Web Search and Link Analysis

Saptarshi Pyne

Assistant Professor
Department of Computer Science and Engineering
Indian Institute of Technology Jodhpur, Rajasthan, India 342030

CSL7390 Social Network Analysis Lecture 36
24th April 2024

Challenges in web search

- Synonymy: Multiple ways to say the same thing
- Polysemy: Multiple meanings of the same term
- From 'needle-in-a-haystack' to abundance of relevant contents
- Continuous addition of new types of contents
- Demand for the 'news search' features
- Need for ranked search results for human users

The field of 'information retrieval' deals with automated search of web and offline documents.

Link analysis can help in ranking search results

- Analysis of in-links
  - However, an in-link may mean different things:
    - Endorsement (expected meaning)
    - Criticism
    - Paid
    - Random etc.
  - If a page has a large number of in-links, we assume that most of the in-links are endorsements

Link analysis can help in ranking search results (contd.)

- We acquire a large sample of webpages that are relevant to the search query

- Each page in this sample receives votes from the other pages of the sample through the in-links of the former

- Finally, the sampled pages are ranked by the votes they have received

References

- **Kleinberg**: David Easley and Jon Kleinberg (2010), ***Networks, Crowds, and Markets: Reasoning About a Highly Connected World***, Cambridge University Press. Pre-publication draft. Book website.
  - Chapter 14: 'Link Analysis and Web Search' Sections 14.1 – 14.5

Thank you