Understanding ERGM models:
Why the first models tried did not work

Mark S. Handcock

Departments of Statistics
and
Center for Statistics and the Social Sciences

University of Washington

email: `handcock@stat.washington.edu`

Joint work with

Martina Morris and the UW Networks Project

Sunbelt HalfTime: February 17, 2005

Building a simple model for a social network

- Start with modeling the density of edges
  - X_{ij} are independent and equally likely
    with log-odds $\theta = \text{logit}[P(X_{ij} = 1)]$
  - Homogeneous Bernoulli graph (Renyi-Erdos model)

$$P(X = x) = \frac{e^{\theta E(x)}}{c(\theta)} \qquad\qquad x \in \mathcal{X}$$

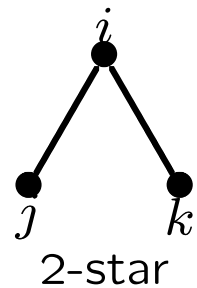
where $E(x) = \sum_{i,j} x_{ij}$, is the number of edges
$c(\theta) = [1 + \exp(\theta)]^g$

Adding the dependence characteristic of social networks

- Edges often occur in clusters - model the *clustering*:

  $S(x)$ is the number of 2−stars

$$S(x) = \sum_{i<j<k} x_{ij} x_{ik}$$



2-star configuration for undirected graphs

- Clusters of edges are often *transitive*:

  $T(x)$ is the proportion of triangles amongst triads

$$T(x) = \frac{1}{\binom{g}{3}} \sum_{\{i,j,k\} \in \binom{g}{3}} x_{ij} x_{ik} x_{jk}$$

  A closely related quantity is the *percent of complete triangles* or *mean clustering coefficient*

$$C(x) = \frac{T(x)}{S(x)}$$

- These models are still simple in the sense that edges in X that do not share an actor are conditionally independent given the rest of the network

