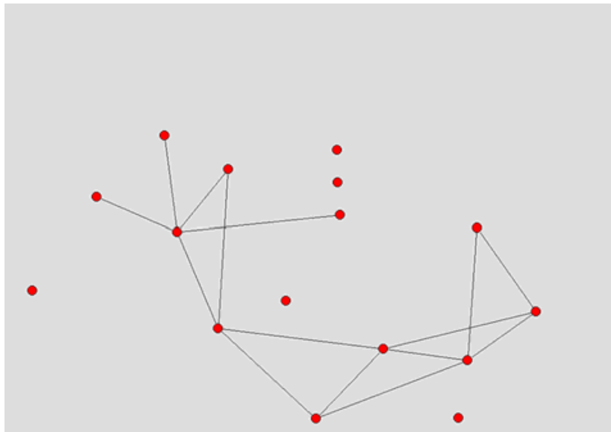
Example:
Florentine families business network
(Padgett & Ansell, 1993)



Markov model estimates:
Florentine families business network

Model containing edges, 2-stars, 3-stars, triangles

Maximum Likelihood estimates

Edge = - 4.27 (1.13)* t = 0.03

2-star = 1.09 (0.65) t = 0.01

3-star = -0.67 (0.41) t = – 0.02

Triangle= 1.32 (0.65)* t = – 0.02

Edge/Triangle model:

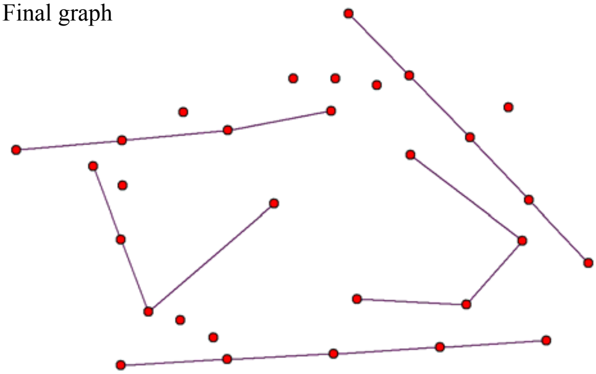
θ = - 3; τ = 0.25

Start from empty graph

Burn-in 50,000

200,000 simulations

Final graph



Edge/Triangle model:

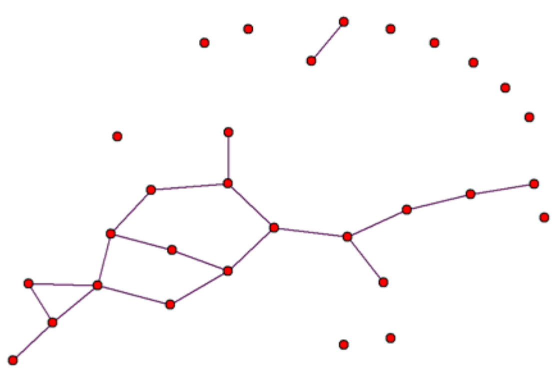
θ = - 3; τ = **0.75**

Start from empty graph

Burn-in 50,000

200,000 simulations

Final graph



Edge/Triangle model:

