Complex Systems 535: Network Theory

Instructor: Mark Newman

Winter 2004

General resources

There is no set text for this course because no one has written one yet. But here are some general references that, between them, cover much of the material.

Course-pack: The course-pack contains a copy of the review article: M. E. J. Newman, The structure and function of complex networks. *SIAM Review* **45**, 167–256 (2003).

Other reviews:

- S. H. Strogatz, Exploring complex networks. *Nature* **410**, 268–276 (2001)
- R. Albert and A.-L. Barabási, Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002)
- S. N. Dorogovtsev and J. F. F. Mendes, Evolution of networks. *Advances in Physics* **51**, 1079–1187 (2002)

Books: A list of worthwhile books in the field is given below. Among these, I'd recommend for graph theory either West or Wilson, and for social network analysis probably Scott. The Ahuja book is excellent if you're interested in the computer programming/algorithms side of things.

- R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, Upper Saddle River, NJ (1993)
- A.-L. Barabási, *Linked: The New Science of Networks*. Perseus, Cambridge, MA (2002)
- B. Bollobás, *Modern Graph Theory*. Springer, New York (1998)
- S. Bornholdt and H. G. Schuster (eds.), *Handbook of Graphs and Networks*. Wiley-VCH, Berlin (2003)
- A. Degenne and M. Forsé, *Introducing Social Networks*. Sage, London (1999)
- S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford University Press, Oxford (2003)
- F. Harary, *Graph Theory*. Perseus, Cambridge, MA (1995)
- J. Scott, *Social Network Analysis: A Handbook*. Sage, London, 2nd edition (2000)
- S. Wasserman and K. Faust, *Social Network Analysis*. Cambridge University Press, Cambridge (1994)
- D. J. Watts, *Small Worlds*. Princeton University Press, Princeton (1999)
- D. J. Watts, *Six Degrees: The Science of a Connected Age*. Norton, New York (2003)
- D. B. West, *Introduction to Graph Theory*. Prentice Hall, Upper Saddle River, NJ (1996)
- R. J. Wilson, *Introduction to Graph Theory*. Addison-Wesley, Reading, MA, 4th edition (1997)

Web sites and software:

- Web site for this course: `http://www-personal.umich.edu/~mejn/courses/2004/cscs535/`
- Physics preprints: `http://aps.arxiv.org/`
- Network software for Windows: `http://vlado.fmf.uni-lj.si/pub/networks/pajek/`
- Network software for Linux/Unix: `http://www.research.att.com/sw/tools/graphviz/`
- Network software for Mathematica: `http://www.combinatorica.com/`

Syllabus

Foundations

Graph theory   Reading: Course-pack, Sections I to III

- Vertices, edges, adjacency matrix
- Directed/undirected graphs, hyperedges
- Bipartite graphs and types of vertices
- Cyclic/acyclic graphs, upper triangular adjacency matrices
- Dual graphs, trees, planar graphs, proof of planarity
- Degree, in-degree, out-degree, mean degree and density, degree distributions
- Paths, powers of the adjacency matrix, geodesics, diameter
- Eulerian/Hamiltonian paths, Königsberg problem
- Components, strongly-connected components, motifs
- Independent paths, connectivity, Menger's theorem
- Weighted networks, max-flow/min-cut theorem, minimum spanning trees, Kruskal's algorithm
- Graph Laplacian, diffusion, spectra
- Random walks, resistor networks

Social network analysis

- Degree centrality
- Closeness, the small-world effect
- Eigenvector centrality, PageRank, Bonacich and Katz centralities, HITS algorithm, co-citation and bibliographic coupling
- Betweenness, flow betweenness, random-walk betweenness
- k-components, k-cliques, k-clans, k-clubs, k-cores, k-plexes, LS-sets
- Transitivity, clustering coefficients, Watts–Strogatz clustering, Burt's redundancy, structural holes, reciprocity
- Valued networks, weak ties, signed networks, structural balance, clusterability
- Structural equivalence, Euclidean distance, Pearson correlations
- Assortative mixing, degree correlations
- Longitudinal network analysis