  $\Rightarrow$ analogous to nearest neighbor ideas in spatial statistics

Example: A simple model with transitivity

$g = 50$ actors $\quad N = 1225$ pairs $\quad 10^{369}$ graphs

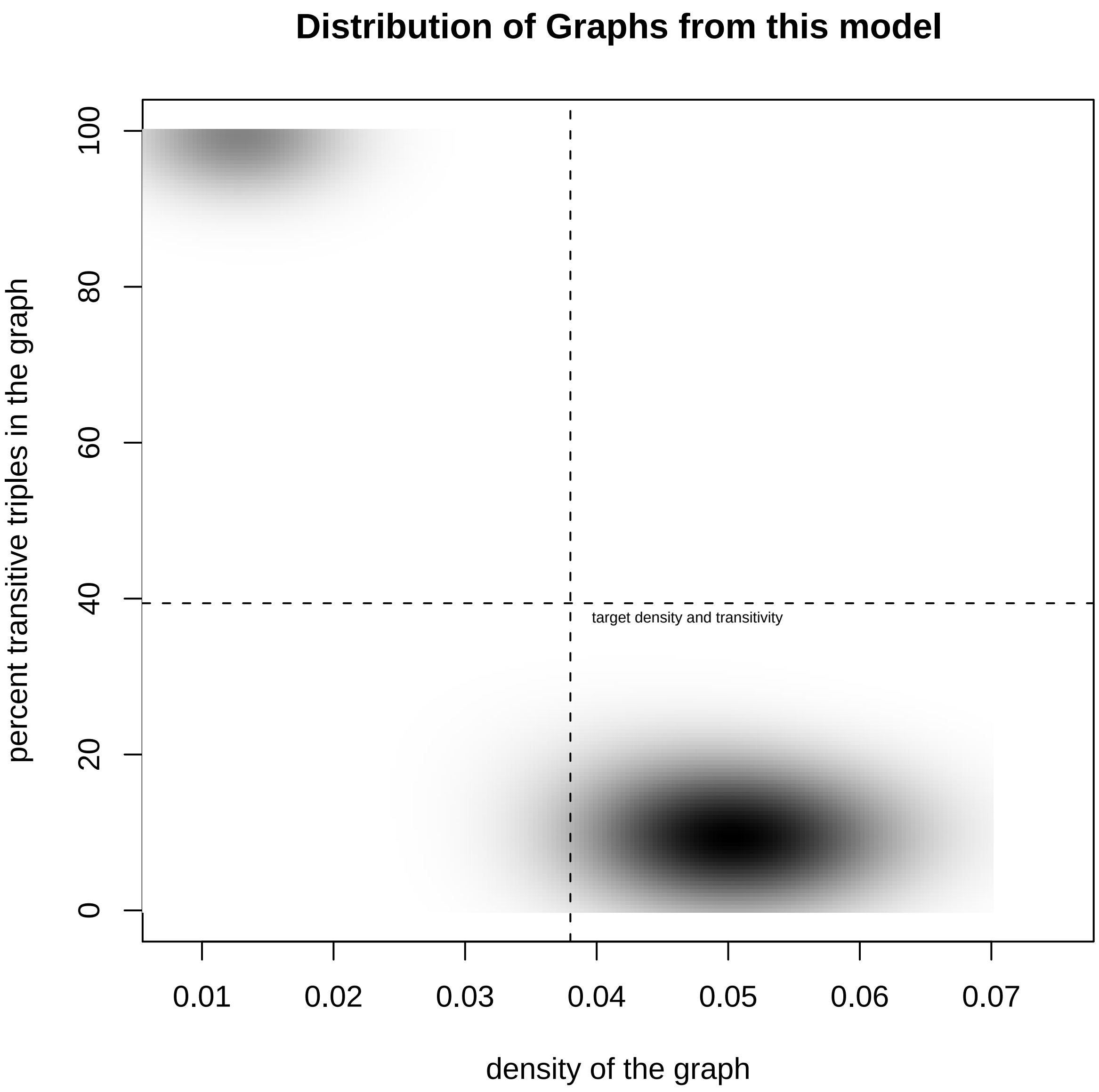
