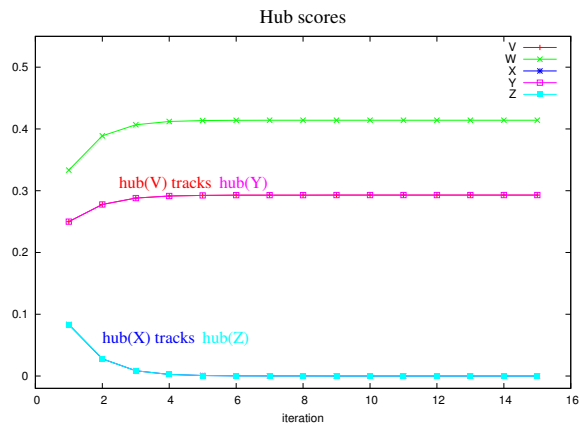
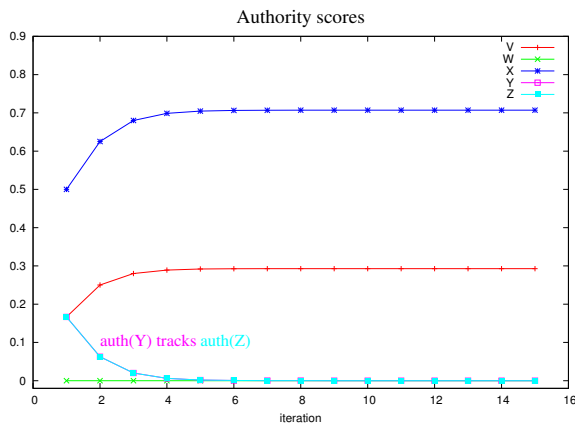


Topics: more on hubs and authorities; comparison to PageRank.

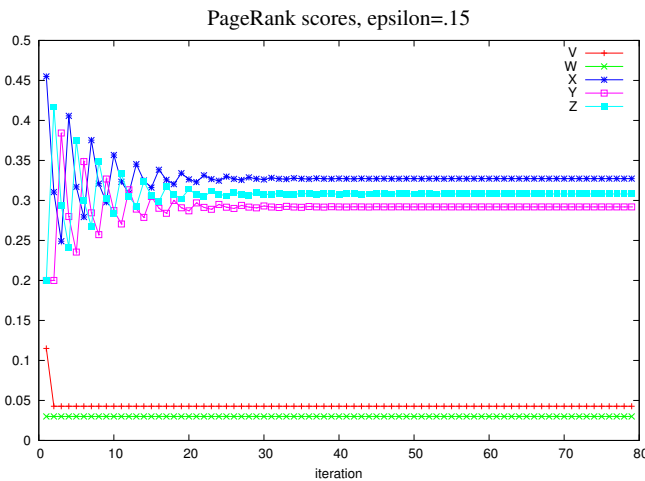
I. Example document set



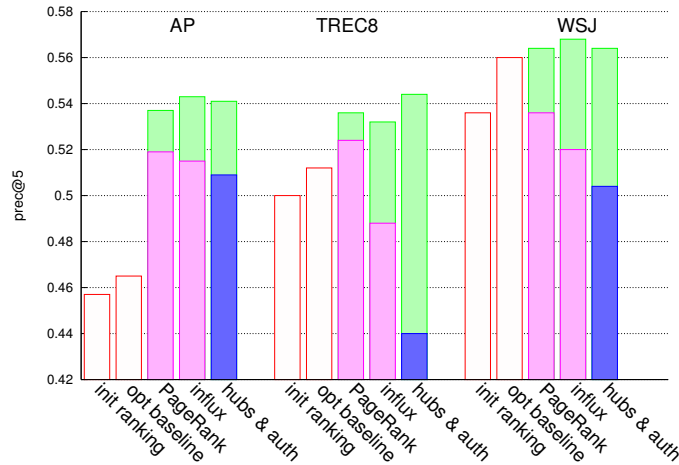
II. Hubs and authorities on our example document set Here, we're assuming that the document collection above comprises a root set in response to some query. Since we are only interested in relative scores, the y-axes in the plots are not aligned.



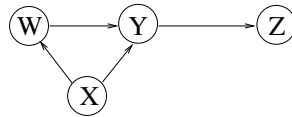
III. PageRank on our example document set Here, we assume that the entire corpus consists of the above five documents.



IV. Link analysis on non-hyperlinked corpora Figure from O. Kurland and L. Lee, “Respect my authority! HITS without hyperlinks, utilizing cluster-based language models”, SIGIR 2006. Links between documents are induced and weighted by their content-analysis-based similarities. Influx corresponds to in-degree in this link-weighted scenario. Green shows the performance improvement when document clusters are used as entities in the network (as opposed to just documents).



V. Self-check: Example hubs and authorities calculations Note that the document network here is *not* the one shown on the previous page. Convergence has not been reached at the end of the table (you can check).



		W		X		Y		Z	
		auth	(hub)	auth	(hub)	auth	(hub)	auth	(hub)
a.	Init	—	(1/4)	—	(1/4)	—	(1/4)	—	(1/4)
b.	TempAs	1/4	”	0	”	2/4	”	1/4	”
c. <i>Norm by 1/4 + 2/4 + 1/4 = 1, so do nothing</i>									
d.	TrueAs	1/4	”	0	”	1/2	”	1/4	”
e.	TempHs	”	(1/2)	”	(3/4)	”	(1/4)	”	(0)
f. <i>Norm by 1/2 + 3/4 + 1/4 = 3/2, or multiply by 2/3</i>									
g.	TrueHs	”	(1/3)	”	(1/2)	”	(1/6)	”	(0)
h.	TempAs	1/2	”	0	”	5/6	”	1/6	”
i. <i>Norm by 1/2 + 5/6 + 1/6 = 9/6, or multiply by 2/3=6/9</i>									
j.	TrueAs	1/3	”	0	”	5/9	”	1/9	”