Complex networks 101

Simone Severini
University College London

CS reading group, Sept 2011

Plan

- References
- Complex networks
- Graphs
- Real networks
- Models
- Fault tolerance
- Dynamics
- Perspectives

References

- The structure and function of complex networks, Newman, cond-mat/0303516.
 Complex networks: structure and dynamics, Boccaletti *et al., Phys. Rep.,* 424 (2006).
- Characterization of complex networks: a survery of measurements, da F. Costa, *et al.* cond-mat/0505185
- Dynamical processes on complex networks, Barrat, Barthélemy, Vespignani, CUP, 2008.
- Complex Graphs and Networks, Chung, Lu, AMS, 2006.
- A course: www.cs.cornell.edu/courses/cs6850/2011sp/ Kleinberg.
- Networks, Crowds, and Markets: Reasoning About a Highly Connected World, Easley, Kleinberg, CUP, 2010.
- Software: GUI: Pajek, Yed, Ucinet; Command line programs: Graphviz, R, Statnet; Libraries: Leda, Igraph.

Complex networks

• Definition: Complex networks are “complex” graphs with “many” vertices (2D lattices?).
• Tools: statistical mechanics (microscopic processes give macroscopic properties), graphs theory, linear algebra, *etc.*
• Fact: a good abstraction for studying systems composed by many parties; often a *necessary approximation*.
• Examples: social, organizational, biological, technological (possibly *ad hoc*), abstract models, *etc.*
• Classes: static, dynamics (*e.g.* growing).
• Questions: (1) determine the structure; (2) metrology; (3) construction; (4) dynamics.
• Folklore: Six degrees of separation, Kevin Bacon number, *etc.*
• Visualization: graph drawing, dynamics, abstract representations.
• Beyond topology: agent based models, dynamical systems.

Graphs

- Types: simple graphs, directed graphs, weighted graphs.
- Local properties: degree, average degree, density, degree distribution (# vertices of a given degree / size)
- Trials: paths, circuits, cycles, Euler circuits and Hamilton cycles.
- Distances: distance, average distance, diameter, betweenness centrality (Floyd–Warshall algorithm).
- Substructures: subgraphs, minors, cliques, clustering coefficient, average clustering coefficient, motifs (more frequent subgraphs w.r.t. random), spanning trees.
- Representations: adjacency matrix, Laplacian(s).
- Spectra of graphs: eigenvalue gap, expansion.
- Random walks: limiting probability distribution, mixing time.

I like: R. Diestel, Graph theory, Springer, 2010.

Real networks

• Universality: (1) degree distribution is a power-law (*i.e.* "scale free" from the log-log plot); (2) betweenness centrality distribution is decreasing; (3) average distance scales logarithmically with the size; (4) clustering coefficient ~ to average degree.
• Mixing patterns: (1) assortative – social networks (people tend to link with people like them); (2) disassortative – biological networks (high degree vertices tend to link to small degree vertices).
• Community structure: (1) modularity ("natural phenomenon" on large graphs?); (2) strong/weak communities.

I like: M. Mitzenmacher A Brief History of Generative Models for Power Law and Lognormal Distributions. Internet Mathematics, vol 1, No. 2, pp. 226-251, 2004.

Models

• Erdös-Renyi random graph (1959): flip a *p*-biased coin for each pair of vertices; Poisson degree distribution $P(k)=e^{-\langle k\rangle}\langle k\rangle^{k}/k!$,
"a.a.s. theorems", properties depend on *p* (phase transitions, giant component, *etc.*); clustering coefficient inversely ~ to the size (clustering coefficient of real networks is const.).

• Watts-Strogatz (1998): rewire the edges of a lattice with probability *p*; transition lattice-random graph; const. clustering coefficient; Poisson degree distribution (no power-law).

• Barabasi-Albert (2000): (1) growth; (2) preferential attachment (the rich gets richer); power-law $P(k)=k^{-\gamma}$.

• Other models: duplication, ageing, deterministic models, blind attachment, *etc.*

• Remarks: no universal exponent, attachment is system dependent, networks evolve in time!

Fault tolerance

- Robustness: complex systems tend to maintain their basic functions even under errors and attack (perturbations); many pre-networks theory studies in interconnection networks (S. Hailes).
- Models and measures: edge/vertex deletion; measures of resilience (variation of average distance, size of the giant component, percolation); flows.
- Facts: ER graph: giant component is destroyed if a critical fraction of nodes is deleted; Scale-free: more error tolerant, more vulnerable to attacks; different types of nodes. General: higher degree nodes; subgraphs.
- Issues: cascading effects; need for system dependant measures.

Dynamics

- Principle: network topology together with a description of network dynamics (add/remove elements or change their state, automata theory).
- Dynamical processes: discrete, continuous, deterministic, stochastic; attractor – states repeated in sequence (single, cycle).
- Boolean networks (Kauffman networks): random/fixed topology; states: frozen, chaotic, self-organized criticality (boundary between order and disorder).
- Epidemics: modelling (agent based models, contact models, meta-population); predictability (susceptible, infected, recovered); targeted immunization & policy (failure of travel restrictions).
- Hamiltonian dynamics: spatial networks.

Perspectives

- Control theory: Kalman-type conditions, controllable graphs.
- Approximation: spanners and sparsifiers.
- Algorithms: search, hierarchies.
- Information theory: information measures and dynamics (entropies); network coding, networks as channels; compression.
- Games on networks: evolutionary dynamics, cooperation, *etc.*
- Inference: Bayesian networks.
- Beyond networks: agent based models.
- Towards our "Graphs Clinic": (1) data; (2) determine a network (or matrix); (3) routine analysis; (4) new tools (?); (5) verification!

I like: A. El Gamal, Y.-H. Kim, Lecture Notes on Network Information Theory, arXiv:1001.3404v4 [cs.IT]
D. Spielman, S.-H. Teng, Spectral sparsication of graphs, arxiv.org/abs/0808.4134.