$$P(X = x) = \frac{\exp\{\theta_1 E(x) + \theta_2 C(x)\}}{c(\theta_1, \theta_2)} \qquad x \in \mathcal{X}$$

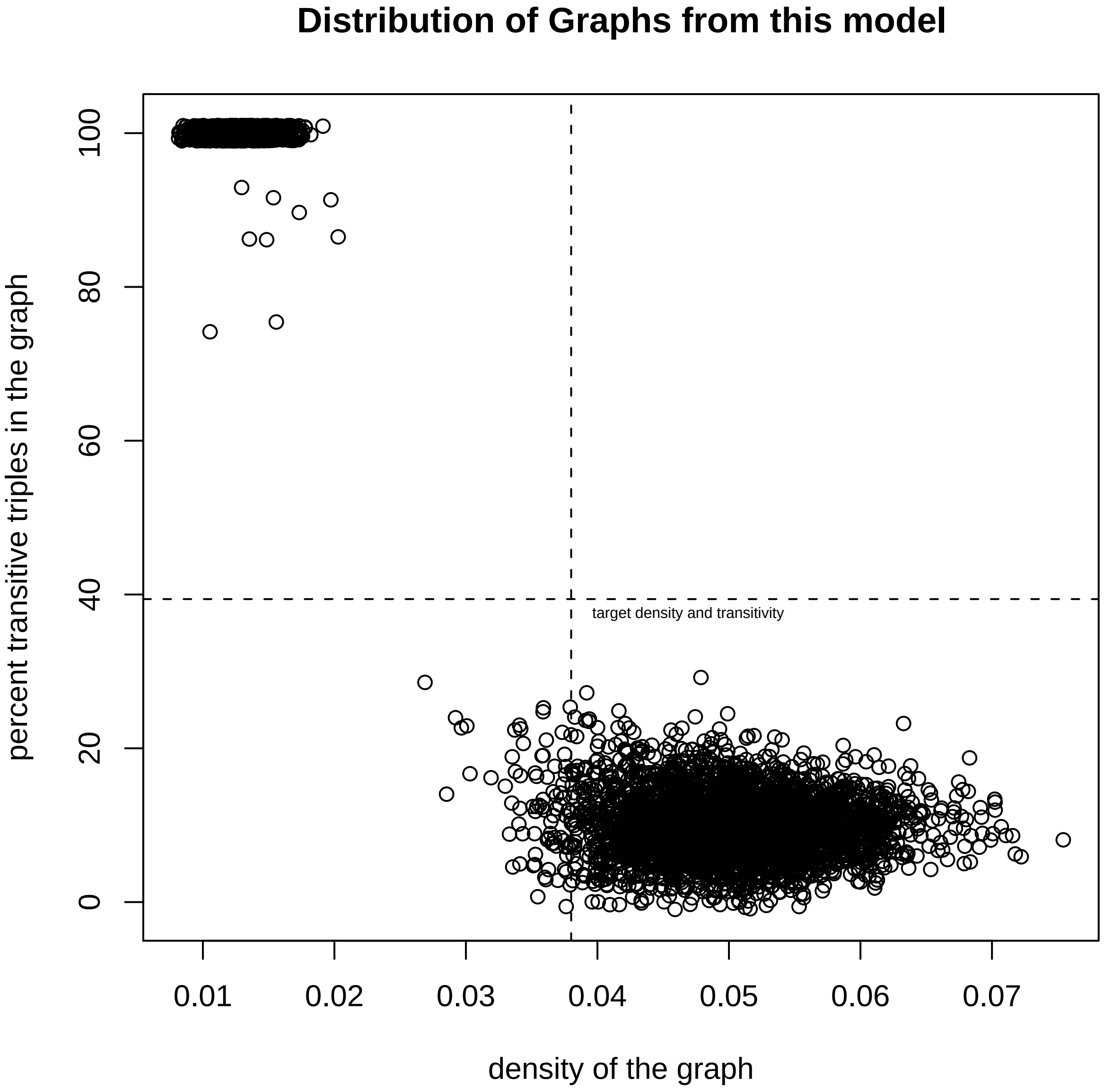
where $E(x)$ is the density of edges $(0 - 1)$