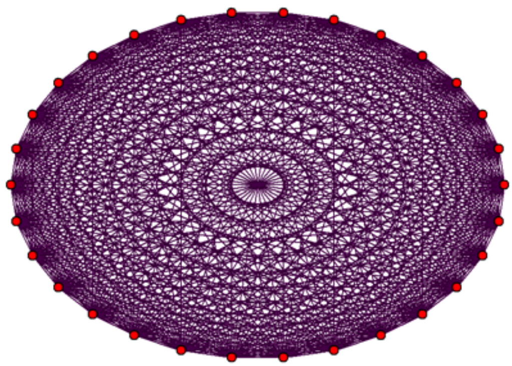
$\theta = - 3; \tau = 0.75$

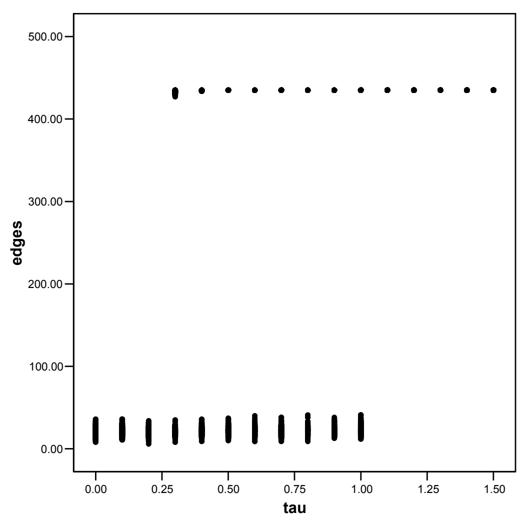
Start from **complete** graph

Burn-in 50,000

200,000 simulations

Final graph





Model degeneracy

For certain parameter values, a model may imply that only one or two graphs with non-zero probability.

Often such graphs are the empty or full graph

(or a graph of complete disconnected components).

These *near degenerate* models cannot be estimated.

Markov models are often degenerate when clustering is high

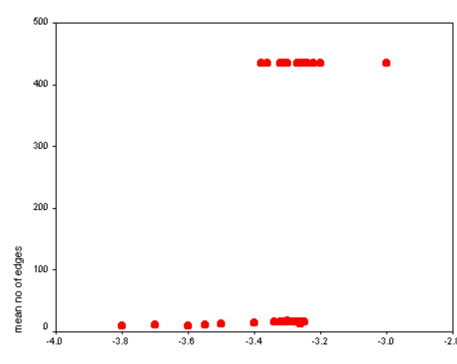
Model degeneracy

Model with only an edge and triangle parameter

$$\Pr(\boldsymbol{X} = \boldsymbol{x}) = (1/c)\exp\{\theta L + \tau T\}$$

Set $\tau = 1.0$

30 node graph



Model degeneracy

Model with edge, triangle parameter and star parameters

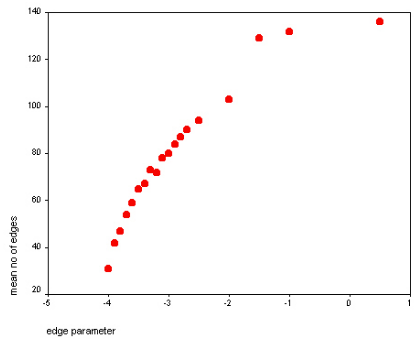
$$\Pr(\boldsymbol{X} = \boldsymbol{x}) = (1/c)\ \exp\{\theta L + \sigma_2 S_2 + \sigma_3 S_3 + \tau T\}$$

Set $\tau = 1.0$

$\sigma_2 = 0.5$

$\sigma_3 = -0.2$

30 node graph



Model degeneracy

For Markov random graphs, the bottom line is:

- Edge/triangle models are NOT going to fit observed data:
  - They imply graphs that are either close to empty or close to complete (*degenerate)*
- Models with edges, triangles, 2-star and 3-star parameters will do better:
  - They may fit data if there is a negative 3-star parameter estimate
  - But they are also likely to fail when clustering is high.

Model degeneracy

Why do Markov models fail when clustering is high in observed networks?

An attempt at an intuitive explanation:

> Markov models suppose that triangles are spread rather evenly throughout the graph (homogeneous models).
>
> They have trouble when triangles tend to clump into *cohesive subsets of nodes.* (The dense part of the graph suggests a very strong triangle effect – but applying this effect to the not-so-dense part tends to “fill the graph up”.)

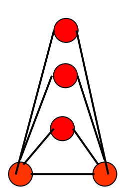
One solution

Go beyond Markov dependence assumptions:

Higher order dependence assumptions producing, most importantly, *alternating k-triangle* statistics

(plus some other higher order statistics.)

3-triangle (T_1)
(T_3)



Snijders, Pattison, Robins & Handcock, 2005;

Robins, Snijders, Wang, Handcock & Pattison, 2005.

©Department of Psychology, University of Melbourne