Authorative Sources in a Hyperlinked Environment

1. What is the problem that the paper wants to solve? Why is it difficult (related works)?
   - Searching the web is difficult: enormous complexity, millions of websites, have to find the most relevant to query
   - Different problems for different queries: for specific queries there are very few pages that contain the relevant information which the user is looking for (scarcity problem), for broad queries there are too many pages that are in some way related to the search query (abundance problem)
   - Even authorative sources (i.e. harvard.edu for information about Harvard University) may not contain the search term on their web page and thus may not appear in search results
   - The author wants to solve the problem of (conferred) authority which may improve search result by ranking websites higher if they have high authority on the search query
2. What is the solution? What is the main idea?
   - Main idea is to consider the world wide web as a graph of websites and links and look at edges in the graph, edges (links) confer authority to other pages
   - Author constructs a subgraph of the WWW for each query which is then used to find the most authorative sites
3. What is the result?
   - Author ranks sites based on their measure of authority
   - Find relevant results globally (rather than locally, on a single site)
   - Discovered an equilibrium between *hubs* and authorities
4. What is the main novelty that enabled the solution?
   - Author builds on the very simple idea that websites with many links from hubs on a topic may be considered an authority on that topic
5. What are the good aspects of the paper? Did you learn something from the paper?
   - Paper leverages a simple (but good) idea to achieve good results
6. What is the impact of the paper?
   - Notion of conferred authority by the paper is very similar to idea behind Google's page-rank algorithm which has worked very well in practice
7. Are there weaknesses/missing parts in the paper? How can you improve it?
   - Paper focuses only on the notion of conferred (link-based) authority to rank search results but excludes other relevant characteristics
   - Subgraph construction relies entirely on other text-based search engines, results may be improved by computing the subgraph differently
8. How can you extend the paper?
   - Find a way to find relevant search results that works well with the author's notion of authority (difficult)
   - Employ graph structures beyond hubs and authorities to improve ranking algorithm
9. How can you apply the technique to other data/problems?
   - Authoritative search result ranking may be used for other data that is similar to the web in structure
   - Author notes scientific citations follow a different structure because they lack an equivalent to hubs in the web