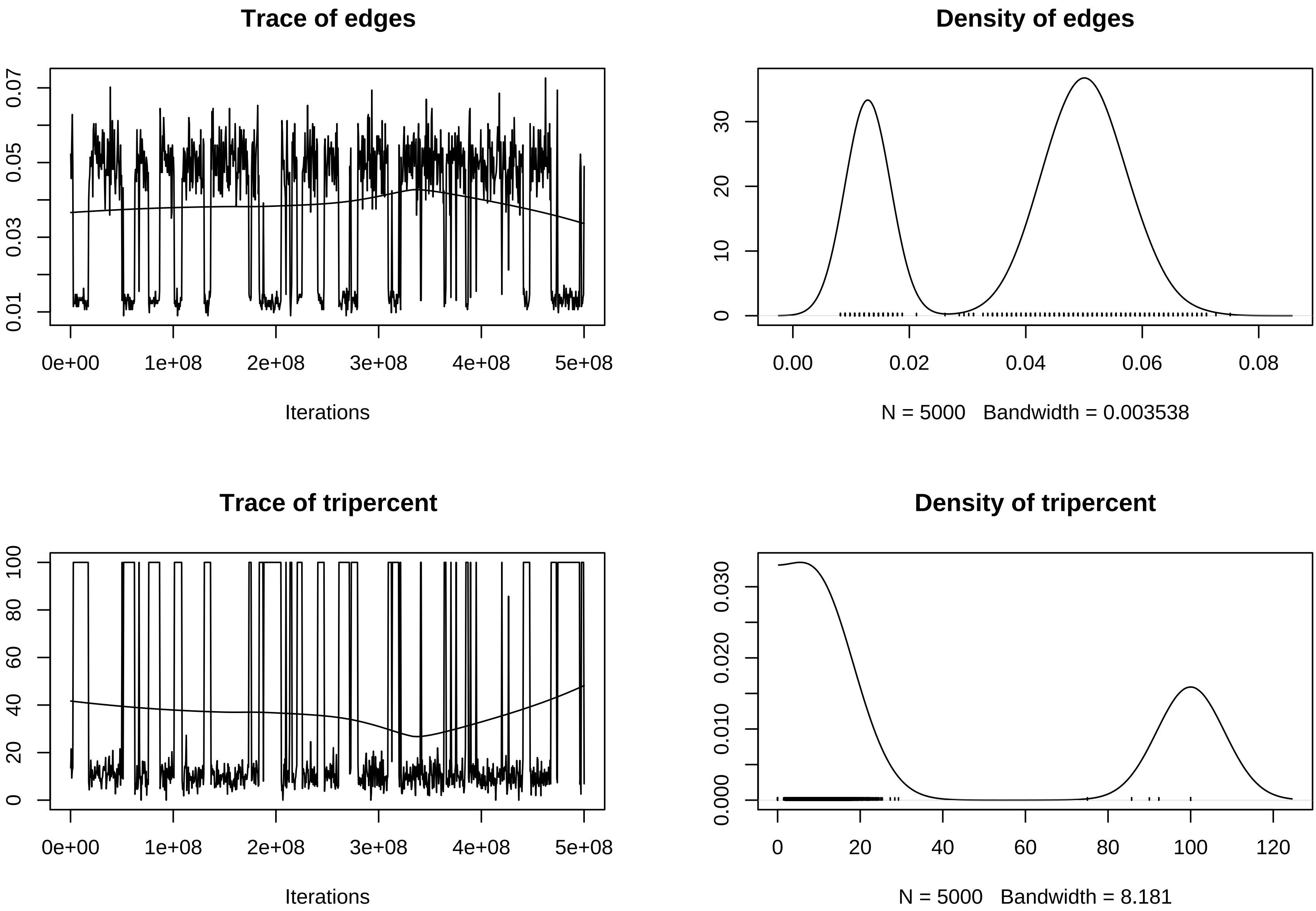
$C(x)$ is the triangle percent $(0 - 100)$

