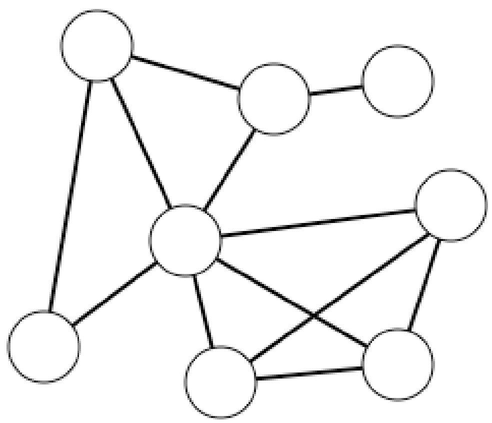
Exponential random graphs and dynamic graph algorithms

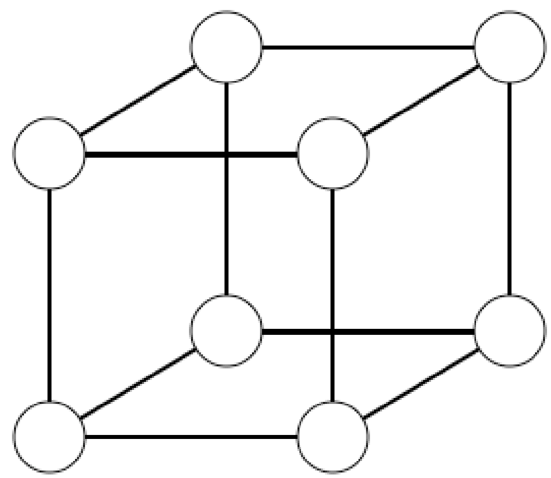
David Eppstein
Comp. Sci. Dept., UC Irvine

What are we trying to do?

- Probabilistic reasoning on social networks
- Appropriately model the different likelihoods of finding different types of network



Likely



Unlikely

Exponential Random Graphs

- Family of graphs with fixed vertex set
- Probability of a graph is proportional to exp(sum of weights of features)
- Different choices of features give simpler or more powerful models

ERG models can be simple…

- Easily subsumes many standard random graph models
- E.g. G(n,p):
  - Edges are independent w/probability p
  - Feature = edge
  - Weight = log(p) - log(1-p)

…but ERG models can also be very powerful

- Powerful enough to represent any distribution over n-vertex graphs
  - Feature = isomorphism with one graph
  - Weight = log(probability of that graph)
- More power requires a more complex set of features

Computational tasks for reasoning with ERGs

- Compute normalizing factor (partition function) for graph probabilities
- Generate random graphs from the model
- Use the model as a prior for max-likelihood data fitting, or modify the feature weights to fit the data

Monte Carlo methods for computing with ERGs

- Start with an arbitrary graph
- Repeatedly propose a small change (e.g., insert or delete a single edge)
- Compute log-likelihood of the modified graph and use it to accept or reject the proposed change

The Algorithmic Lens

- Social scientists and statisticians determine the sorts of models that best describe their data
- Algorithms researchers (e.g. me) figure out how to make the model run quickly
- Faster algorithms lead to the ability to use more accurate models

Algorithmic rephrasing of the computational task

- Maintain a dynamic graph subject to edge insertions and deletions
- As the graph changes, keep track of its computational properties efficiently (faster than recomputing them from scratch)
- The properties we track should be the ones needed for ERG feature vectors

A brief survey of dynamic graph algorithms

- Sparsification (E., Galil, Italiano, Nissenzweig, JACM '92):
  - Replace dense graphs by tree of sparse subgraphs
  - Applies to many problems including maintaining connected components
  - Replaces #edges by #vertices in running times of update algorithms

A brief survey of dynamic graph algorithms

- Fast dynamic connectivity (Holm, de Lichtenberg, Thorup, JACM 2001):
  - Maintain connected components, number of connected components, or a spanning tree (so can use #components as ERG feature)
  - Update time O(log n log log n)
  - Complicated, of interest to search for more easily implemented variants

A brief survey of dynamic graph algorithms

- Distance and reachability in graphs
  - Of likely use in ERGs (e.g. to model small-world properties of these graphs)
  - Some dynamic graph algorithms are known but more theoretical than practical

A brief survey of dynamic graph algorithms

- Graphs in the plane and on surfaces
  - E. et al, J. Algorithms 1992
  - E. et al, J. Comp. Sys. Sci. 1996
  - E., SODA 2002
- Of possible interest for integrating social networks with geographic data

Not-yet-dynamized graph algorithms useful for ERGs

- Low-degree orientations of sparse graphs (Chrobak, E., Theor. Comp. Sci. 1991)
  - Assign directions to the edges of the graph so that each vertex has O(1) outgoing edges
  - Enable fast search for small subgraphs (e.g. list all cliques in linear time)
  - May be found in linear time

Not-yet-dynamized graph algorithms useful for ERGs

- Finding all maximal complete bipartite subgraphs in a sparse graph (E., IPL 1994)
  - Allows concise representation of all four-vertex cycles (quadratically many cycles may be represented in linear space and time)
  - Based on low-degree orientation

Not-yet-dynamized graph algorithms useful for ERGs

- Subgraph isomorphism: finding all copies of some small pattern graph in a larger graph (E., J. Graph Th. 1993 and J. Graph Algorithms 1999)
  - Commonly used as ERG features
  - Known fast algorithms rely on special graph properties e.g. planarity

Not-yet-dynamized graph algorithms useful for ERGs

- Centrality: measuring the importance of a node in a social network
- Fast approximation of closeness centrality based on small world hypothesis (E., Wang, J. Graph Algorithms 2004)
- Other centralities e.g. betweennness

Conclusions

- ERG are important model for social nets
- ERG computation naturally involves dynamic graphs
- Many existing dynamic graph algorithms known, not fully adapted to ERG problems
- Much opportunity for further study of dynamic graph algorithms in ERG setting