Empirical studies of networks

Reading: Look again at Sec. II in the course-pack

Social networks

- Early studies, Moreno/sociometry, southern women study, etc.
- Methods of observation: direct observation, affiliation data, interviews and questionnaires, ego-centered networks and snowball sampling
- Direct observation: Zachary's karate club
- Interviews: school friendship networks, the AddHealth study and dating patterns
- Archival data: Padgett's Florentine families, Krebs' terrorist network, email networks, logs, address books, telephone call graphs
- Affiliation networks: boards of directors, CEOs and clubs, collaboration networks, scientists, film actors, ReferralWeb
- Ego-centered networks: Killworth's estimates of personal networks, contact tracing and the Colorado Springs study
- Other methods: Milgram's experiments, the INDEX experiments, Klovdahl's random walk studies, criminal and covert networks, online communities

Technological networks

- Internet, routing, traceroute, BGP, autonomous systems
- The power grid
- The telephone network
- Electronic circuits, software call graphs, dependency networks
- Roads, airlines, railways, cargo and postal networks
- River networks, blood vessels, roots

Information networks

- Citation networks, citation indices, Citeseer, Citebase, patents, law cases
- World wide web, web crawlers, search engines, blogs and wikis
- Peer-to-peer (P2P) networks, filesharing, Napster, Gnutella, KaZaA
- Certificate signing and trust networks
- Recommender networks, Firefly, Ringo, HOMR, Amazon
- Keyword indices
- Word networks, WordNet, thesauruses

Biological networks

- Food webs
- Metabolic networks, protein interaction networks, genetic regulatory networks
- Neural networks

Algorithms

Computer algorithms for the analysis of networks

- Summary of available standard software: Pajek, UCInet, JUNG, Netminer, NetDraw, GraphViz, MAGE, Krackplot, R, Combinatorica, AGD, LEDA, Visone
- Representations of networks: adjacency matrices, edge lists, adjacency lists, hybrid representations
- Trivial algorithms: calculation of degree, clustering coefficient, degree distributions, rank/frequency plots (cumulative histograms)
- Shortest path algorithms: breadth-first search, finding components, depth-first search, closeness centrality, Freeman betweenness, edge betweenness, Dijkstra's algorithm and binary heaps
- Eigenmodes, eigenvector centrality, algebraic connectivity, matrix diagonalization, the multiplication method and the Lanczos algorithm
- Graph partitioning and clustering: spectral partitioning, hierarchical clustering, union/find algorithms, Girvan–Newman algorithm and variants, optimization methods and the Kernighan–Lin algorithm
- Singular value decomposition, latent semantic indexing, collaborative filtering
- Maximum flow, augmenting path algorithm, community-finding algorithms of Flake *et al.* and Wu and Huberman

Models of networks

Mathematical models of networks and their solutions

Reading: course-pack Sections IV to VII. (Don't read them all at once—it will take us some weeks to cover all the material they contain.)

- The Erdős–Rényi random graph: definition of G_{mn} and G_{np}, heuristic derivation of the phase transition

- Generalized random graphs and the configuration model: degree and excess degree distributions; the generating function method, component sizes, the phase transition, the giant component; directed graphs, bipartite graphs, and graphs with degree correlations
- Exponential random graphs: graph Hamiltonian, partition function, free energy, expectation values; simple examples; Monte Carlo simulation, parameter estimation; the 2-star model, the model of Holland and Leinhardt
- The small-world model: scaling theory, mean-field theory
- Growing graphs: cumulative advantage, preferential attachment, Price's model, and the rate-equation method; the Barabási–Albert model; redirection models and triadic closure; nonlinear kernels, fitness models, and other variations on the theme; empirical evidence for preferential attachment and triadic closure; vertex copying models
- Branching networks and the models of Banavar *et al.* and West *et al.*

Processes taking place on networks

Reading: course-pack Sec. VIII

- Graph resilience, site percolation theory
- Epidemic processes on networks: SIR, SIS, SIRS models; compartmental models and moment closure, percolation theory, mean-field theory; correlated networks, vaccination, acquaintance vaccination, computer viruses
- Network navigation, search on networks
- Constraint satisfaction, coloring, satisfiability, vertex cover, Davis–Putnam algorithm, analysis of simple heuristics, survey propagation
- Voter models, coupled oscillators, cascading failures, game theory, Landau theory, random Boolean networks, neural net models, dynamical stability (May's theorem)