- If we set the density of the graph to have about 50 edges then the expected triangle percent is 3.8%
- Suppose we set the triangle percent large to reflect transitivity in the graph: 38%
- By construction, on average, graphs from this model have average density 4% and average triangle percent 38%

- If the model is a good representation of transitivity and density we expect the graphs drawn from the model to be close to these values!


Distribution of Graphs from this model
percent transitive triples in the graph
density of the graph
target density and transitivity
0
20
40
60
80
100
0.01
0.02
0.03
0.04
0.05
0.06
0.07


Distribution of Graphs from this model
percent transitive triples in the graph
density of the graph
target density and transitivity
0
20
40
60
80
100
0.01
0.02
0.03
0.04
0.05
0.06
0.07



Trace of edges
Iterations
Density of edges
N = 5000 Bandwidth = 0.003538
Trace of tripercent
Iterations
Density of tripercent
N = 5000 Bandwidth = 8.181

How can we tell if a model class is useful?

Many aspects:

- Is the model-class itself able to represent a range of realistic networks?
  - *model degeneracy:* small range of graphs covered as the parameters vary (Handcock 2003)
- What are the properties of different methods of estimation?
  - e.g, MLE, psuedolikelihood, Bayesian framework
  - *computational failure*: estimates do not exist for certain observable graphs (Snijders 2002; Handcock 2002)
- Can we assess the goodness-of-fit of models?
  - appropriate measures and tests (Hunter, Goodreau, Handcock 2005)

Measuring Model Degeneracy

idea: A random graph model is *near degenerate* if the model places almost all its probability mass on a small number of graph configurations in $\mathcal{X}$.
e.g. empty graph, full graph, an individual graph, no $2-$stars, mono-degree graphs

