CS 685 The Structure of Information Networks February 25, 2005

Reaction Paper 2

Due Date: March 11, 2005

The second reaction paper will cover the topics of community finding on the web, sociological analysis of communities in networks, and methods for aggregating multiple rankings of web pages for metasearch. The topics are defined by the list of papers below.

If you covered one of these topics (community finding, for example) in the first paper, you are welcome to write your second paper on topics of the first paper, or to write another paper on community finding.

Recall that a reaction paper will cover two related papers associated with this segment of lectures; at least one of these must not be on the given list. The reaction paper should be about three pages of discussion related to the papers. The discussion should go beyond a mere summary of the two papers. For instance, you might discuss some weaknesses in the papers, and suggest alternative approaches; formulate different hypotheses than the papers and suggest experiments that could be used to validate the hypotheses; discuss related work from a different field that the authors were not aware of, or a potential application of the work to a different field.

The papers listed below are the same as those listed on the web site under items 3 & 4; links to these papers can be found on the web site.

- **Finding communities.** How can we automatically find "communities" of common interests in information networks?
  - Spectral methods
    - S. Deerwester, S.T. Dumais, T.K. Landauer, G.W. Furnas and R.A. Harshman. Indexing by latent semantic analysis. Journal of the Society for Information Science, 41(6), 391-407 (1990).
    - C. Papadimitriou, P. Raghavan, H. Tamaki, and S. Vempala. Latent Semantic Indexing: A Probabilistic Analysis. 17th ACM Symposium on the Principles of Database Systems, 1998.
    - D. Gibson, J. Kleinberg, P. Raghavan. Inferring Web communities from link topology. Proc. 9th ACM Conference on Hypertext and Hypermedia, 1998.
    - P. Drineas, R. Kannan, A. Frieze, S. Vempala and V. Vinay. Clustering large graphs via the Singular Value Decomposition. Machine Learning 56:9-33, 2004.
    - Y. Azar, A. Fiat, A. Karlin, F. McSherry and J. Saia. Spectral Analysis of Data. 33rd ACM Symposium on Theory of Computing, 2001.
    - D. Achlioptas, A. Fiat, A. Karlin, and F. McSherry, Web Search via Hub Synthesis. 42nd IEEE Symposium on Foundations of Computer Science, 2001, p.611-618.
    - A. Y. Ng, A. X. Zheng, and M. I. Jordan. Link analysis, eigenvectors, and stability. International Joint Conference on Artificial Intelligence (IJCAI), 2001.

    * A. Y. Ng, A. X. Zheng, and M. I. Jordan. Stable algorithms for link analysis. 24th International Conference on Research and Development in Information Retrieval (SIGIR 2001).

  - Combinatorial methods
    * R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. Trawling the web for emerging cyber-communities. 8th International World Wide Web Conference, May 1999.
    * R. Agrawal, R. Srikant. Fast algorithms for mining association rules. 20th Int'l Conference on Very Large Databases (VLDB), 1994.
    * G.W. Flake, S. Lawrence, and C.L. Giles. Efficient identification of web communities. Proc. of the 6th International Conference on Knowledge Discovery and Data Mining (KDD), 2000.
    * G. Flake, S. Lawrence, C. L. Giles, and F. Coetzee. Self-organization and identification of web communities. IEEE Computer, 35:3, March 2002.
    * G. Flake, K. Tsioutsiouliklis, and R.E. Tarjan. Graph Clustering Techniques based on Minimum Cut Trees. Technical Report 2002-06, NEC, Princeton, NJ, 2002.
    * M. Girvan and M. E. J. Newman. Community structure in social and biological networks. Proc. Natl. Acad. Sci. USA 99, 8271-8276 (2002).
    * M. Granovetter. The strength of weak ties. American Journal of Sociology, 78(6):1360-1380, 1973.

- **Rank aggregation and meta-search.** Given multiple rankings of web pages, can we combine these in a principled way into a single ranking?
  - K. Arrow. Social choice and individual values. Wiley, 1951.
  - H.P. Young. Condorcet's theory of voting. American Political Science Review 82(1988), pp. 1231-1244.
  - Cynthia Dwork, Ravi Kumar, Moni Naor, D. Sivakumar. Rank aggregation methods for the Web. 10th International World Wide Web Conference, May 2001.
  - Erik Selberg, Oren Etzioni. The MetaCrawler architecture for resource aggregation on the web. IEEE Expert, 1997.
  - Weiyi Meng, Clement Yu, King-Lup Liu. Building efficient and effective metasearch engines. ACM Computing Surveys 34(2002).