Task-3

Top-5 page rank results for G1

Sustainable_energy	0.0010440914311
Wind_power	0.001026080118
Plug-in_hybrid	0.001026080118
Solar_energy	0.001026080118
Hydroelectricity	0.001026080118

in-links to the top 5 pages listed above (from G1 directed graph)

Sustainable_energy Cogeneration Efficient_energy_use Green_building Greenpower Energy_conservation
Wind_power Cogeneration Efficient_energy_use Green_building Sustainable_energy Energy_conservation
Plug-in_hybrid Cogeneration Efficient_energy_use Green_building Sustainable_energy Energy_conservation
Solar_energy Cogeneration Efficient_energy_use Green_building Sustainable_energy Energy_conservation
Hydroelectricity Cogeneration Efficient_energy_use Green_building Sustainable_energy Energy_conservation

we can see that Sustainable_energy has 5 in-links of which it has Greenpower in-link which has 44 out-links count(from G1 graph). As per the page rank calculation in the algorithm,

newPR[p1] += d*float(PR[q])/L

where p1 is page from set of all pages P,q is an in-link to the page p1 and L is the the out-link count of q.

Since GreenPower has lesser out-link count,the page rank value of Sustainable_energy increases to the highest thereby.

Similarly the Wind_power,Plugin_in_hybrid,Solar_energy and Hydroelectricity have second highest page ranks (same) since all of them 5 in-links. Sustainable_energy is one of their in-inks with highest page rank,this increases the page rank value to be the second highest for Wind_power,Plugin_in_hybrid,Solar_energy and Hydroelectricity. The number of in-inks also has a contribution to high page rank values. Higher the in-inks count(indicating higher popularity) higher will be the page rank. Also we can notice that the only difference in the in-link set of Sustainable_energy and others is the listing above is that Sustainable_energy has Greenpower whereas the others have Sustainable_energy in their in-inks set. Other unlinks are the same. Also the outline count for

Sustainable_energy 459. In conclusion, Both in-link count and the out-link count factors together determine the page rank values.

Top-5 page rank results for G2

WT21-B37-76 0.00267941005759
WT21-B37-75 0.00152591699385
WT25-B39-116 0.00146949352594
WT23-B21-53 0.00137232204363
WT24-B40-171 0.0012450717567

in-link & out-link count of the top 5 pages sorted by page rank

	in-link count	out-link count
WT21-B37-76	: 2568	:5
WT21-B37-75	:1704	:1
WT25-B39-116	:169	:1
WT23-B21-53	:198	:1
WT24-B40-171	:270	:209

Topics
WT21-B37-76 - The Economist(index.html)
WT21-B37-75- The Economist(Copyright.html)
WT25-B39-116-(Contents page)
WT23-B21-53-(Web Development Team)
WT24-B40-171-(www.chch.com:80/news/eveningnews/archive/)

Top 5 in-link count of the page set

WT21-B37-76:2568
WT18-B29-37:2269
WT01-B18-225:2260
WT23-B27-29:1940
WT21-B37-75:1704

WT21-B37-76 has the highest page rank since its in-inks count is the highest and out-link is 5 indicating that its most popular in the given set. Similarly WT21-B37-75 is the second highest in-link count. On the other hand WT25-B39-116(Content page) has lower in-link count and yet in the top 5. The in-link count is not the only deciding factor for

high page ranks. It could be so that some of its in-links could point to sink nodes resulting in a loop. Same is the case the 4th and the 5th rank pages.
As we can see above, the top 5 in-link count of the page are not the same as the top 5 in page rank result.