- Example: The *2-star* model

$$P(X = x) = \frac{\exp\{\theta_1 E(x) + \theta_2 S(x)\}}{c(\theta_1, \theta_2)} \qquad x \in \mathcal{X}$$

is near-degenerate for most values of $\theta_2 > 0$

Figure 4: Regions of the parameter space of μ

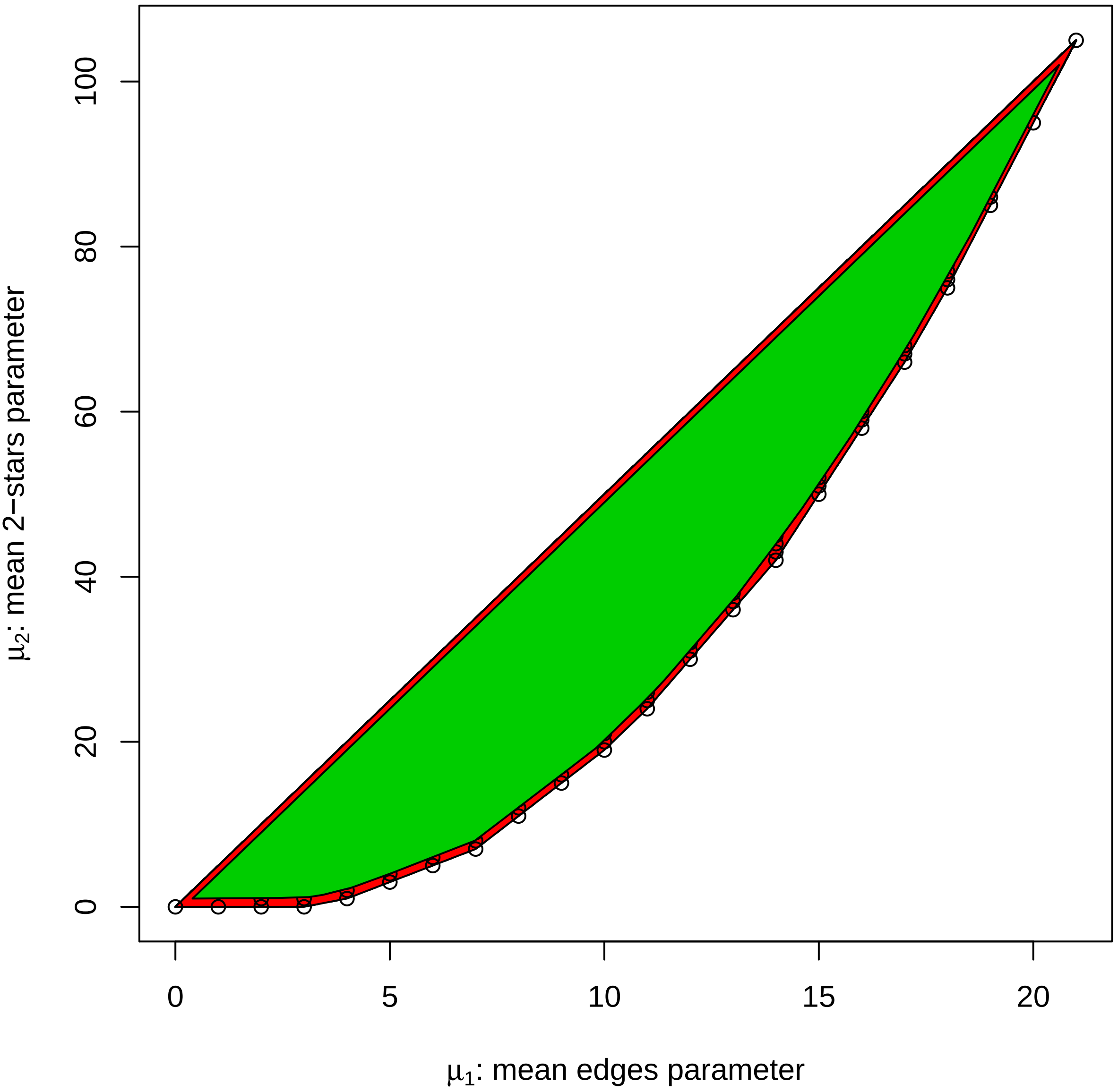
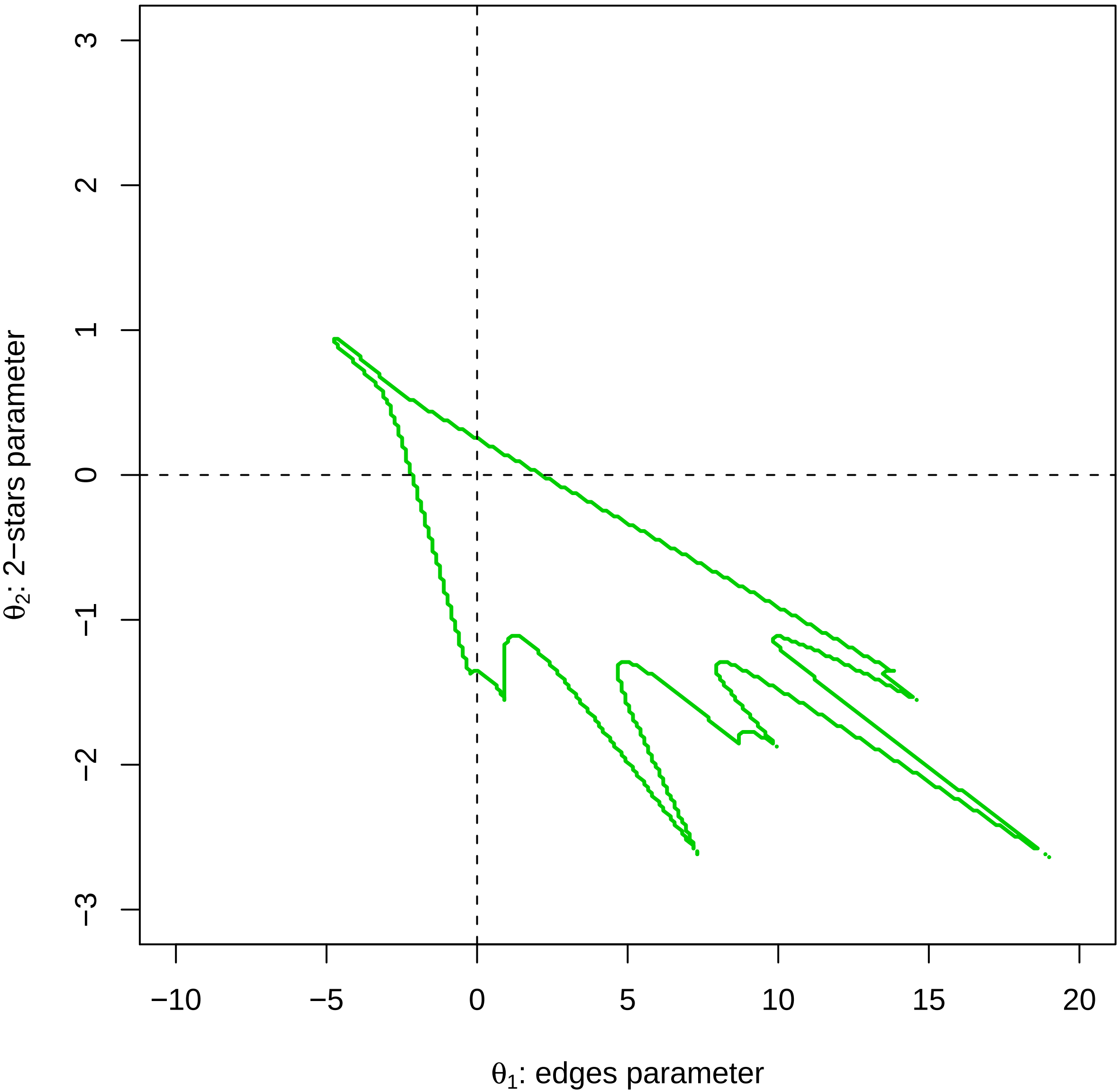


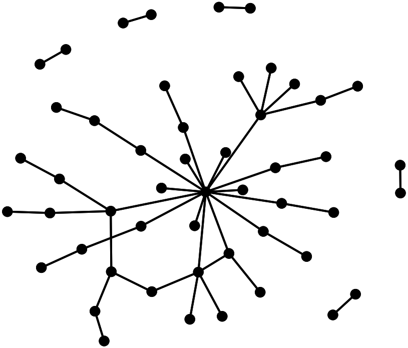
Figure 5: Regions of the parameter space of θ



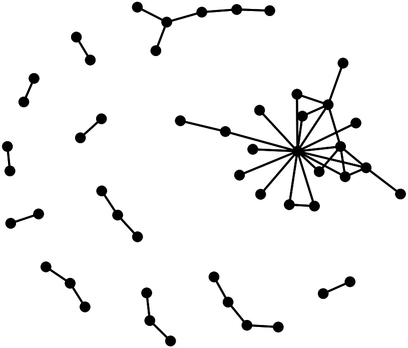
Illustrations of simple models that work

- village-level structure
  - $g = 50$
  - mean clustering coefficient = 15%
  - degree distribution: Yule with scaling exponent 3.
- larger-level structure
  - $g = 1000$
  - mean clustering coefficient = 15%
  - degree distribution: Yule with scaling exponent 3.
- Attribute mixing
  - Two-sex populations
  - mean clustering coefficient = 15%
  - degree distribution: Yule with scaling exponent 3.

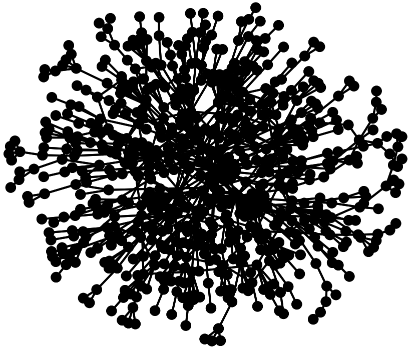
Yule with zero clustering coefficient conditional on degree



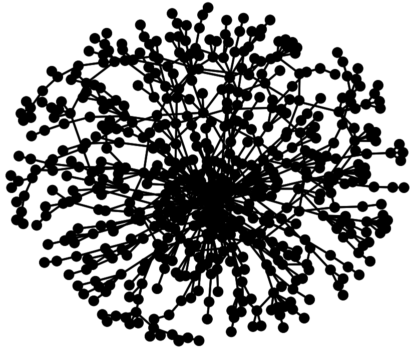
Yule with clustering coefficient 15%

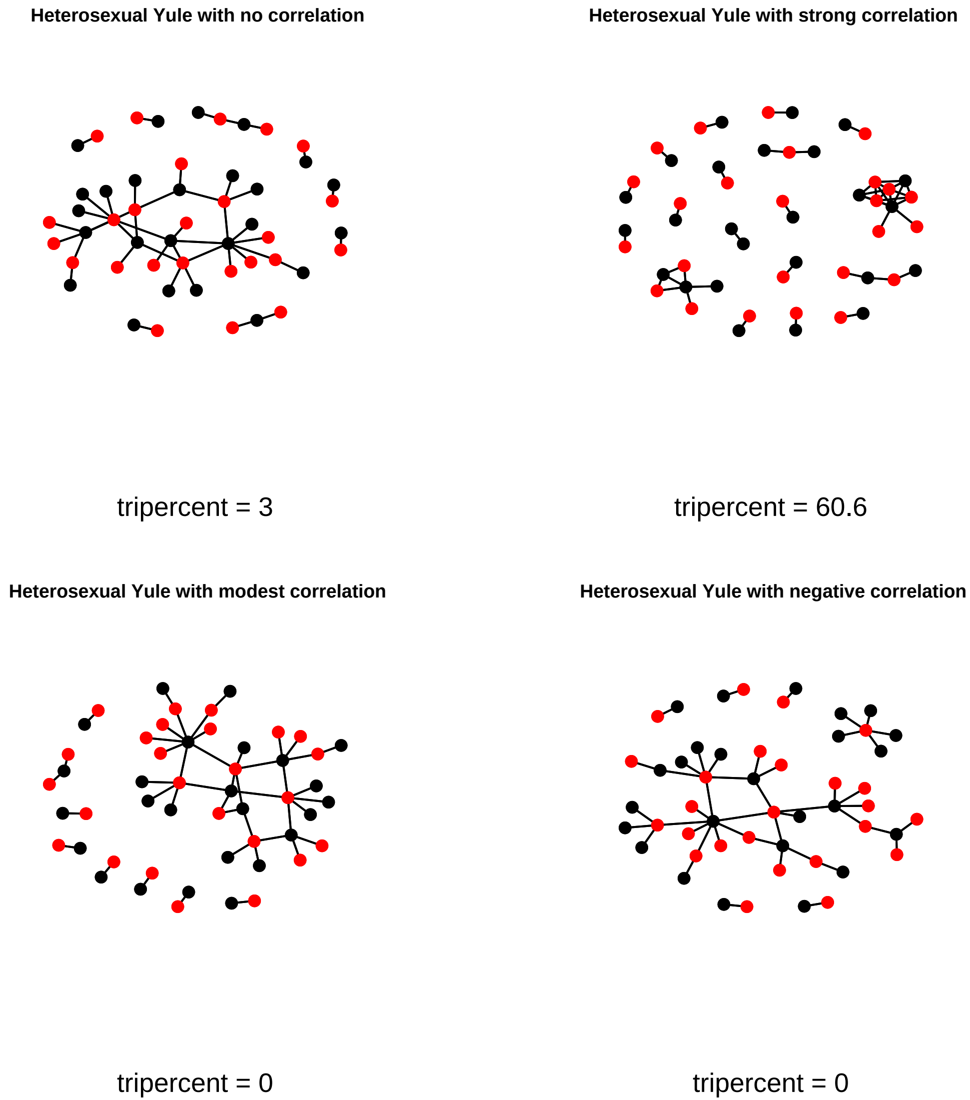


Yule with zero clustering coefficient conditional on degree



Yule with clustering coefficient 15%




Heterosexual Yule with no correlation
tripercent = 3
Heterosexual Yule with strong correlation
tripercent = 60.6
Heterosexual Yule with modest correlation
tripercent = 0
Heterosexual Yule with negative correlation
tripercent = 0

Conclusions and Challenges

- The Choice of Models depends on the objectives
- some seemingly simple models are not so.
- simple models are being used to capture clustering and other structural properties
- The inclusion of attributes is very important
  - actor attributes
  - dyad attributes e.g. homophily, race, location
- latent class and trait models are important
  - an underlying latent "social space" of actors
    - $\Rightarrow$ Hoff, Raftery and Handcock (2002)
    - $\Rightarrow$ Hoff (2003, 2004 ,...)
  - latent class models are very promising
    - $\Rightarrow$ Nowicki and Snijders (2001)
  - latent class and trait models

What has been done in the last five years?

- Generating random networks with (constrained) structure
  - Snijders (1991, 2002); Handcock (2000; 2005)
- Network sampling design and analysis
  - Frank (1972); Thompson and Frank (2000)
- The likelihood framework can naturally adapt to sampling and missing data
  - Handcock (2000; 2005)

Some Key references:

- Leinhardt (1977)
  Book: "Social Networks: A Developing Paradigm"
- Holland and Leinhardt (1981) JASA
- Frank and Strauss (1986) JASA
- Wasserman and Faust (1994)
  Book: "Social Network Analysis: Methods and Applications"
- Morris (1997)
- Doreian and Stokman (1997)
  Book: "Evolution of Social Networks"
- Besag (1974-2000) JRSSB, etc
- Newman (2003) SIAM Review
- Morris (2003)
  Book: "Network Epidmiology"

Models for dynamic social networks

- Continuous-time Markov models
  - Wasserman (1977), Holland and Leinhardt (1977)
  - Leenders (1995)
- Actor Oriented: fusion of rational choice and continuous-time Markov models
  - Snijders (1996), Leenders (1995)
- Models of Network Growth
  - Motivated by Simon (1955)
  - Citation networks: Price (1965; 1976)
  - WWW: Albert and Barabási (1999)

Visualization of dynamic social networks

- Benefits:
  - Intuitive way to display networks (Moreno 1932; 1934)
  - Helps people see a map of social space
  - A concise presentation of a great deal of data.
- Costs:
  - Lack of standards for how to display can create misleading images
  - Displays of large networks tend to reveal only the grossest properties of the network
- Linton Freeman has contributed greatly to this
  - For history and development of the field, see

`http://eclectic.ss.uci.edu/~lin/gallery.html`